uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Marina Sokolova
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Computer Science)
GRADE / DEGREE

School of Information Technology and Engineering
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Learning from Communication Data: Language in Electronic Business Negotiations

TITRE DE LA THÈSE / TITLE OF THESIS

Stan Szpakowicz
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Jean-Pierre Corriveau

Diana Inkpen

Stan Matwin

Jian-Yun Nie

Gary W. Slater
LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES /
DEAN OF THE FACULTY OF GRADUATE AND POSTDOCORAL STUDIES

Learning from Communication Data: Language in Electronic Business Negotiations

by

Marina Sokolova

Thesis submitted to the

Faculty of Graduate and Postdoctoral Studies

In partial fulfillment of the requirements

For the degree of Doctor of Philosophy in Computer Science

Ottawa-Carleton Institute for Computer Science

School of Information Technology and Engineering

University of Ottawa

Ottawa, Ontario, Canada

Your file Votre référence
ISBN: 978-0-494-15045-0
Our file Notre référence
ISBN: 978-0-494-15045-0

Геннадию, Ольге и Михаилу Соколовым.

Памяти Маргариты и Леонида Сафоновых
и Марии Саркисянц.

For my husband, Gene, daughter, Olga, and son, Mikhail.

With loving memory of my parents, Margarita and Leonid, and grandmother, Maria.

ABSTRACT

When people communicate, language is one of the means of reaching the goal of communication. Negotiations by electronic means is an example of communication where language is the principal deal-making tool. Negotiators use language to persuade, threaten and query, aiming to get the largest piece of the pie, to reach a compromise or to find prospective partners. Here is a sample from electronic negotiations, with the original spelling, punctuation and capitalization:

Seller *Dear BuTTerFLy Thanks for your offer. I see there are still some things that have to be thought about. We both come along with payment upon delivery.I could imagine a price of $3.98 and delivery 45 days, but unfortunately with the returns i cant make you any other offers. I hope you quite like this offer. Im sure an agreement will be found. Im looking forward to your respond, daisy.*

Buyer *To my dearest friend daisy... Thank you for your quick respond, I quite like your second offer. However I'll be more than happy if the price goes down to 3.71$ and the delivery would be within 30 days (about the payment and the return I don't have any problems with them). I'll really appreciate it if you accept the offer I just made, but if you don't, I'm sure somehow we'll come up with an agreement. yours faithfully BuTTerFLy!!!!!*

We apply statistical modelling and build a semantic lexicon to find the characteristics of e-negotiation data which make it unique. We find language patterns that signal of negotiator roles and success or failure of negotiations.

Research in human communication shows that it is very difficult to find the characteristics of unsuccessful activities and communication corresponding to them. The interesting and promising result of this dissertation comes in the form of identifying two sets of features that characterize successful and unsuccessful communication respectively. We use these sets to represent negotiations and then classify the negotiation outcomes. The results show the advantage of the proposed feature selection approach compared with the popular statistical selection.

We apply our research to the largest available collection of electronic negotiations and, when appropriate, to data of face-to-face negotiations. In the dissertation we employ methods developed for Corpus Linguistics, Natural Language Processing and Machine Learning. We investigate the ability of the methods to model and classify the data. Throughout the dissertation we examine hypotheses on language, learning and the process of electronic negotiations.

Acknowledgements

My most grateful thanks go to ***Stan Szpakowicz*** who has supervised my dissertation research. I have benefited from his expertise in Natural Language Processing and Artificial Intelligence fields and willingness to share this expertise with students. During our regular meetings, numerous research discussions and joint work on several conference and journal papers Professor Szpakowicz has taught me how excellent a supervisor has to be.

My special thanks go to ***Mario Marchand***. Professor Marchand was my first supervisor in Canada and lead my Machine Learning studies. The Decision List Machine, showing the high accuracy performance in this dissertation, has resulted from our joint work. Committee, ***Jean-Pierre Corriveau***, ***Diana Inkpen***, ***Stan Matwin***, and ***Jian-Yun Nie***, the external examiner, gracefully agreed to review this work. I extend my thanks to them for the inspiring and helpful comments.

Many parts of this research benefited from collaboration with ***Mohak Shah***, my colleague and office-mate. I thank Mohak for our intensive and productive work, support of my ideas and his own constructive proposals.

From negotiation roles to Computational Linguistics to the use of LaTeX– this is a non-exhaustive scope of our dialogues with ***Vivi Nastase***, my collaborator and friend. *Mulţumiri tu foarte mult!*

My husband, ***Gene Sokolov***, our daughter, ***Olga Sokolova***, and son, ***Mikhail Sokolov***, have helped to balance Ph.D. studies with life outside the university. One of the wonders of the world is how our family has managed these four years with children and me being full-time students! My wholehearted thanks go to all of us.

I also want to thank ***Nathalie Japkowicz*** – for many interesting discussions on and beyond Machine Learning and co-supervision at the beginning of my Ph.D. studies, ***Irina Bezrukova*** and ***Svetlana Kiritchenko*** – for keeping up the spirit of our *alma mater* – Faculty of Computer Science of Moscow State University, ***Fernanda Caropreso*** – for moral support during my Ph.D. research, ***Daniel Amyot, Alan Williams*** and ***Jianping Yao*** – for their advices and suggestions on different aspects of research life, ***Suzanne St-Michel*** – for the administrative assistance during my years at SITE, ***Michel Racine*** and the system staff – for a reliable and friendly support for experiments, the InterNeg team led by ***Gregory Kersten*** – for providing the *Inspire* and *SimpleNS* data, and ***Jeanne Brett*** and ***Wendi Adair*** – for providing the *Cartoon* data.

This work has been supported by

– the Natural Sciences and Engineering Research Council of Canada through a doctoral scholarship,

– Ontario Government through graduate scholarships,

– University of Ottawa through National Excellence and Excellence scholarships,

– the Canadian Society for Computational Studies of Intelligence through an award and a scholarship,

as well as by a project on Electronic Negotiations funded by the Initiative on the New Economy major grant of the Social Sciences and Humanities Research Council of Canada, Discovery grants funded by the Natural Sciences and Engineering Research Council of Canada, and the Faculty of Graduate and Postdoctoral Studies of University of Ottawa.

Contents

List of Tables

List of Figures

Glossary of Acronyms

AI	Artificial Intelligence
BL	Baseline
BSN	buyers in successful negotiations
BUN	buyers in unsuccessful negotiations
CM	Communication model
CMC	Computer-mediated communication
DLM	Decision List Machine
DM	Data Mining
DS	Decision Stamps
DT	Decision trees
GTK	Good-Turing model with Katz smoothing
i.i.d.	independently identically distributed
IBK	Instance-based classifier with k instances
KN	Kneser-Ney model
LDOCE	Longman Dictionary of Contemporary English
ML	Machine Learning
NB	Naive Bayes
NLP	Natural Language Processing
NSS	Negotiation-support systems
POS	Part-of-speech
SRCC	Spearman's Rank Correlation Coefficient
SSN	sellers in successful negotiations
SUN	sellers in unsuccessful negotiations
SVM	Support Vector Machine
WSJ	Wall Street Journal
XIP	Xerox Incremental Parser

Chapter 1

Introduction

Everything happens for a reason.

Murphy's Law

In this chapter we introduce the topic and goal of our studies, present the motivation leading of the research, state the main contributions and describe the possible applications of this work.

The guide to the remainder of the thesis, accompanied by the dissertation map, and the list of publications conclude the chapter.

1.1 The overview of the problem

Natural Language Processing (NLP) and Machine Learning (ML) methods can study vast amounts of data and extract new knowledge about the data domain. This has turned such methods into necessary and important tools in many research areas. Bioinformatics and economics are among the examples of successful applications of ML and NLP.

Human communication is an exciting field of research that opens vast opportunities in various branches of science. It suggests problems whose solution requires developing new techniques and methods. Computer Science – Text Mining, NLP and ML – studies the data of human communication, gathered in spoken and written dialogue, in email, and so on.

In our work we regard human communication as a complex phenomenon that reflects the underlying activities and depends on them. It is affected by the means which produced it. Its outcome, or the absence of outcome, is the effect of many influencing factors. We consider the outcome of communication a success if and only if the outcome of the underlying activities has been a success. In this work we study how to learn the success, or the lack of it, of communication, and the different roles that the communicating parties play. We have applied our method to human-to-human electronic negotiations. Human-to-human electronic negotiations supply us with the largest available collection of data recorded in human communication. The method is general, that is, it can apply to data gathered in fields other than electronic negotiations.

We have two kinds of results. First, we have proposed new sets of features for the textual data in communication and stated and solved the new tasks of text classification. Second, we have learned the language patterns that contribute to success and failure in electronic negotiations and the language patterns that are used in different negotiation roles.

We discuss the motivation in Section 1.2 and present the goal of our work in detail in Section 1.3.

1.2 Motivation and research context

The number and diversity of communications involving humans rapidly grow. This is due to the globalization of modern life and to the intensively increasing variety of communication means and their use. Behavioural Science, Communication Theory, Linguistics and Computer Science study human communication. Nevertheless, the complex nature of communication leaves numerous opportunities for new approaches and results. In this work we explore the opportunities opened to Natural Language Processing, Machine Learning, and Text Mining by exploiting the person-situation context of human communication.

Communication becomes special, and quite interesting, when it accompanies a business or personal activity, e.g., purchase of goods, writing a paper with a co-author, and so on. The data gathered through such communication provide information about the activities as well as the communication process itself.

A method we propose starts with the analysis and modelling of the data and then employs a set of NLP and ML tools. We apply our method to the data of electronic business negotiations where the activities themselves are well-defined and structured and, thus, it is feasible to trace and interpret them. At the same time, electronic means influence the communication process interfering with the activities that this communication accompanies. Thus we seek two intermingled types of information

– about the activity domain and about human communication through electronic means.

We have to find a set of relevant features in order to gain a new knowledge when apply Machine Learning methods. In this work we propose feature selection based on the knowledge of the negotiation process. We show that features found in this way guarantee a high quality of learning, in the sense of both quantitative (mathematical) and qualitative (negotiation) evaluation of results.

In our work we perform text mining of the data obtained through electronic negotiations. Hearst (2003) has defined the goal of Text Mining as "to discover heretofore unknown information, something that no one yet knows and so could not have yet written down." We obtain, and present here, new results about the language in electronic business negotiations, connecting language with the negotiation outcomes – success or lack of success – and social roles within negotiation – buyers and sellers.

To the best of our knowledge, nobody has applied symbolic and statistical NLP and ML methods to the study of large amounts of the text data exchanged in negotiations – electronic, phone or face-to-face, or any mixture of those. Nor were any such studies performed to analyze text data from electronic business communications. This work is the first to perform both types of analysis.

1.3 The goal and the general plan of this work

Our goal is to establish a method of study of human communication. We achieve this goal by proposing a method which employs NLP and ML techniques. We apply it to the domain of human electronic negotiations: we acquire new knowledge about this domain.

The method we propose has several steps. The first step defines learning tasks and evaluates learning perspectives of data. This step utilizes the multiple viewer approach where electronic negotiations are represented by independent viewers, e.g. log-in records, messages, and so on. The next step builds models based on the data characteristics. In our case it relies on statistical and symbolic NLP techniques. There follows a step of learning the outcomes and roles in communication. It includes feature selection and employs a variety of ML methods. Each step works with the output of the preceding step and contributes to learning.

In the course of our research we have worked on knowledge discovery in a new application field. We have introduced new procedures of feature selection and new tasks for text classification. We have also exploited and then discussed the applicability of existing NLP and ML techniques to a new type of textual data.

We support our hypotheses by extensive experiments. The empirical results are reported throughout the dissertation. We analyze the results with respect to the application field – electronic negotiations. We also investigate new NLP and ML research directions opened by learning from the data supplied by electronic negotiations. The proposed future work focuses on the specific problems of the data.

1.4 Electronic negotiations and their data

We use human-to-human electronic negotiations as a case study. Rojot (1991) defines negotiation as "a process whereby two or more parties attempt to settle what each shall give and take, or perform and receive, in a transaction between them". Negotiations by email or other electronic means are called electronic negotiations (e-negotiations).

The form of e-negotiation data depends on the systems through which they are collected. E-negotiation systems target various domains (e.g., legal or business) and various users (e.g., negotiators or facilitators). They differ in the type of negotiation support and the negotiation parameters, e.g., the number of participants, deadlines, success criteria. We also distinguish communication support and functional support (Holsapple and Whinston, 2000). In this work we concentrate on communication support.

Communication support does not intervene with the negotiation, while functional support influences the negotiation process by assisting users in formulating, evaluating and solving negotiation problems. Passive e-negotiations systems offer no functional support, active systems – some, while pro-active systems offer a high degree of functional support. Figure 1.1 shows such systems. The stronger functional support a system provides, the more structured the data become. Table 1.1 presents the data that different types of e-negotiation systems supply.

For the analysis of human communication the text messages that come from passive systems are the richest: they contain all the information exchanged during negotiation. Messages acquired from a pro-active system can be deprived of the essential information if the user heavily relies on the system's negotiation abilities; texts that make up a (possibly very small) part of highly structured

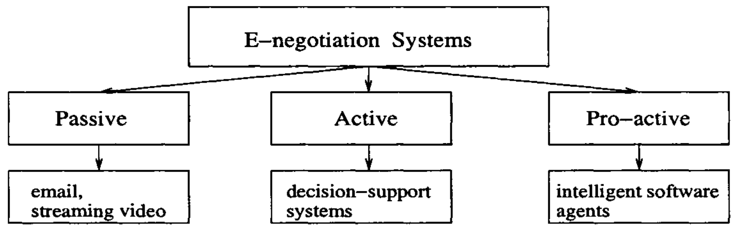

Figure 1.1: E-negotiation systems

Systems	Passive	Active		Pro-active		
Negotiation support	Communication	Communication	Functional	Communication	Functional	
			Process-oriented		Process-oriented	Decision-making
Data	text messages	text messages formal offers	questionnaires, history records, numerical data	text messages formal offers	questionnaires, history records numerical data	defined by the system's protocol

Table 1.1: Services that e-negotiation systems provide, and the corresponding data.

e-negotiation data are outside the scope of our research. Active systems can provide messages that lie between the two extremes. They may contain sufficient information about negotiations because the users negotiate without direct interventions from the system. On the other hand, the users must adapt their interactions to pre-set negotiation parameters, such as deadlines or the number of participants.

1.5 The organization of the remainder of the dissertation

The thesis is organized as follows:

- Chapter 2 surveys NLP and ML methods that we use. It reviews previous research on feature selection. We present our suggestion for employing of NLP and ML methods in the study of communication data. Bibliographical notes conclude the chapter.

- Chapter 3 reviews previous research on negotiations and electronic business negotiations which we consider relevant to our work. We survey AI studies of electronic negotiations. The areas of research that are of special interest for us are listed in the chapter's conclusion.

- Chapter 4 analyzes e-negotiation data. It explains data pre-processing and introduces learning tasks and data classes used in this work. It investigates the characteristics of different types of e-negotiation data. The chapter reports the corpus analysis results. An analysis of the vocabulary used in e-negotiation and face-to-face negotiation texts follows. Conclusions about the general characteristics of the negotiation data and their connections with the negotiation process finish the chapter.

- Chapter 5 reports the results obtained through statistical and semantic data modelling. It analyzes the predictability of the negotiation data and characteristics of the statistical models. The results of tagging and parsing and their connections with the negotiation process are reported. The chapter presents the procedure of lexicon-building from data, which produces a semantic lexicon.

- Chapter 6 introduces language patterns. We report the results of learning. We find patterns supporting a hypothesis about different intensity of successful and unsuccessful negotiations, patterns supporting a hypothesis about different self-assertiveness in successful and unsuccessful negotiations, patterns indicating politeness and analyze their use in successful and unsuccessful negotiations. We find patterns supporting hypotheses about different behaviour of buyers and sellers. A comparison of the patterns concludes the chapter.

- Chapter 7 reports the use of semantic categories in feature selection. We discuss the use of the semantically tagged words to represent the data for classification of negotiations. We call those words *negotiation-related* and use them in the classification problems. The negotiation data are classified with respect to the negotiation outcomes (a binary classification problem), the negotiation roles (a binary classification problem) and with respect to negotiation roles and negotiation outcomes (a multi-class classification problem). Analysis of classification results completes the chapter.

- Chapter 8 explains how the negotiation strategies and the influence strategies are reflected in language. We concentrate on the language expressions of permission, possibility, necessity, volition, and rejection/denial. We explore how negotiators use language to manipulate the evaluation of events. We show that the strategies are related to the use of modals, mental and volition verbs, negations, and adjectives. We call such word types *strategy-related* and use them to represent the data in the classification problems. The chapter demonstrates the advantage of strategy-related features in representation of unsuccessful negotiations. The chapter shows the importance of different segments of the negotiation data for the outcome of classification. The summary of results concludes the chapter.

- Chapter 9 recapitulates the contributions made in this work and lists the directions for future work. The chapter provides the grounds for suggested directions and gives the starting points.

For the reader convenience we have a Glossary of Acronyms and an Index that list the major notions of the dissertation. The Glossary appears before the Table of Contents and the Index appears after Bibliography. We present a diagram of the dissertation chapters in Figure 1.2. The arrows show main connections between them.

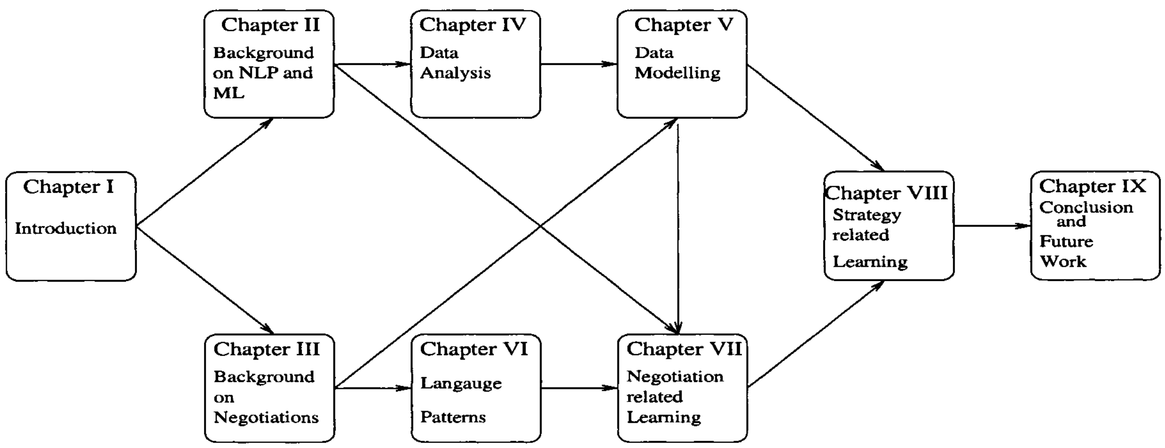

Figure 1.2: A map to the dissertation.

1.6 List of publications

Chapter 2' ML background and the results from chapters 4, 5, 6, 7, 8 partially appear in publications:

refereed – M. Sokolova, M. Shah and S. Szpakowicz, (2006), "Comparative Analysis of Text Data in Successful Face-to-Face and Electronic Negotiations", *Group Decision and Negotiations*, 15(2), 22 pages, to appear; an extended version of the FINEXIN paper;

Mario Marchand and M. Sokolova, (2005), "Learning with Decision Lists of Data-Dependent Features", *Journal of Machine Learning Research*, 6, p. 327 – 451;

M. Sokolova, V. Nastase, M. Shah and S. Szpakowicz, (2005a), "Feature Selection for Electronic Negotiation Texts", *Proceedings of RANLP'05*, p. 518 – 524;

M. Sokolova, V. Nastase, S. Szpakowicz and M. Shah, (2005b), "Analysis and Models of Language in Electronic Negotiations", *Issues in Intelligent Systems. Models and Techniques*, ISBN 83-87674-91-5, p.p. 197 - 211, EXIT, Warszawa;

M. Sokolova, V. Nastase and S. Szpakowicz, (2004a), "Language in Electronic Negotiations: Patterns in Completed and Uncompleted Negotiations", *Natural Language Processing (Proceedings of ICON'2004)*, p. 142 – 151;

M. Sokolova, M. Shah and S. Szpakowicz, (2005c), "Comparative Analysis of Text Data in Successful Face-to-Face and Electronic Negotiations", *Proceedings of FINEXIN 2005, Workshop on the Analysis of Informal and Formal Information Exchange during Negotiations*, p. 33 – 42;

M. Sokolova and S. Szpakowicz, (2005), "Analysis and Classification of Strategies in Electronic Negotiations", *Advances in Artificial Intelligence (Proceedings of AI'05)*, 18, 145 – 157;

M. Shah, M. Sokolova and S. Szpakowicz, (2004a), "The Role of Domain-Specific Knowledge in Classifying the Language of E-negotiations",*Natural Language Processing (Proceedings of ICON'2004)*, p. 99 – 108;

M. Sokolova, S. Szpakowicz and V. Nastase, (2004c), "Using language to Determine Success in Negotiations: A Preliminary Study", *Advances in Artificial Intelligence (Proceedings of Canadian AI'04)*, p. 449 – 453;

M. Sokolova, M. Marchand, N. Japkowicz and J. Shawe-Taylor, (2003), "The Decision List Machine", *Advances in Neural Information Processing Systems (Proceedings of NIPS'2002)*, 15, p. 921 – 928;

non-refereed – M. Shah, M. Sokolova and S. Szpakowicz, (2004b), "Using Domain-Specific Knowledge to Classify E-negotiations", InterNeg Report 07/04;

M. Sokolova and S.Szpakowicz, (2004), "Language in Electronic Business Negotiations", *Proceedings of InterNeg International Seminar "Markets, negotiations and dispute resolution in the new economy"*;

M. Sokolova, S. Szpakowicz and V. Nastase, (2004b), "Automatically Building a Lexicon from Raw Noisy Data in a Closed Domain", InterNeg Report 01/04;

Chapter 2

Background Review

A good place to start from is where you are.

Murphy's Law

2.1 Introduction

In this chapter we present an overview of NLP and ML research which we consider relevant to this dissertation. The NLP methods are presented in enough detail to let researches working in other fields to understand the discussion; the same applies to the presentation of the ML methods. In both cases, we justify the choice of methods and show their appropriateness for this work. The appropriateness is shown on the example of e-negotiation data. We present feature selection in a separate section because of its importance for this work. Moreover, a separate presentation is a prelude to our feature selection approach which combines knowledge-based NLP and ML techniques.

We discuss the relevant research on electronic negotiations in Chapter 3. Section 3.4.3 lists the applications of Artificial Intelligence to electronic negotiations and presents reinforcement learning.

Section 2.2 presents Natural Language Processing methods and Section 2.3 presents Machine Learning methods. Both sections provide references to the further parts of this dissertation, where the methods are used. Section 2.4 reviews previous research on feature selection. Section 2.5 lists evaluation measures. Section 2.6 reports conclusions of the background study. Section 2.7 contains bibliographical notes.

2.2 Natural Language Processing

In this section we present the NLP tools used in our research. In applying NLP techniques, our goal is to minimize the manual intervention in the procedures. Our aim is to explain why we have used these tools and to show their positive and negative characteristics with respect to our work.

2.2.1 Corpus analysis methods

We start by reviewing the N-gram-based corpus analysis methods; for a comprehensive overview see (Kilgarriff, 2001). The reliability of word frequencies makes the statistical results a trustworthy measure although for a deeper comparison of two corpora we consider other features, such as part-of-speech distribution and lexical semantics.

We explore vocabulary richness of the corpora based on the information on rare words, *hapax legomena* ($V(1,N)$) and *dis legomena* ($V(2,N)$); $V(i,N)$ is the number of word types that occur i times in the text with N word tokens. We calculate the growth rate of the vocabulary

$$P(N) = \frac{V(1,N)}{N} \tag{2.1}$$

and Sichel's characteristic

$$S(N) = \frac{V(2,N)}{N} \tag{2.2}$$

$P(N)$ is a common measure of the richness of the vocabulary of a text. $S(N)$ also measures the richness of the vocabulary and its homogeneity. Sichel (1986) and Thomos et al. (2004) state that $S(N)$ in addition to $P(N)$ provides more reliable results than only $P(N)$.

Corpus comparison looks for differences and similarities between corpora. Finding and estimating similarities is a more difficult task. Different sizes of the comparable corpora suggest the use of rank-based methods such as Spearman's Rank Correlation Coefficient (SRCC)

$$\rho = 1 - \frac{6\sum d^2}{N(N^2-1)} \tag{2.3}$$

where N is the number of words and d is the difference in the ranks of the same word in two corpora (Oakes, 1998). SRCC has shown accurate – and better than other methods – results in comparing corpora with respect to similarity. However, it tends to overemphasize words with low frequencies. These words do not contribute much to general trends of the language and are not of special interest to our work. In section 4.5 we use SRCC to compare words which have high frequency and contribute to general trends of the language, thus avoiding the pitfalls of SRCC.

Rayson and Garside (2000) propose a significant relative frequency difference technique. We use it to find words on which corpora differ. We compute the log-likelihood statistics LL for each word w:

$$LL(w) = 2 * ((a * log(\frac{a * (a + b)}{c})) + (b * log(\frac{b * (a + b)}{d}))) \tag{2.4}$$

where a and c are the number of occurrences of w and the number of word tokens in the first corpus respectively; b and d – in the second corpus. The higher the $LL(w)$ is, the larger the difference between the frequencies of the word w in two corpora. Although the standard use of this method is two-tailed/non-directional, we have introduced the obvious constraint $a \leq b$ and use this modification as as one-tailed/directional. The directional version allows us to extract words which are frequent in one corpus and rare in the other, and find out on which words the two corpora differ. The non-directional version is used when we need to represent both corpora through words on which they differ. We call the found words *indicative words.* One example of the use of indicative words is when texts used to build both the corpora are represented as bags of words and then classified as belonging to one corpus or another by ML algorithms. We present experiments and report results in Sections 7.7 and 8.8. It is worth noting that the poor performance of LL on low frequency words does not affect our results.

We compare the use of a word in different corpora by calculating its mutual information (Church and Hanks, 1989). Mutual information MI relates the two corpora and a word w:

$$MI(w) = log(\frac{a * (c + d)}{c * (a + b)}) \tag{2.5}$$

where a, b, c and d are the same as in Formula 2.4. We use MI on the common words because of its reliable performance on frequent events. However, MI tends to overemphasize rare events, thus it cannot be a reliable measure for low-frequency word (Kilgarriff, 2001; Oakes, 1998).

We employ those methods and report the empirical results in Sections 4.5, 7.7, 6.2, 8.6 and 8.8.

2.2.2 Statistical Natural Language Processing methods

In this section we briefly present the statistical NLP methods used in our research. We employ statistical modelling because it accommodates the vocabulary growth of an unrestricted language (see section 4.5). Again we show why we used these tools in our work and what are their strong and weak qualities with respect to the task. To build the first reliable portrait of our data we build N-grams $w_1 \ldots w_N$, where $N = 1, 2, 3, 4$. We study N-gram frequencies because in spite of being a simple data representation they are a necessary and important step in understanding data and for selecting appropriate tools to work with the data: for each corpus originating from a specific genre or source, the N-gram frequency distribution is one of the essential characteristics (Tweedie and Baayen, 1998).

As the next step we construct N-gram models by computing

$$P(w_k|w_1^{k-1}) \approx P(w_k|w_{k-N}^{k-1}) \tag{2.6}$$

where $P(w_k|w_1^{k-1})$ is the probability of the word w_k appearing after the sequence of words $w_1 \ldots w_{k-1}$ We have used a unigram model for data-dependent automatic spelling correction described in section 5.4.2, applying it with *ispell*, an off-the-shelf spell-checker. In Chapter 6 N-gram models are used to find language patterns.

In recent years statistical language modelling has found a variety of applications in studying text data; for a general overview of the field see Rosenfeld (2000) and for an overview and analysis of practical applications see Chen and Goodman (1998). The classical uses of statistical models are found in machine translation, information retrieval, speech recognition. We have employed statistical language modelling to analyze the level of uncertainty and predictability that the (electronic and face-to-face) negotiation data bears.

N-gram models are arguably the most widely used models for statistical and language modelling purposes. An important concern in such modelling is *smoothing* that helps account in advance for previously unseen N-grams. For the Good-Turing model (Good, 1953) we use Katz smoothing (Katz, 1987) (GTK model) that generally performs well on data with a low type-token ratio. The model – given by Equation 2.7 – uses the frequencies of frequencies for N-grams . The probability P_{GTK} of the previously seen N-gram

$$P_{GTK}(w_1...w_n) = \frac{r^*}{N} \tag{2.7}$$

where N is the number of tokens, r^* is the estimate of the N-gram frequency r:

$$r^* = (r+1)\frac{n_{r+1}}{n_r} \tag{2.8}$$

n_r is the number of N-grams that occur exactly r times in the training data. This estimation approximates N-grams with frequencies > 5 because of the Katz' assumption that the large counts are usually reliable. If N-gram has occurrence $k \leq 5$ then Katz approximation combines higher order N-grams with lower order N-grams. It is found according to the formula:

$$\frac{(k+1) - \frac{(k+1)n_{k+1}}{n_1}}{1 - \frac{(k+1)n_{k+1}}{n_1}} \tag{2.9}$$

We use Katz smoothing for a few theoretical and practical reasons. Katz smoothing generally performs well on data with a low type-token ratio. It is easier to implement. It has very few parameters and does not require a validation set for adjustable parameters. The size of our data is another argument for Katz smoothing.

Another statistical model which has caught our attention is the model by Kneser and Ney (1995). The Kneser-Ney model (KN model) has plugged-in smoothing where the probability of the unseen N-gram is substituted by the weighted probability of the $N-1$-gram. The probability P_{KN} of the previously seen N-gram

$$P_{KN}(w_1...w_n) = \max\{c(w_1 \ldots w_{n-1}) - D, 0\} + DN_{1+}(w_1 \ldots w_{n-1})P_{KN}(w_n|w_2 \ldots w_{n-1}) \tag{2.10}$$

where $c(w_1 \ldots w_{n-1})$ is the number of occurrences of N-gram $w_1 \ldots w_{n-1}$, $D = \frac{n_1}{n_1+n_2}$, where n_i is the same as above, $N_{1+}(w_1 \ldots w_{n-1})$ is the number of different $N-1$-grams preceding w_n.

The KN model makes the unigram probability proportional to the number of different words that it follows, not the number of occurrences of the word in the corpus. The probability of the previously unseen word is $\frac{1}{N}$, where N is the number of tokens. Chen and Goodman (1998) state that it performs well with respect to collocations. This makes the Kneser-Ney model appropriate for modelling data obtained from a closed domain where certain collocations are frequently used, e.g. *delivery time* and *last offer* are frequent in purchase of goods; this is the main attraction of this model to our research since we work in a closed domain with a fixed topic.

We report practical results in Sections 5.2 and 5.3. In Section 9.5.3 we use the statistical model to analyze how uncertainty and predictability vary with respect to gathering data during different phases of negotiations.

2.2.3 Symbolic Natural Language Processing methods

We employ NLP techniques in an "in-house" semi-automatic procedure for extracting and building a corpus-based semantic lexicon; see Section 2.2.4 for the details on a semantic lexicon. We introduce bootstrapping procedures for learning language patterns. Chapters 6 and 7 present these procedures and the results of their applications. In these procedures we have used an off-the-shelf spelling corrector *ispell* (isp, 2001).

We use tagging to study part-of-speech distribution in the corpus. The accuracies of the most successful approaches and of the baseline are high, $\approx$ 96% and $\approx$ 91% respectively; for examples see (Marcus et al., 1993; Carroll et al., 2003) . Thus, we expect the tagging results to find the most likely syntactic category from among the possible categories for a particular word in a sentence to be accurate.

We employ the transformation-based Brill's tagger (Brill, 1995) which has common features with rule-based taggers and with stochastic taggers. The former tag ambiguous words according to rules, while the latter construct the rules from a previously tagged training corpus. Brill's tagger is based on rules that specify with what tag a word should be labelled. However, the rules are automatically induced from the data. The learning process is supervised because it requires a correct tag for each word in the training set. We provide the details of Brill's algorithm to show that because of the learning component it is a proper tagger for the data with the characteristics different from the standard data; see Sections 4.5 and 4.7 for the data analysis.

To go beyond words we apply parsing to analyze semantic information present in the data. We run the Xerox Incremental Parser (Chanod and others, 2001). During three steps of parsing it uses rule-based disambiguation based on Hidden Markov Model taggers, layered chunking rules and dependency relationships that are specified by the grammar. The incremental behaviour is acquired by selecting and throwing away inconsistent parts of the tree. Then the parser suggests the components that connect the tree and are grammatically correct. Further in the dissertation we use XIP to build the intra-sentence dependencies. The dependencies show linguistic relations between words in a sentence. The specific dependencies and their use are reported in Sections 8.5, 7.10 and 8.12. We prefer XIP to free MiniPar (Lin, 2001) and Link (Sleator and Temperley, 1991) parsers because of its robustness: it parses sentences challenging for the other two parsers. Next, XIP builds more dependency relations than MiniPar. These dependencies are organized in levels, unlike in the Link parser.

We present the results of applying symbolic NLP methods in Sections 4.5.2, 5.4.2 and throughout Chapter 8.

2.2.4 Semantic Lexicons

In the dissertation we perform knowledge-based learning from texts. The learning process requires tools that assist in text understanding. One of such tools extensively used in knowledge-rich learning environments is a semantic lexicon.

Semantic lexicons are built to support text understanding and inference (Hahn et al., 1999). The lexicons present categorical information about the meaning of terms and might present significant relations between terms. The term can be a word or a phrase. The meaning of a term is determined by a task which language understanding attempts to solve. In this work we consider a meaning that is taken in the context of communication (Matthews, 1995). It differs from a literal, or grammatical, meaning determined outside of context (Frawley, 1992). In recent years, due to the increase of knowledge-based research in bioinformatics and medicine, the semantic lexicons were built and widely used for applications in these fields; for an overview see (Thompson and Mooney, 2003).

Building a lexicon starts with the use of annotated corpus as training data and then continues with learning semantic information from it. Unfortunately, not many semantically annotated corpora are available and in some disciplines there are no such annotated corpora at all. This situation promotes search for general-purpose semantic resources. Examples of such resources are WordNet (Fellbaum, 1998), ontologies (Nirenburg and Raskin, 2004) or Roget's Thesaurus (Kirkpatrick, 1998). The resources satisfying the task of building a lexicon are selected to annotate training data.

The learning procedure depends on both the training data and the desirable output. There is no well-established rule by which the learning process abides. However, the underlying agreement is that time- and effort-consuming manual intervention should be kept at the necessary minimum.

In Chapter 5 we build a semantic lexicon using the electronic version of Longman Dictionary of Contemporary English (LDOCE) (Summers, 2003). We choose LDOCE among other resources because it gives the necessary semantic information through assigning semantic field tags to a word.

2.3 Machine Learning

The machine learning aims to find description of the data from its sample of a limited size (Herbrich, 2002). Natural language text is one of the most important and challenging applications for machine learning (Shawe-Taylor and Cristianini, 2004).

In this section we present the *supervised* ML problem and analyze ML methods and justify the use of these methods for our purposes. In supervised learning the training input examples are given with their class labels. We give a brief overview of main learning tasks with respect to Web-based text data. Recall from the chapter's *Introduction* that we describe reinforcement learning in Section 3.4.3.

2.3.1 The classification problem

We start with the statement of the classification problem which belongs to supervised learning. An input example $\vec{x} = (x_1, \ldots, x_n)$ needs to be classified into one (and only one) of j classes (or groups) $C_1, \ldots, C_j$. The existence of the classes is known a priori. Classification is concerned with the relationship between the class-membership label y and the input vector $\vec{x}$.

In the binary classification problem the output takes on symbolic values $y \in \{0, 1\}$ corresponding to two classes, negative and positive respectively. In the multi-class classification problem the output takes on values $y \in \{c_1, ..., c_m\}$, corresponding to m classes. A function $f(\vec{x})$ maps the input example $\vec{x}$ onto the class label y.

We use the above definition to define classes in e-negotiation data; see Section 4.4. We solve classification problems throughout this work.

For better understanding of the classification results we remind how learning occurs during binary classification. The learning process consists of the following steps: a learner builds a classifier; the classifier generalizes from characteristics provided by the positive examples and specializes from characteristics provided by the negative examples, where the former includes examples and the latter excludes examples. Hastie et al. (2001) provides details and examples on machine learning paradigms.

Let us consider the case when the positive examples are classified much better than the negative examples. This shows that the classifiers generalize rather well. At the same time classifiers do not specialize well enough. This means that the classifiers extract sufficient information from positive examples but cannot extract sufficient information from negative examples. This corresponds to the situation when:

1. the positive class is either homogenous or consists of a few well-represented subclasses;

2. the negative class is divided into several small subclasses and some subclasses of the negative class are under-represented.

Then a classifier can miss under-represented subclasses; Japkowicz (2001) gives several examples. We employ similar reasoning in Section 7.9 for the analysis of successful and unsuccessful negotiations and in Section 8.10.3 for the analysis of different negotiation segments.

2.3.2 Learners

To apply ML methods for supervised learning and to choose which methods will learn better on our data, we consider the data characteristics and further use of the learning results.

Here we state the main characteristics of the negotiation data with respect to ML techniques:

- the data are textual and, possibly, numerical;
- the data are high-dimensional when negotiations are represented through texts or through the negotiation process;
- the data are noisy.

According to the listed characteristics, the learning algorithms should perform well on the textual and high-dimensional data, be noise-resistant and fast. Here we emphasize that we do not claim that our data is sparse. Although sparsity is expected to be one of the main characteristics of the textual high-dimensional data, this rule applies only partially to the e-negotiation data. The text data of e-negotiations are sparse for low frequency terms and dense for high- and medium-high frequency words (excluding function words). If the data are represented as bags of word occurrences then for the low-frequency words rare "ones" are scattered among many "zeros", while for the high frequency words there are almost no "zeroes" and occurrences are rather high.

Another requirement on ML algorithms that comes from the first goal of our studies – to investigate the data – is the easy understandability of the algorithm output for humans. The "easy-to-understand output" and the "need to run fast" requirements eliminate the use of neural networks whose output is difficult to decipher and that run relatively slowly.

We have to say that usually in text categorization and classification ML methods employ the seeds of manually labelled data; for references see Section 2.7. We do not have access to manually labelled text data, nor do such data exist in large amounts. We seek methods that (almost) eliminate manual intervention.

We select ML methods such that for each method at least two of the following requirements are met:

– the method's output can be easily interpreted by humans;

– the method should be resistant to noise and work well on different data distributions;

– the method has been shown previously to give high accuracy in text classification.

Decision-based classifiers produce output that is easy to understand and they are often used with good results in text classification tasks (Gabrilovich and Markovitch, 2004). For classification procedures we use decision-based classification algorithms such as Decision Trees (DT). We use them in two forms. Decision Stumps (DS) (Witten and Frank, 2000) are basically the first level of nodes of Decision Trees. C5.0 is a decision tree learner that classifies entries by separating them into classes according to information gain $G(a, y)$ of the attributes

$$G(a, y) = H(t) - H(t|a) = H(t) - (p_L H(t_L) + p_R H(t_R)) \tag{2.11}$$

where $G(a, y)$ is used for class label in binary classification problem, a is the splitting attribute, y is the value of a, t is the distribution of a, p_L, p_R, t_L, t_R are proportions of elements and the distributions of left and right nodes in that order (Quinlan, 1993). In general case, the tree needs not be binary.

In spite of the characteristics of decision-based classifiers we do not want to restrict ourselves to only one type of classifiers. We employ the instance-based classifier (IBK), namely k-Nearest Neighbour (kNN) (Duda et al., 2000), to classify the data. For the input x it calculates the output $\hat{\mathrm{Y}}(x)$ by averaging the labels of k closest points x_i in the training set:

$$\hat{\mathrm{Y}}(x) = \frac{1}{k} \sum_{x_i \in \mathrm{N_k}(x)} y_i \tag{2.12}$$

where $\mathrm{N_k}(x)$ is the neighbourhood of x and y_i is the label of x_i. When the distance metric is fixed, this classifier has only one adjustable parameter k, the number of neighbours of the example x for which the label should be found. Its output is also easy to understand.

Kernel methods, especially Support Vector Machines (SVM), have been successfully used for text classification (Shawe-Taylor and Cristianini, 2004). They satisfy the second requirement as well: kernel classifiers are resistant to noise and work well on data with various distribution. On the other hand, their output is difficult to understand. SVM builds a hyperplane that separates training examples from two classes, with the largest possible separation (Cristianini and Shawe-Taylor, 2000). The search for the hyperplane is done by solving a constrained optimization problem in dual representation:

$$\vec{w} = \sum_j \alpha_j c_j \vec{d_j}, a_j \geq 0, \tag{2.13}$$

where $\vec{w}$ represents the hyperplane, $c_j \in \{1, -1\}$ is a class label, α_j are found by solving a dual optimization problem. Support vectors $\vec{d_j}$ correspond to $a_j \geq 0$. SVM with Latent Semantic Kernels are among the most popular SVM modifications for text classification (Cristianini et al., 2001; Shawe-Taylor and Cristianini, 2004). However, kernels for texts normalize text length, remove stop words and do not consider word occurrence. The data pre-processing requires stemming. These characteristics make the text kernels not suitable for the classification of human communication texts. Further we present the results for linear SVM corresponding to the parameter values that give the best results from an exhaustive search of possible parameter values.

We also apply probabilistic classifiers. Fast and easy to implement Naive Bayes (NB) (Duda et al., 2000) has been used as a baseline for text classification (Rennie et al., 2003). NB assigns the class label $l^* = \arg\max_l P(l|example)$ to an entry *example* , where $P(l|example)$ is the probability of the class label l on the entry *example*:

$$P(l|example) = \frac{P(l)P(example|l)}{P(example)} \tag{2.14}$$

or

$$posterior = \frac{prior * likelihood}{evidence} \tag{2.15}$$

Evidence does not affect the selection of l^*. Further simplification comes from the assumption that for each entry *neg* the attribute values f_i are conditionally independent given *example*'s class. This assumption makes it possible to decompose *likelihood* into the product of probabilities $P(f_i|l)$. The resulting rule for assigning the class label is the following:

$$l^* = \arg\max_l P(l) \prod_i P(f_i|l) \tag{2.16}$$

We apply Decision List Machine (DLM) because of its high accuracy on the "real life" data where it often outperforms SVM (Sokolova et al., 2003). DLM represents the class of data-dependent classifiers (Marchand and Shawe-Taylor, 2002) whose basic idea is to construct features from the original data and then use these features to build a classifying function. Let $\mathbf{x}$ denote an arbitrary n-dimensional vector of the input space X which could be arbitrary subsets of $\Re^n$. We consider binary classification problems for which the training set $S = P \cup N$ consists of a set P of positive training examples and a set N of negative training examples. We define a *feature* as an arbitrary Boolean-valued function that maps X onto $\{0, 1\}$. Given any set $\mathcal{H} = \{h_i(\mathbf{x})\}_{i=1}^{|\mathcal{H}|}$ of features $h_i(\mathbf{x})$ and any training set S, the learning algorithm returns a small subset $\mathcal{R} \subset \mathcal{H}$ of features. Given that subset $\mathcal{R}$ and an arbitrary input vector $\mathbf{x}$, the output $f(\mathbf{x})$ of the Decision List Machine (DLM) is defined to be:

If $h_1(\mathbf{x})$ then b_1

Else If $h_2(\mathbf{x})$ then b_2

...

Else If $h_r(\mathbf{x})$ then b_r

Else b_{r+1}

where each $b_i \in \{0, 1\}$ defines the output of $f(\mathbf{x})$ if and only if h_i is the first feature to be satisfied on $\mathbf{x}$ (*i.e.* the smallest i for which $h_i(\mathbf{x}) = 1$). The constant b_{r+1} (where $r = |\mathcal{R}|$) is known as the *default value*.

We analyze the preliminary classification results and select only the learners producing the classifiers with the highest accuracy. To make a fair comparison of all classifiers we do not perform any additional data preprocessing such as scaling.

In addition to using the existing learning algorithms, we have designed and implemented an "in-house" bootstrapping learning procedure. The procedure allows us to learn patterns from the text data. The description of the procedure is given in Section 6.2. The procedure is straightforward, easy to implement and requires only the basic additional information.

Before going further, we say that in all the experiments we employ L_2 metric and exhaustive search among the adjustable classification parameters.

2.3.3 Learning from Web-based text data

In the dissertation we work with text data gathered through a Web-based system. Thus, our work is included into a broader phenomenon – learning from Web-based text data (Cooley et al., 2000).

The Web-based data attract enormous attention as the object of studies for information retrieval and extraction, text categorization and text classification communities (Sebastiani, 2002). The number of text classification publications has grown rapidly, with the increasing diversity of topics. Here we introduce only the topics relevant to the dissertation.

The Internet text data can be divided into two broad categories: the mass media type category, e.g. articles, advertising, and the communication category, e.g. email, instant messaging. ML and NLP methods perform different tasks with respect to the data from the former category. Mostly these tasks involve mining the large volumes of documents e.g. well-structured and edited texts written according to the established norms albeit relaxed in comparison with the pre-Internet era (Joachims, 2002). The application of NLP and ML methods to the latter category – Computer-Mediated Communication – concentrates on topic classification and filtering and often works with relatively small collections of email texts; for the detailed overview of email classification topics see Kiritchenko and Matwin (2001).

Recently the *Enron* data has attracted attention of the research community. The public *Enron* data are the collection of email exchanged by the *Enron* employees (Klimt and Yang, 2004). The published papers address email filtering (e.g., Martin et al. (2005) use the *Enron* email records to identify spam), statistical modelling (e.g., Bickel et al. (2005) use an interpolated N-gram model to complete sentences), social networks (e.g., McCallum et al. (2005) present a social network analysis of *Enron* discussion topics combined with the sender-recipient relationships). To the best of our knowledge nobody has studied the *Enron* email texts as the purposeful interpersonal communication process.

Kosala and Blockeel (2000) present the survey of Web mining. The authors, among other topics, summarize research on the mining of what we refer as text data sets. Laender et al. (2002) and Sebastiani (2002) survey the methods used in information extraction and text classification and give many references related to the Web text data. We list specific references in Section 2.7.

With respect to the classification goal the text classification is divided into two broad categories: topic and non-topic classification. The latter includes genre and gender classification, sentiment and emotion classification, etc. Mostly these classification tasks apply to the document and email data. We list the related references in the chapter's *Bibliographical notes* and throughout the dissertation. For references on feature selection for text classification and other specific methods refer to Section 2.7.

The emerging field of Web-based text data is text mining. It seeks, locates, extracts and generalizes new knowledge (Hearst, 2003). The knowledge mining from bio-informatics texts is the most commonly known up to date application of text mining; for example, see GENIA project (Tsujii and others, 2005). We introduce new classification tasks in Section 4.4.

2.4 Features and feature selection

For the application of learning methods the data should be represented by its features. The selection of features representing data is as important as the selection of learners themselves. Feature selection is especially important when the number of data features is substantially larger than the number of available examples. This always happens when we work with texts. The problem of feature selection becomes harder when texts are unedited transcripts of freely written communication.

2.4.1 Features in Natural Language Processing

In Natural Language Processing the number of available features and their variety are extremely large (McEnery and Wilson, 2001; Oakes, 1998). A few examples of features include:

- individual words, e.g. an, Moon, computer;
- morphemes, e.g. -s, -ed, -ies,
- words classes,e.g nouns, prepositions;
- grammatical features, e.g. clauses, passives;
- computed measures, e.g. cross-entropy, frequency of an event;
- constructed features, e.g. combinations of the individual features;
- style characteristics, e.g. misspellings, abbreviations, business terminology.

To make this review feasible – and related to the current work – we discuss features used in *learning from Web-based texts.* We roughly divide the text research on two groups: a) research conducted on the text level, and b) research conducted on the sentence level.

On the text level the problems of text classification and information extraction attract the most attention. Features used in these fields are mostly, but not always, individual words or word categories, where categories satisfy different criteria. Sebastiani (2002) gives a detailed overview of text classification and analysis of the most common features. For specific examples see Lin et al. (2003) and Yakushiji et al. (2004) (the use of terms in information extraction) and Caropreso et al. (2001) (the use of N-grams in text classification).

There is no golden rule what approach to take in search for features, even within one class of problems, e.g. text classification. Advantage of automatic vs manual feature identification depends on a specific task. Pang et al. (2002) give an example of an advantage of automatically over manually found features. On the contrary, Zhang et al. (2004) provide an example where manually found features give better learning results.

On the sentence level the situation is even more diverse. Depending on the tasks, features might be syntactic patterns (Riloff et al., 2005), annotated parts of speech (Nasukawa and Yi, 2003) or even contain non-language symbols (Wolska and Kruijff-Korbayov, 2004). Hutchinson (2004) presents a comparison of two types of features. The first type contains knowledge-rich features (linguistic; obtained from parse trees, with additional information). The second type contains knowledge-poor features (lexical co-occurrences). The acquiring of the meanings of discourse markers is the goal of learning. The results show that lexical co-occurrences provide better *best results* than more sophisticated linguistic features. However, the linguistic features provide better *worst results*. The author attributes this to the connection between linguistic features and specific meanings.

The first step, and a major challenge, is to find the **initial set of features** relevant to a problem that NLP researchers want to solve. In most cases, the problem of finding relevant features relies mostly on the intuition of a researcher; for more examples refer to Section 2.7. Then statistical analysis and Machine Learning methods are usually used to select a (small) subset from the initial set of features; see the next section.

2.4.2 Feature selection methods

In the context of (machine) learning, feature selection helps increase classification accuracy by finding relevant features among all the available features. In this section we briefly discuss the major approaches to feature selection.

Attempts to perform feature selection can be broadly categorized as:

1. pure statistical study of the data;

2. the presence of some expert intervention.

To escape dependency on human expert knowledge, machine learning methods have been developed to automatically select features; for definitions and general overview see (Blum and Langley, 1997). For a detailed overview of feature selection methods and their recent improvements refer to (Guyon and Elisseeff, 2003). Selection is based only on the statistical characteristics of the features. The methods function autonomously without any expert intervention and can again be roughly categorized as *embedded*, *filter-based*, *wrappers-based*, and *hybrid*, a combination of filter- and wrapper-based. An embedded selector gives the work of selecting features to a learning algorithm itself. A filter filters out irrelevant features before the application of a learning algorithm (Blum and Langley, 1997). We employ filters in Chapters 7 and 8. A wrapper is applied in conjunction with the learning algorithm. It removes recursively the features that get small weights from the classifier. Wrappers have been shown to outperform filters, but only at a great computational expense (Stracuzzi and Utgoff, 2004). Wrappers are slow. This drawback makes them unsuitable to our research goals, so we have not included them in this work.

We summarize ML feature selection approaches in Figure 2.1. The classification of genes and

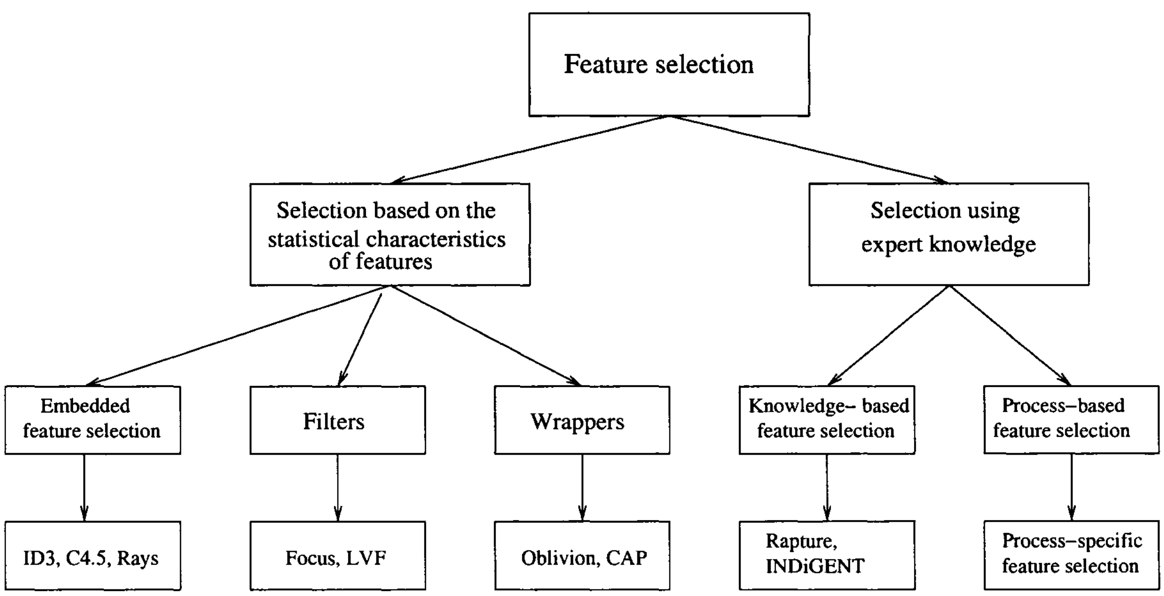

Figure 2.1: Summary of feature selection approaches

text classification are two ML application areas where automated feature selection is well-studied and widely used; see Marchand and Shah (2005) and Cristianini et al. (2001) for the examples of successful feature selection in gene and text classification respectively.

The other category of feature selection approaches are those that use some expert knowledge. For instance, knowledge-based feature selection embeds human knowledge in a formal setting. Knowledge-based feature selection is guided in part by expert knowledge about the classification domain. It performs a generic search in the feature space in order to select those features that best classify data from the domain, when considered in conjunction with the given domain theory. There are a number of systems that use neural network encodings to refine a variety of representation types including finite state automata, certainty-factor rule bases, first-order Horn clause logic, and a simplified form of Horn-clause logic; for overview refer to Burns and Danyluk (2000). When the application domain originates from a multi-dimensional human activity that, as any other human activity, includes random elements, formal logic is unsuitable and insufficient to describe the expert knowledge about the domain. In these cases, expert knowledge about the domain needs to be presented using a descriptive model.

Many (although not all) feature selection algorithms perform some form of search in the space of variable subsets as part of their operation; see Cormen et al. (2000) for details. A forward selection algorithm begins with the empty set and searches for variables to add. A backward elimination algorithm begins with the set of all variables and searches for variables to remove. The former is sometimes also referred to as an incremental, when the latter as a decremental approach to feature selection. Optionally, forward algorithms may occasionally choose to remove variables and backward algorithms may choose to add variables. This allows the search to *backtrack* and recover from previous poor selections. The advantage of forward selection is that the size of the subsets will remain relatively small helping to speed up the data classification. On the other hand, backward elimination has an advantage in relatively easily recognizing irrelevant variables. Removing a relevant variable from an otherwise complete set should cause a decline in the evaluation while adding a relevant variable to an incomplete set may have little immediate impact.

We use feature selection methods in Chapters 7 and 8.

2.5 Evaluation measures

In the dissertation we compare performance of statistical models, feature selection methods and classifiers.

We employ t-test to compare two samples. We employ the Analysis of Variance (ANOVA) method when we compare more than two samples. ANOVA tests if there are significant differences between the means of the sample characteristics we want to compare. In order to do this, the variance between the sample characteristics is calculated and then compared with the variance within each sample. Samples originate from different populations if the variance between them is significantly greater than within each sample; for details on t-test and ANOVA see Oakes (1998).

We need to evaluate the goodness of fit for statistical models on our data. The standard measure for evaluating statistical models is cross-entropy

$$-\frac{1}{n}\sum_{i=0}^{n} log(P(w_i)) \tag{2.17}$$

where n is the number of words in the test set and $P(w_i)$ is the probability of the appearance of the word w_i in the test data. $P(w_i)$ approximates the unknown probability distribution that the data actually have (Charniak, 1993). By definition, cross-entropy estimates the average number of bits necessary to encode the test data using an algorithm associated with $P(w_i)$. The use of compression makes cross-entropy different from entropy. The latter estimates the average number of bits necessary to encode each of the words of test data using the algorithm with optimal encoding. Obviously, cross-entropy is the upper bound on entropy. A model with the lower cross-entropy on the test set models the data better. With respect to language, cross-entropy is one of the estimators of the complexity or predictability of a language. A lower cross-entropy means that the data is predictable while a higher cross-entropy indicates high uncertainty in the data.

There exists several measures of the performance of classifiers. Statistical McNemar's test and the resampled paired t-test compare the relative performance of two classifiers (Dietterich, 1998a). Research in Computational Learning Theory proposes the generalization error bounds. The bounds can be data-dependent, e.g., sample compression bounds (Marchand and Sokolova, 2005), used to estimate the performance of data-dependent classifiers, or data-independent, e.g. the well-known Vapnik-Chervonenkis bounds (Vapnik, 1998) used to estimate the performance of classifiers when they belong to a continuous set. However, the purpose of the dissertation – with respect to Machine Learning – is to apply ML tools on the text data and empirically estimate how well they can identify the label of a previously unseen data entry. Therefore we find sufficient to estimate the classifier performance by calculating its accuracy on the test data:

$$Accuracy = \frac{correctly\ identified\ examples}{all\ examples} \tag{2.18}$$

For relations between the generalization error and the accuracy of a classifier see Cherkassky and Muller (1998).

We employ the standard text categorization metrics precision, recall, F-measure.

$$Precision = \frac{correctly\ identified\ positive\ examples}{all\ returned\ positive\ examples} \tag{2.19}$$

$$Recall = \frac{correctly\ identified\ positive\ examples}{all\ positive\ examples} \tag{2.20}$$

$$F - measure = \frac{(\beta^2 + 1) * precision * recall}{\beta^2 * precision + recall} \tag{2.21}$$

Precision, depending on true positives and false positives, and recall, depending on true positives and false negatives, are antagonistic. In our work precision and recall are equally important. For the remainder of the dissertation β is set to one because the equal weights of precision and recall. Equally-weighted F-measure tends toward results with more true positives. For the applications of various F-measures refer to Kazawa et al. (2005).

We use ten- and five-fold cross-validation to estimate the accuracy of classification. In K-fold cross validation the cases are randomly divided into K mutually disjoint sets of approximately equal size. The concept is learned from the examples in $K - 1$ sets and is tested on examples from the remaining set. This is repeated K times, once for each set (that is, each set is used once as a test set). The average error rate over all K sets is the cross-validated error rate. Five-fold cross validation takes $K = 5$ and ten-fold cross-validation takes $K = 10$. The ten- and five-fold cross-validations are chosen because of their high accuracy and reliability of results (Cherkassky and Muller, 1998).

Although we use different classifiers in our work, we do not employ the specific methods of comparing classifiers (Dietterich, 1998a). These methods concentrate on evaluation of the learners, whereas our task is to evaluate the applicability of the combination of selected features, their representations and classifiers to represent and learn from e-negotiation data.

We employ regression analysis (Duda et al., 2000) when the number of results requires generalization. We calculate polynomial regression to extract the commonalities when a large number of results are obtained while solving the same problem. Equation 2.22 calculates a m-order polynomial:

$$\hat{\mathrm{Y}} = \hat{\beta}_0 + \sum_{j=1}^{m} \mathrm{X}^j \hat{\beta}_j \tag{2.22}$$

where $\hat{Y}$ is a prediction of the output Y, $X = (X_1, \ldots, X_n)$ is the input vector, $\hat{\beta} = (\hat{\beta}_0, \ldots, \hat{\beta_m})$ is the vector of coefficients. The *least squares* method (Hastie et al., 2001) is the most popular method to calculate the coefficients. The coefficients β have to minimize the *residual sum of squares*:

$$RSS(\beta) = (\mathbf{y} - \mathbf{X}\beta)^{\mathrm{T}}(\mathbf{y} - \mathbf{X}\beta) \tag{2.23}$$

where $\mathbf{X}$ is a $k \times n$ matrix with each row is an input vector, T denotes matrix transpose, $\mathbf{y}$ is a k-vector of the outputs in the training set. If $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ is nonsingular[1] then the unique solution is given by

$$\hat{\beta} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y} \tag{2.24}$$

which then is substituted in Equation 2.22. Note that with the fixed input and output training vectors Equation 2.22 becomes linear with respect to coefficients. Thus, it does not require the extensive amount of data.

We use evaluation measures throughout the dissertation.

2.6 Conclusion

After reviewing the work done on applications of ML and NLP methods, especially to data obtained through the Web, we have concluded that the following problems either were not addressed or have not received proper attention:

1. the problem of learning the outcomes of human activities from freely written text data, especially non-topical learning of short informal texts;

2. application of Natural Language Processing and Machine Learning methods to Computer-Mediated Communication text data;

3. non-topical text classification of freely written human communication texts;

4. Text Mining of human communication texts.

We address the listed problems in the remainder of our dissertation either by obtaining results or proposing future work.

[1] A square matrix $\mathbf{A}$ is nonsingular if there exists a square matrix $\mathbf{B}$ such that $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$, where $\mathbf{I}$ is identity matrix.

2.7 Bibliographical Notes

In this section we list annotated related work.

Sichel (1986), Thomos et al. (2004) explore the statistical measures of vocabulary richness.

Tweedie and Baayen (1998) provide a detailed analysis of corpora measures and their dependence on corpora sizes.

Rayson and Garside (2000) employ word frequencies to compare corpora.

Kenne and O'Kane (1996) build a statistical model in a task-oriented spoken dialogue; they implement the role-dependent information in the speech recognition model for court procedures.

Solsona et al. (2002) build a language model for a dialogue system.

Popescu-Belis et al. (2004) build a corpus from a dialogue data.

Jokinen et al. (2001) apply cognitive maps to study dialogue management.

Sahlgren (2002) compares the application of transformation-based learning and vector-space approaches to dialogue management.

Mast et al. (1996) employ semantic information to classify dialogue acts.

Demetriou and Atwell (2001) suggest the use of LDOCE as a knowledge base for domain-independent semantic tagging.

Hutchinson (2004) and (2005) studies discourse relations and signals of relations between discourse units.

Jarmasz and Szpakowicz (2004) compare the measure of semantic similarity by WordNet and Roget's Thesaurus.

Nastase and Szpakowicz (2003) introduce an algorithm of combining WordNet' structures with the LDOCE word definitions.

Forman (2003) empirically compares feature selection methods on a number of text classification problems.

Kohavi (1995) explores wrappers and their positive impact on decision trees performance.

Opitz (1995) refines domain theories through the addition of rules between the input and output features (The Regent system).

Mahoney (1996) genetically refines the topology of hidden nodes in the networks and attempts to improve the set of input features by adding relevant (as judged by Quinlan (1986) information gain metric) if the network fails to reach a certain level of accuracy(The Rapture system).

Forman (2004) discuss the problems of feature scoring methods used in multi-class classification.

Caropreso et al. (2001) suggest a statistical N-gram selection procedure used for text representation in text classification. The paper also provides a overview of N-gram selection methods in text classification.

Gabrilovich and Markovitch (2004) improve the classification accuracy of SVM on the newsgroup data by reducing the number of irrelevant features. The relevance and irrelevance of features is determined by building the entropy profile of a feature set.

Cowie et al. (1999), Cowie et al. (2000) study how Neural Nets recognize emotions in the text data.

Devillers et al. (2003) explore the dialogue data labelling beneficial for the emotion detection.

Mullen and Collier (2004), as well as Nasukawa and Yi (2003), Nigam and Hurst (2004), Pang et al. (2002), apply ML tools to analyze sentiments in text data.

Koppel et al. (2002) classify the gender of the text authors using their style.

Riloff (1996), Riloff et al. (2005) learn syntactic patterns and then use them to classify texts, with respect to topics as diverse as terrorist attacks and private opinions.

Lin et al. (2003) apply bootstrapping for pattern extraction and semantic classification.

Yangaber and Grishman (2000) show that the language of medical reports is a restricted language.

Cristianini et al. (2001), Wang et al. (2003), Zhang and Rudnicky (2002) apply Latent Semantic Analysis to classify and model documents respectively.

Suzuki et al. (2004) use SVM for feature selection in question classification.

Aery and Sharma (2004) use graph theory for email classification.

Boparai and Kay (2002) and Scheer (2004) study how on-line conversation can be supported by email classification.

Cohen (1996) addresses rule-based approach to classify email.

Cohen et al. (2004) classify email with respect to a dialogue it supports.

Kiritchenko et al. (2004) use temporal features in email classificaiton.

Murakoshi et al. (2000) construct the structure of email communications.

Chapter 3

Background on Electronic Negotiations

Inside every large problem is a small problem struggling to get out.

Inside every small problem is a larger problem struggling to get out.

Murphy's Law

3.1 Introduction

The novelty of the application domain of negotiations and its sub-domain electronic negotiations (e-negotiations) prompts us to make an extended background search on current negotiation and e-negotiation research. In this chapter we present an overview of research which we consider relevant to our dissertation. Negotiation-specific knowledge and results obtained on behaviour in negotiations constitute a substantial part of the background. The negotiators' behaviour is strongly influenced by the roles they play and by the negotiation environment and available means. We specifically address research on e-negotiations and negotiation support systems.

In this chapter we show how the delivery of the negotiation and influence strategies is related to and reflected in the language of negotiators. The reason for this consideration lies in the language being a means of communication. On the other hand, the language is being influenced by the process which this communication depicts (Gries, 2003). Pragmatic information exchanged in communication may be associated with the language forms in which it was expressed. We look for parts-of-speech whose use corresponds to agreement, refusal, exchange of information, argumentation, persuasion and substantiation. We seek how the negotiator's indirect and direct influence on the counter-part is implemented in language. We summarize the overview in the chapter's *Conclusion.*

Chapter 4 presents our case study and shows how language trends described in this chapter can be found. We use the results of the current chapter in chapters 5, 6, 7 and 8 and in Section 9.6.

In the main sections of the chapter we cite only the fundamental references. This is done to preserve a smooth reading flow. We list annotated related work in *Bibliographical Notes* at the end of the chapter.

3.2 The communication model

(Hargie and Dickson, 2004), based on (Argyle et al., 1981) introduce a conceptual model of communication. They call it simply the communication model (CM). The model identifies six elements of interpersonal communication:

person-situation context, goal, mediating process, response, feedback, perception.

The person-situation context includes how people devise the situation, what goals they formulate, what meanings they assign to the events and how they choose and exchange the patterns of conduct. The key elements of the person-situation context include

Environment in which the communication happens,

Means by which the communication holds,

Goal at which the communication aims,

Topic of the communication discussion,

Rules governing the communication, including those conducted by the ***roles*** of the participants.

The usefulness of the person-situation context for learning from data will become clear after a motivating example. Let us consider phone conversations aiming to set a date and a venue for a dinner which involves a few friends. The first, and foremost, problem to solve is how to define the set of tasks. Next, we must find out how to solve each task, and, finally, solve them.

For example, consider *arranging a dinner party for a company* by phone. Here, the environment is *personal*, the means is *phone*, the goal is *to make an agreement* which consists of solving simultaneously several issues (e.g. where and when the party should be held, the number of courses served, etc), the topic is the *discussion of the dinner party.* Rules are settled by the socially acceptable norms of friendly discussion on phone, roles, a host and guests, might not be well defined until the venue of the dinner has been agreed on.

The key elements provide us with the list of the possible tasks, e.g., the effect of the phone conversation on the discussion, the characteristics of the successfully reached agreement, the interconnections between the issues, how the social rules are implemented and so on.

As soon as the set of tasks has been identified we are ready to establish a method capable of attending to the tasks. Generally, the method takes a multiple view of the problem. As an initial step, we evaluate the available data. In the dinner discussion example we (possibly) have: a) the records of the phone calls e.g., their time, frequency, duration and maximum, average and minimum time between calls; b) the transcripts of conversations from which the numerical data – suggested dates, number of menu options, expenses – can be extracted. All of those data have different characteristics – distribution, ranges, etc. – that have to be identified. When we finish this step, we know about the complete data and their parts. In the next step, we keep the multiple view of the data; this time we use it to model the data. Naturally, we use the data from the first step with the most promising characteristics. We apply various models to see which of them brings us closer to the task solution. For example, modelling spoken conversation includes the duration of calls, the turns in which a host and a guest have called, the number of interactions during one call and so on. The modelling phase assists in identifying the *sets of candidate features* through which the data are represented during further steps. The Information Extraction step follows the modelling. For example, we might extract information about further calls. The extracted information is used for learning. Here we have to identify which methods, DM, ML or NLP, or their combinations, are applicable to learning from this data. We consider the learning tasks identified earlier. In the final step we interpret the learning results.

We list the key elements for electronic negotiations in Section 3.4.4. In Chapters 5 and 7 we use the person-situation context of negotiations in data modelling and feature selection.

3.3 Negotiations

In this section we present a short background report on the related negotiation research. We characterize negotiation by its qualities and stages. We present the strategies employed by participants during the negotiation. We focus on negotiators' behaviour during negotiation and on the influence strategies they use to reach their goals. We also list research on human behaviour in electronic negotiations.

3.3.1 Overview of negotiations

Negotiation occurs in political, business and casual settings. Traditionally negotiations are studied by the negotiation theory which combines Decision Analysis, Game Theory and Behavioural Science. Decision Analysis provides qualitative support and prescriptions for negotiation strategies and tactics, identifies the required information and proposes solutions. Game Theory provides quantitative modelling of negotiations. Behavioural Science studies negotiation as interpersonal communication and concentrates on negotiators' attitudes, perceptions, qualitative description of the negotiation process and their outcome (Putnam and Roloff, 1992). Negotiation Theory integrates Decision Analysis and Game Theory to build formal models of negotiations meaningful to the negotiation practitioners.

When the negotiation's goal is to claim resources then it is known as *distributive* negotiation (bargaining). Thompson and Nadler (2002) claim that a positive bargaining zone must exist, i.e. there must be a range of agreements suitable for both parties for distributive bargaining to reach an agreement (compromise). In *integrative* negotiations, negotiators have different preferences for the same issue and multiple issues are discussed (Kersten, 2000). In such negotiations both parties can gain, which leads to a "win-win" negotiation. The goal of distributive negotiation is to reach an effective compromise, and the goal of the integrative negotiation is to create a compromise. Bilateral integrative negotiation is one of the most common type of negotiations although there is no strict boundary between integrative and distributive negotiation. When all issues but one are temporarily fixed, an integrative negotiation becomes a distributive one.

Hargie and Dickson (2004) state that as a special type of human communication, negotiation has the following qualities of a process: dynamic, multi-dimensional, irreversible and purposeful. We use these qualities in Chapters 6 – 8 of the dissertation.

3.3.2 Business negotiations

In this section we give a general description of the negotiation phases and strategies. Everything we say here applies to business negotiations. Dupont and Faure (2002) see negotiation as a sequence of stages:

pre-negotiation the main activities at this stage are formulating goals, identifying key issues, gathering information, defining negotiation strategies and planning for settlement;

opening the main activities are personal introduction and setting an agenda;

exploration the main activity is to examine one another's positions;

bargaining the main activities include making offers and counter-offers and persuasion of the counterpart to make concession;

settlement the main activities include trial closure, closure and documentation of agreement;

post-negotiation the main activity is the analysis of negotiation.

The opening stage of negotiation is of key importance for the development of the whole negotiation. At this stage negotiators make decisions about their counterpart, how co-operative or competitive they are going to be and whether the social relationship will benefit the negotiation.

The strategic approach to negotiations states that the outcome of negotiations is the result of the negotiators' strategic choices. There are four main strategies that may be employed by a negotiator:

1. *unilateral concession*, when one negotiator accepts demands and offers of the other;

2. *individual gain*, when one negotiator is interested only in maximizing his gains without considering the needs of the other party;

3. *competition*, which usually is connected with distributive bargaining when both parties aim to get a higher share of benefits than the other;

4. *co-operation*, which is usually connected with integrative bargaining when the goal is to achieve the best possible deal for both parties.

In language, these strategies are exhibited in agreement, refusal, questioning, answering and exchange of offers.

We explore the importance of different negotiation stages and strategies in Chapter 8.

3.3.3 Research on behaviour in negotiations

Druckman and Hopman (2002) study behaviour in negotiations by applying content analysis, which concentrates on the formal negotiations, leaving aside informal meetings. Content analysis isolates negotiation from broader, i.e. political, legal, or diplomatic, context. These constrains of content analysis correspond to the approach we use in the dissertation. Bargaining process analysis (BPA) (Walcott and Hormann, 1978) is a system used by content analysis to code specific actions during the negotiation process. We show in Table A.1 BPA's behavioural categories and corresponding actions.

During negotiation, participants express personal power by employing influence strategies, intended to make the counterpart concede. Such strategies are mainly divided into direct and indirect ones, which are expressed by various types of appeal (Brett, 2001). Direct influence strategies are used when the participant says what she wants the other party to do. In indirect influence strategies such requests are implied and often masked by asking for sympathy. The influence strategies are exhibited in such negotiation moves as **argumentation, persuasion, threats, and substantiation**, and general behaviour as questions, reactions, offers, exchange of information.

Negotiators are sensitive to the competitive or cooperative climate established by the counterpart and to strategies employed by the counterpart, and they adjust their strategies accordingly. Drake (2001) has shown that measuring information exchange within a negotiation pair, or a dyad, violates a statistical assumption of independence.

Cellich and Jain (2004), Thompson and Nadler (2002) state that the choice of the communication medium influences negotiations. In face-to-face negotiations truth-telling is higher than in other types of negotiations, increasing the likelihood of a mutually beneficial agreement. For individualistically oriented negotiators the absence of a face-to-face situation results in lower use of pressure tactics, less impasse, and achievement of higher joint profit. This corresponds to showing less personal power (Ströbel, 2000). In email communication the lack of social context cues leads to more forthright behaviour, on the edge of socially undesirable behaviour, either because of reduced evaluation anxiety or reduced attention to social norms. The positive factor of email communication is its tendency to cut across social hierarchies.

Currently research suggests that professional background, *role* requirements, negotiation issues, and the cultural orientation of the opponent are contextual features that moderate the influence of culture on the negotiator's behaviour. The most extensively documented contextual factor is a role , or sets of rights, obligations, and normative expectations attached to social positions (Rubin, 2002). In intercultural negotiations, buyer-seller roles were better predictors of work strategies than culture. Sellers may perceive greater dependence on buyers than the reverse. Sellers may perceive the need to contact with a given buyer. In contrast, buyers may perceive that if a profitable agreement is not possible with a given seller, then a number of alternative sellers are available.

We found only one book that directly lists language examples used in negotiations (Hansberger, 1985). The book is a practical guide for the French how to negotiate in English. It contains 888 examples of language patterns for conducting face-to-face negotiation. Examples are divided into seven broad groups:

Practical arrangements and ceremonial manners, Bargaining, Discussing, Deciding, Frequent subjects in a negotiation, Technical points, Commenting on negotiations

Each group is divided into subgroups; see Table A.2 adapted from (Hansberger, 1985). The groups of language examples in his book correspond to those in Brett (2001), Hargie and Dickson (2004), Perloff (2003), the main references on the negotiation strategies for our research.

Throughout the dissertation we partially address the listed problems with regards to electronic negotiations.

3.4 Electronic Negotiations

In this section we discuss the main components of electronic negotiations, similarities and differences with face-to-face negotiations found so far by other researchers. We also address the major trends in current AI research on electronic negotiations.

3.4.1 Electronic means

Here we briefly describe electronic means used in e-negotiations. Electronic means provide the environment necessary to fulfill functions specific to negotiations and are based on theories of individual decision-making, communication and negotiation. Schoop (2003) suggests that electronic means represent either quantitative automated negotiation, where decisions are made by the system, or non-automated negotiation support, where decisions are made by negotiators. The former are represented by electronic auctions and intelligent software agents, and are outside the scope of our research. The latter are represented by process-oriented systems, i.e. communication systems and negotiation support systems. Figure 3.1 shows the relations between e-negotiation means.

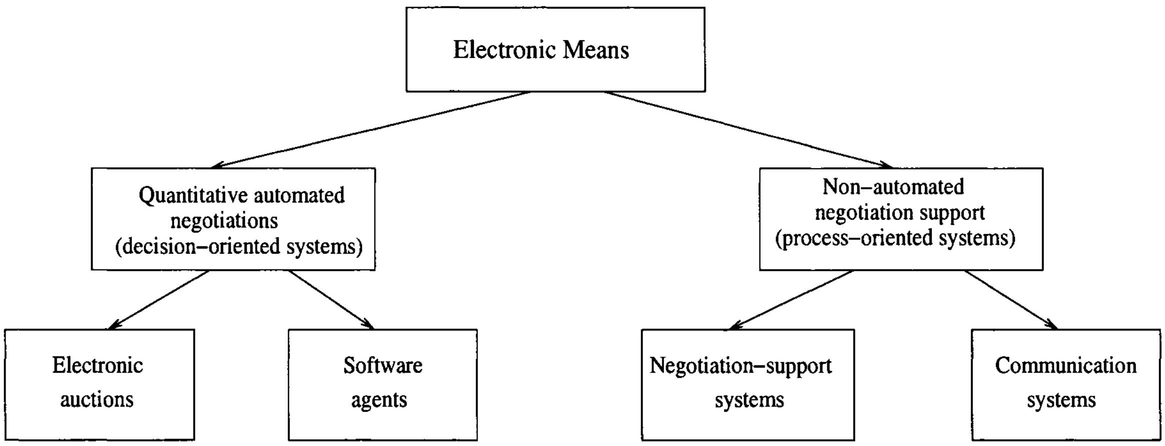

Figure 3.1: Types of electronic means in e-negotiations

Holsapple and Whinston (2000) discuss in details the decision-support functions. The decision-support functions are : decision problem formulation, decision-maker specification, offer and message construction and evaluation, counterpart analysis, what-if and sensitivity analysis, history and process analysis, knowledge seeking and use, negotiation protocols, strategics and tactics. Negotiation support systems (NSS) vary according to the negotiation phases they support, problem-solving levels present in the decision-making process, and decision-analytical methods they use; see Figure 3.2.

The negotiation methodology recognizes three negotiation phases: pre-negotiation, negotiation, and post-negotiation. There are three problem-solving levels: (i) the need and value level, at which negotiators specify their preferences, (ii) the cognitive level, at which needs are matched against opportunities and threats, and activities are chosen, (iii) the tool and calculation level, at which activities are coordinated with the visualizing, computational, and analytical support. The decision-analytical methods include mostly game theory, applied mathematics and operational research.

The communication functions of electronic means are: transport and storage of information, in-

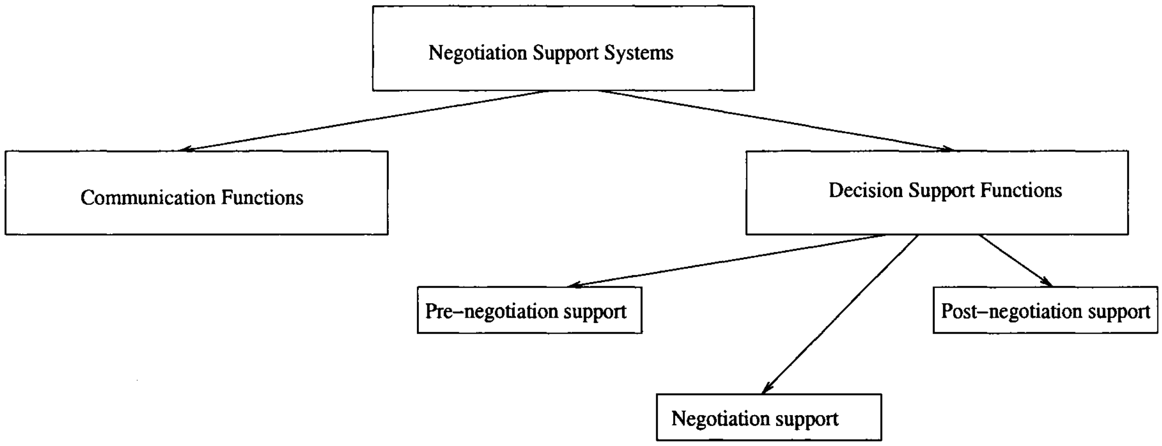

Figure 3.2: Functions of NSS in e-negotiations

formation search and retrieval, data formatting and presentation, user-system interaction. Communication systems, such as email or chat boards, supply negotiators only with means of exchange of information; for example, see Hendry (2004).

In this dissertation we work with the data provided by the communication functions of electronic means.

3.4.2 Research on e-negotiations

We looked at the studies conducted on process-oriented e-negotiation systems, including email. We found that the previous studies of e-negotiations, except Kersten and Zhang (2003) and Nastase and Kersten (2005), have staged small-scale experiments. The experiments were usually carried out by homogeneous groups of graduate students, with the number of participants seldom more than 100. The longest negotiations took less than 10 days. We conclude that no reviewed studies involve large amount of data because gathering such data requires either a substantial number of negotiators or long-term negotiations. This means that the generalization capability of such research is relatively low. Kersten and Zhang (2003) and Nastase and Kersten (2005) ran experiments with a large amount of data, but no textual data were analyzed. At the same time, many questions remain. Although the question of culture and e-negotiations was debated many times and in many studies, usually it concentrates on either the background of participants or their communication habits. The negotiator's behaviour according to assigned roles in the negotiation process does not receive the attention corresponding to their significance in e-negotiations.

Thompson and Nadler (2002) concentrate on the effect of information technology on social behaviour. The e-negotiations were conducted by MBA students from two US universities. The subject of negotiations was a purchase of company cars; it was a bilateral negotiation, negotiators did not have preliminary knowledge about their counterparts. Negotiations were conducted through simple communication channels, which did not provide any negotiation support. The duration of negotiations was limited to 10 days. Unfortunately, the authors do not provide any statistical information about the data contributors, not even their number. Nor do they talk about the existence of any available negotiation support. The key findings on e-negotiations are the following:

- multi-issue offers are a common practice,
- brief exchange of personal information on the pre-negotiation stage reduces the likelihood of a deadlock,
- negotiators attempting to build rapport engender more positive emotion and trust than do those who attempt to dominate.

The authors have found four major biases that affect e-negotiation:

1. *temporal synchrony*; e-negotiators usually discount the asynchronous aspect of e-negotiation and ignore its implications;
2. *burned bridge*; e-negotiators tend to engage in risky interpersonal behaviour and may use ultimatums that lead to a high probability of failure;
3. *squeaky wheel*; the tendency of adopting an aversive emotional style to achieve negotiation goals;
4. *sinister attribution*; the likelihood of making dispositional assumptions increases disproportionately to the rate of counternormative e-behaviour.

Key findings are that non-task related, relation-focused communication sets the stage for cooperation and trust that facilitates mutually beneficial agreements.

Morris et al. (2002) also studied the positive impact of the exchange of personal information during e-negotiations. All 106 e-negotiators were MBA students, and the e-negotiation time was limited to

7 days. The authors have applied statistical analysis to the answers on pre- and post-negotiation questionnaires.

Ströbel (2000) studied the influence of electronic means on the negotiation power. The following sources of power are used in negotiations: resource control, information power, personal power (attractiveness, emotion, integrity, persistence and tenacity). The general assumption for highly transparent markets, including electronic markets, is that both parties always have the same level of information, hence no parties can benefit from excessive information power. The sources of personal power are not present in an electronic exchange, except for texts of messages. The avoidance of threats, positional statements and other messages related to personal power promote integrative solutions.

In a small scale experiment Buddress et al. (2003) empirically contrasted email negotiations with face-to-face negotiations. The studies were conducted in a class of 72 students studying negotiations. The class was equally divided into two groups, face-to-face and email negotiators. Both groups were divided into buyer-seller dyads and given the same negotiation task. The face-to-face negotiators expressed higher level of satisfaction with both the outcome and the process, they also prepared more thoroughly than email negotiators.

The longer e-negotiation takes, the more complex the structure of the e-negotiation process becomes. Simpler e-negotiation may involve exchange of well-structured business documents (predefined contracts, retail transactions). A complex e-negotiation process comprises numerous offers and counter-offers and has a high degree of uncertainty, which "results from the instability of the process environment, and from the unpredictability regarding the dynamic behaviour of the organizational elements. The probability for changes of the situation and behaviour as well as the extent to which they occur, play a central role. Through the establishment of rules (infrastructure), uncertainty is reduced via feedback loops, since specific patterns of behaviour are prescribed for certain situations. The more difficult it is to predict future developments and situations, the higher the costs to build up a useful infrastructure tend to be." (Gebauer and Scharl, 1999).

We investigate the influence strategies and their implementation in Chapter 8. We study the roles of buyers and sellers in e-negotiations in Chapters 6 and 7.

3.4.3 Artificial Intelligence and e-negotiations

In recent years AI methods have contributed to the quantitative modelling of negotiations. This is mostly work on decision-making during negotiations; the overview of decision-making in bilateral negotiations is presented by Li et al. (2004).

Reinforcement learning is one of the ways in which AI is involved in negotiations; for example, see (Vega-Redondo, 2003). Reinforcement learning originates from control theory and models a dynamic environment. The data are state-action-reward triples. On each step the reinforcement algorithm must find an action to maximize the expected reward over time. This differs it from the supervised machine learning algorithms where the algorithm seek the optimal – with respect to the chosen criteria – solution on each step (Herbrich, 2002).

Although the AI community actively investigates the process and the outcome of e-negotiation, research concentrates on decision-oriented agent-based systems; see, for example, Namatame et al. (2002). Usually research deals with the planning of negotiations, focusing on the theoretical aspects of negotiation strategies (Bowling, 2005; Park et al., 2004) or the development of agent systems (for early, and now classical, results refer to (Shaw and Garland, 1996)).

"Real life" electronic negotiations involve many people with varying cultural and educational background. Those negotiations do not attract the attention of AI researchers. Apart from our research only two initiatives using DM addressed the question of human activities in e-negotiations. Kersten and Zhang (2003) and Nastase and Kersten (2005) applied DM techniques to the records of negotiations, working only with the records of exchange of formal offers. Among their findings, the following results on the behaviour of e-negotiators are very interesting for us: if offer exchanges are made during the early stages of the negotiation, it is more possible to reach an agreement; offers sent in the last day before the deadline reduce the probability of achieving the agreement. The Authors did not consider the language aspect of negotiations. We, on the other hand, focus on the interpersonal context of negotiations and study how this context can contribute to learning the negotiation process.

3.4.4 The key elements of e-negotiations

We identify the key elements of CM with respect to the person-situation context of electronic business negotiations.

Environment (business)

Means (electronic system)

Goal (make a purchase)

Topic (purchase of goods)

Rules , including

- imposed by roles within a negotiation (*buyer and seller*)
- protocol of negotiations (*set by a system*)
- criteria of success (*set by a system*)

As it is seen from the list, a few key elements depend on the system through which a negotiation takes place.

3.5 Characteristics of computer-mediated communication

This section presents the characteristics of interpersonal Web-based communication. The Web-based interpersonal text data is studied by such different fields as linguistics, text classification, etc. From the point of view of communication it is studied by Communication Theory under the umbrella of Computer-Mediated Communication (CMC) research.

Murray (2000), Herring (2001), Yates and Orlikowski (1993) have shown that the text data is characterized by short sentences and facilitators which are intended to make it easier for the receiver (reader) to comprehend the message. The texts exchanged through the Web tend to be syntactically informal, highly erroneous and poorly edited. Texts of messages imitate speech, using sound imitations **Hm,Uh-ha** and dots, letter repetitions **sooon**, capitalization **THANKS**. Examples found in electronic negotiation data are shown in Table 3.1.

Among the work on CMC texts, Climent et al. (2003) analyze the linguistic characteristics of the text data obtained from contributors whose background is similar to the background of the contributors of our data. The study concentrates on non-standard linguistic features of CMC texts, especially email exchanged in newsgroups. The study deals with machine translation (Catalan to Spanish) and does not apply NLP and ML techniques. There are 533 data contributors, university students mostly

Characteristics	Examples	Characteristics	Examples
Foreign words	offert, da, niet, ich	Misspelling	negociation, receve
Informal words	u, gotta, wanna	Incorrect spacing	every body, deliverytime
Speech imitation	hellooo, sooorry, hahaha	Miscapitalization	YOu, monday
Mistyping	abuot, abd	Abbreviations	pls, thnx, btw

Table 3.1: Special characteristics of Web-based communication

specializing in computer science. The authors have empirically classified the linguistic characteristics that cause translation errors into three broad areas: intentional non-standard features, unintentional non-standard features, and terminology. Statistical analysis shows that the text data contain a large amount of intentional and unintentional non-standard features and terminology specific to the domain of discourse.

According to (Climent et al., 2003), CMC users may express affective and socioemotional information using intentional misspelling, lexical surrogates for vocal segregates, grammatical markers, strategic capitalization, and visual arrangements of text characters into "emoticons". Intentional misspelling often includes the repetition of a vowel or consonant to represent the accentuation of a word or phrase for affect, as in the phrase, "sssoooooo good!". Lexical surrogates function as parenthetical metalinguistic cues, as "hmmm" might represent a paraverbal expression of thoughtfulness or "yuk yuk" might express self-deprecating laughter. Grammatical markers include gratuitous capitalization as well as repeated exclamation points and question marks to add affective emphasis. Emoticons refer to short combinations of textual characters which, if turned clockwise, resemble various facial expressions. This exactly corresponds to what we call non-standard features.

The similarity of CMC texts to both speech and written texts affects the choice of a text unit. In written texts researches agree on such well-defined units as sentence, clause or phrase, but in spoken language the usefulness of different units, and even their existence, is questionable (Miller, 1995).

We analyze CMC features in the text data of e-negotiations in Section 4.7. We define the unit of e-negotiation text data in Section 7.3.

3.6 Language and negotiation strategies

In this section we survey the studies of language used in negotiations. We connect their results with the results of linguistic analysis. Our focus is on the language implementation of influence strategies used in negotiations.

3.6.1 Expression of necessity, possibility, permission, and volition

We could not find any research on the negotiation and influence strategies and their implementation in the language of e-negotiations. Analysis provided in this section is based on the research of face-to-face negotiations, hence the term "speech" is used. To avoid unnecessary invention of new terms we borrow from Communication Theory "powerful" and "powerless" speech (Boparai and Kay, 2002), although we neither appreciate nor support these terms.

Powerful speech is consistent and direct, but its specific characteristics usually are not defined. In research powerful speech is defined as the one that does not have characteristics of powerless speech (Perloff, 2003). The characteristics of powerless speech include hesitations and noninfluences *uh, well, you know*, hedges *sort of, I guess*, tag questions *That plan will cost us too much, don't you think?*, disclaimers *I'm not an expert.* Examples show that powerless speech partially corresponds to the speech imitations in CMC (hesitations and noninfluences), the use of verbs *guess, suppose* (hedges) and not-negatives (disclaimers) and the excessive use of qualifiers. Note that hedging opinions is one of the markers of polite speech (Brown and Levinson, 1978), although "polite" is not equal to "powerless". The positive correlation between powerful speech and persuasion shows relations between the influence strategies and the language (for an overview see Burrell and Koper (1998)).

One of the indicators of argumentation is an openness to feedback from the counterpart. The mental verbs used in the positive statements *I/we think/know/consider* aim to obtain feedback from the counterpart (Perkins, 1983). Such statements suggest careful deliberation and reflective weighing. They also represent psychological being of the negotiator at the present moment.

The directness of the speech can be studied through the use of the modal auxiliary verbs. The modal auxiliary verbs, or modal , have both logical and pragmatic meanings. They express permission, possibility and necessity as the representatives of logic, and condescension, politeness, tact and irony as the representatives of practice (Leech, 1987). In the negotiation process, the use of modal verbs partially corresponds to argument, persuasion, substantiation and appeal which constitute the influence strategies.

The modals are divided into two main categories, primary and secondary (Leech, 1987); see Table 3.2. In Table 3.3 we list the most common meanings of the primary modals in positive statements.

Primary modals	can, may, must (need), will, shall, have (got) to
Secondary modals	could, might, ought to, would, should

Table 3.2: Primary and secondary modals

Meanings are listed in the order of commonality, from the most common to the least common. The secondary modals tend to be more hypothetical, polite, tentative, formal and indirect than the primary modals. The secondary modals refer to the past, while the primary modals refer to present. In general, the difference between the primary and the secondary modals can be stated thus: the secondary modals are more conditional than the primary modals. Also, the use of the secondary modals makes the statement more polite than the use of the primary modals (Brown and Levinson, 1978).

Modals	**Meaning**
can	possibility, ability, permission
may	possibility, permission, exclamatory wish
must (need)	obligation or requirement (speaker's authority), logical necessity
will	prediction, willinness, insistence, intention
shall	prediction, intention
have (got) to	obligation or requirement, logical necessity

Table 3.3: The meaning of the primary modals

Below we point out the uses of the primary modals which we consider to be important in the studies of influence strategies. In analyzing the data we seek to find out how the negotiators appeal to their counterparts, either they exercise authority openly or prefer democratic imperative. Consequently, we will be looking for the following patterns:

- When *can* is used with *we* and third-person subjects, it implies tactful imperative. Such situation occurs when the speaker does not want to exercise the authority openly and suggests that there are certain possibilities. In sentences of the form " *You can* **verb** ..." *can* denotes permission.

- *May* is mostly used as *can*, with slight emphasis on factual possibility instead of theoretical possibility as in *can*. It is interesting that *may* is used ten times less than *can* in dialogues and three times less in narrative texts (Lebrun, 1965).

- When *must* denotes obligation, this implies that the speaker has an authority over the person mentioned in the clause. *I must* and *we must* denote self-obligation, when the speaker has power over himself or herself.

- The meanings of *have to* correspond close to the ones of *must*, with more objectivity in it, as though obligation comes from an outside source.

- The use of the intentional meaning of *will* is the most interesting for the studies of implicit promises. Combined with first-person subjects, it conveys that a decision has been made and the fulfilment of the intention is guaranteed.

Modals express power differently if they appear in *if*-clauses. If-clauses have the same meaning as questions, especially if they include *any, anyone anything, ever.* In this case self-obligation, possibility of permission are given to the listener, or the receiver of the question/if-clause. We look at the use of modals with negations in section 3.6.2.

We compare the use of the secondary and the primary modals. In the use of modals we search for trends in negotiators' messages which reflect the influence on the counterparts, have explicit or implicit signs of power distribution and self-positioning. We compare the use of the modals with respect to the negotiation outcomes and with respect to the negotiators' roles. See section 8.2 for empirical results.

Verbs expressing volition of a speaker – volition verbs – are divided into several groups with respect to the speaker's intentions about the subject of discussion S and communications with the counterpart (Rudanko, 1989).

First we look at verbs which imply positive volition of a speaker/negotiator:

- Verbs meaning that a negotiator wants S to be realized or to hold: *hope, like, prefer, want, wish*, although these verbs do not necessarily mean that the negotiator intends to realize S.

- Verbs corresponding to the negotiator's intention to realize S and communicating the intention either refer to the negotiator's actions only, as *agree, offer, promise, volunteer*, or implicitly refer to a second person, as *apply, ask, claim, demand.*

- Verbs corresponding to a decision but not necessarily communicated for the realization of S are *afford, aim, choose, decide, determine, intend, look, mean, plan, propose.*

- Verbs showing the negotiator's endeavor in addition to volition and intention are *arrange, attempt, bother, help, learn, make, manage, move, proceed, push, seek, set out, speed, try.*

Next we consider verbs which express negative volition with respect to the realization of S:

- Verbs expressing communication are *decline, refuse.*

- Verbs expressing negative intention not necessarily communicated are *delay, hesitate, neglect, omit, postpone.*

In section 8.2 we look how negotiators in different classes communicate their volition, intention and reference to the counterparts.

We use modals and volition verbs as features in the classification problems in Chapter 8.

3.6.2 Expression of negation

The viewpoint of a negotiator can be expressed in positive and negative ways. Among the indicators of the negative viewpoint are so-called *fuzzy* and *formally and semantically* negative expressions, or negatives, among implicit indicators of negative context are *any* and *ever* ; see Table 3.4. We assume that the negatives express rejection including refusals and denial, explicit and implicit.

Implicit negatives	**Fuzzy negatives**	**Formally and semantically negatives**		
		Not-negation	No-negation	Affixal
any	few,	not,	never,	a-
ever	hardly,	-n't	neither,	dis-
little,		nobody,	in-	
rarely,		no, none,	non-	
seldom		nor,	un-	
		nothing	-less	
		nowhere	-out	

Table 3.4: Implicit, fuzzy, formally and semantically negatives.

Tottie (1991) gives an overview of the research on negatives which can be summarized as follows: many negatives co-occur with mental verbs, e.g., know, think, want, mean, suppose, like, consider, understand, etc, i.e. verbs indicating mental processes, and with modals. The higher use of mental verbs in implicit denials, as in *I don't think I'll ...*, in spoken language than the written language is attributed to the "face-saving" version of negation, which helps continue communication. Another type of implicit denial is a statement functioning as a question, as in *I don't know if you ....* We want to see how these indicators of cooperativeness are connected with the outcomes of negotiations.

The negatives are used differently in written language and speech, with the average frequency 27.6 per 1000 words in spoken language and the average frequency 12.8 per 1000 words in written language. The quantitative results were calculated on the Survey of English Usage corpus for the spoken language and on the Lancaster corpus of British English for the written language. One of the important differences between the written and spoken language comes from the amount of direct interactions, which are highly present in spoken communications and low in written communications. In case of written communications the sender only imagines a receiver.

The affixal negations are much more frequent in written than in spoken language, 4.16 per 1000 words vs. 2.23 per 1000 words. The affixal negatives amount to 8% of all negatives in the spoken texts, and 33% of all negatives in the written texts. One of the differences in the use of affixal negatives lies in the distinction between prepared and unprepared discourse, which directly corresponds to the amount of direct interactions. Another difference comes from the denser filling of ideas per interaction in the written communications than in the spoken communications.

Together no-negation and not-negation are used more frequently in written than in spoken language, although the not-negation is used more frequently in spoken language and no-negation in written language. One of the differences in the use of no-negation and not-negation comes from specific and general, as in *because we believe in no wars* vs. *because we don't believe in wars.* Another difference comes from the difference between a simple statement of fact, as in *Jack is not a student,* with an emotional statement, as in *Jack is no student.* In general, not-sentences are more vague and variable with the focus of negation than no-sentences. In spoken language no-negations sum up a previous discussion, either in the form of an existential sentence, i.e. *there is no need,* or with other indicators. In both written and spoken languages not-negation corresponds to explicit negation and no-negation corresponds to implicit negation. It is interesting that in both written and spoken language pronouns use used with no-negations twice more than nouns.

Modals with negations represent two groups: modal negation and main verb negation (Leech, 1987). Cannot, may not, have not in all senses are examples of the modal negation, meaning no permission, no ability, no necessity. "Will not *Verb*" is an example of the main verb negation.

In our study of rejection and decline patterns we investigate two aspects of their use in the e-negotiation data. We investigate which commonalities the e-negotiation data exhibit with spoken language and which features are similar to written language. On the other hand we investigate how classes of negotiators, different with respect to the outcomes and the roles, use expressions of rejection and decline; see Section 8.3 for empirical results.

We use negations as features in the classification problems in Chapter 8.

3.6.3 Expression of comparison and classification of events

Events are identified by nouns. The adjective is an attribution of a quality, which it describes, of the noun it modifies (Warren, 1984). A negotiator uses adjectives to evaluate an event and to manipulate this evaluation in a way that suits him the most (Mulholland, 1991). Negotiators use gradability of adjectives to compare events, e.g., *faster* delivery, *fastest* delivery, *more important* issue, *most important* issue. Adjectives can be intensified *very cheap, extremely cheap*, thus the evaluation of the event is also intensified *very good offer*, or attenuated *not very important, less urgent*, leading to the attenuation of the event *less expensive model* (Watson, 1994). Gradability of adjectives corresponds to common knowledge that certain qualities, states, effects vary in intensity, amount and number.

The identifying adjectives are *actual, certain, favourite, only, particular, principal, sole, specific, very* and the superlative adjectives. A speaker uses identifying adjectives when they, supposedly, indicate the sole or most prominent feature of the event, and this fact is assumed to be evident to the listener. Phrases with identifying adjectives tend to be definite. In the comparative form *better, more cheap* adjectives initiate the pull effect (Rusiecki, 1985): *better (than)* and *more cheap (than)* pull towards *good* and *cheap.*

We use the material cited in this section to explore the connection between the influence strategies and the negotiation outcomes. Our results are reported in Chapters 6 and 8.

3.7 Conclusion

After reviewing the work done on negotiations, including e-negotiations, we have concluded that:

1. Although AI intensively studies e-negotiations, these studies are mostly applied to agent-based negotiations and focus on decision-making during negotiations; no ML or NLP studies have considered e-negotiations as communication between participants;

2. All studies of behaviour in e-negotiations were either small-scale or did not use text data.

3. No models of e-negotiation text data were built using NLP or statistical methods; no learning of text data was done using ML methods.

4. No studies were performed on how language used in e-negotiations relates to the model of negotiations, i.e. the language and the outcome of e-negotiations, the language and the roles, the language and the NSS system, etc.

In this chapter we have explored how the negotiation and influence strategies are reflected in language. We have shown that the implementation of strategies results in, but is not restricted to, the use of modal verbs, mental verbs, volition verbs, negatives and adjectives.

We address these problems in the remainder of our dissertation, by either obtaining results or listing suggestions for future work.

3.8 Bibliographical Notes

Argyle et al. (1981) propose the elements of the communication situation.

Hargie (1997), Hargie et al. (1999) employ the elements to build a communication model.

Bazerman et al. (2000), Carnevale and Pruitt (1992), Hargie et al. (1999) explore behaviour in negotiations including negotiation strategies.

Walcott and Hormann (1978) analyze bargaining process and introduce Bargaining Process Analysis (BPA), a technique widely used in the negotiation theory.

Simons (1993) has found that language patterns of the first part of negotiation efficiently predict the negotiation outcome.

Drake (2001), Rubin (2002), Sjostedt (2003) study the importance of roles (e.g. buyers and sellers) in negotiations.

Sjostedt (2003) shows how roles influence the choice of strategies during negotiations.

Sutter et al. (2003), Watson (1994) present gender studies in negotiations.

Druckman and Hopman (2002) introduce content analysis built on BPA.

Dupont and Faure (2002), Mulholland (1991) consider the process and stages of negotiations.

Faure (2002), Salacuse (1998), Smith (2000) discuss culture and negotiations.

Hansberger (1985), Hargie (1997), Mulholland (1991) suggest practical guidance in improving communication and negotiation skills.

Brown and Levinson (1978) study politeness in negotiations.

Buddress et al. (2003), Ströbel (2000) study the effectiveness of e-negotiations.

Kersten (2000), Ströbel (2000) consider integrative and distributive e-negotiations, including negotiation strategies exploited during e-negotiations.

Morris et al. (2002) explore the importance of socializing before and during negotiations.

Schoop (2003) suggests the approach based on the speech acts helps in conducting e-negotiations.

Zeleznikow (2002), Bakos (1998), Chu-Carroll and Carberry (2000) provide examples of negotiation support systems that work in such domains as legal, economic and dialogue systems respectively.

Vetschera et al. (2004), Smith (2000) study culture in e-negotiations concentrating on collectivism vs individualism and high context vs low context characteristics of the society that the negotiators belong to.

Hu (2003), Gebauer and Scharl (1999) investigate the increasing volume of on-line communication in routine business.

Climent et al. (2003) explore Newsgroups messages exchanged by university students from the perspective of machine translation.

Carey (1980), Herring (2001), Murray (2000), Walther (1992), Yates and Orlikowski (1993) investigate the communication aspects of the messages exchanged through Web emphasizing the special features. Those features originate from economizing on typing effort, mimicking spoken language features and such simplifications as short sentences, excessive use of abbreviations, omitting articles and non-standard spelling.

Fais and Ogura (2001) apply discourse analysis to the machine translation of e-mail.

Crowston and Kammerer (1998), Pujolar (2000),Yates (1997), Mitra et al. (2005) focus on how gender among other factors affects computer-mediated communications.

Angell and Heslop (1994) give practical advice on email writing style.

Chapter 4

Data Analysis

Whenever you set to do something, something else must be done first.

Murphy's Law

4.1 Introduction

In this chapter we introduce e-negotiation data. We define data classes based on the characteristics of negotiations we aim to learn. Next, we perform a detailed analysis of e-negotiation data. From the complete e-negotiation data we extract the data gathered from different sources. We analyze each type of data with respect to its predictive and learning abilities.

We justify the importance of the text data compared with the numerical data. The data noise is analyzed. We perform corpus analysis and show that the vocabulary used in e-negotiation text messages grows as the vocabulary of unrestricted languages.

The analysis is performed on the e-negotiation data and on face-to-face negotiation data. The results show that both corpora are closer to English spoken dialog data than to standard written and business English.

The results of this chapter assist us in selection of the modelling techniques which we use in Chapter 5 to build statistical and semantic models of the data. We partially use the analysis results to find language patterns in Chapter 6 and support hypotheses in Section 9.3.

4.2 The data of e-negotiations

In this section we present data gathered by two e-negotiation systems, *Inspire* and *SimpleNS*. We introduce the systems, the scope of their use and discuss in detail negotiation support that they provide. We also give the examples of the data.

4.2.1 The Inspire system and its data

The largest data gathered in e-negotiation have been collected by the NSS *Inspire* (Kersten and others, 2002). We have accessed 2550+ negotiations. *Inspire* is a research and teaching tool mainly used in college and university programs in a number of countries. Also, *Inspire* is available in the public domain from the Internet. No restrictions are imposed on possible users. *Inspire* supports users with a Web-based medium for conducting negotiations, gives access to on-line manuals, provides automatic evaluation of the negotiation process, and keeps the history of each negotiation. All negotiators should log in to the system to conduct negotiations. Negotiators use nicknames to log in to *Inspire*. Negotiation starts after both negotiators have filled pre-negotiation questionnaires. During negotiation negotiators exchange formal numerical offers and may exchange freely written text messages.

Inspire supports all three phases of negotiations, providing preference assessment in the pre-negotiation phase, analysis of alternative offers, offer and message exchange, counter-offer evaluation in the negotiation phase, and assessment of the efficiency of the compromise (Pareto-optimality) in the post-settlement phase. All offers are monitored and evaluated. Built-in visualization tools are available to view the history and process of negotiations. Its main decision-analytical tool is the negotiator's utility function. The construction of the function is based on hybrid conjoint measurement and discrete optimization (Kersten and Noronha, 1999). Briefly, the utility function is calculated for each negotiator based on the preferences for negotiated issues, the preferred options for those issues and the overall evaluation of the preferable package.

The negotiations conducted through *Inspire* have the following fixed parameters:

- the negotiation topic is the purchase of bicycle parts,
- the issues to negotiate are: price, delivery time, payment time, and the conditions of the return of the spoiled parts
- for each issue there are several numerical values which the negotiators must use,

- each negotiation is bilateral,
- the negotiators send to each other standard offers (tables with numerical values), on forms supplied by *Inspire*,
- the virtual purchase should happen within three weeks,
- negotiation is successful if a virtual purchase has happened within the designated time, and negotiation is unsuccessful otherwise.

The following activities are optional for the *Inspire* negotiators: exchange freely-form messages that either accompany offers or are exchanged between offers, and answer post-negotiation questionnaires.

Working with the *Inspire* data, Kersten and Zhang (2003) used data mining to classify 1525 negotiations as success or failure based on various factors including the characteristics of the negotiations. Each negotiation was represented by the number of offers sent, regularity with which offers were sent, time when the offers were sent, with special attention paid to the time of the last offer, and so on. Their results are presented in Table 4.1. The Precision, Recall and F-measure values are not available. Before classifying the outcomes Nastase and Kersten (2005) have pre-processed the Inspire data, filtering out the negotiations for which the system was unable to calculate the utility function. To represent the data they used four different sets of features extracted from the negotiation records; see Table 4.2 for the best accuracy result and corresponding recall. The results show the very high positive recall and the very low negative recall.

Classifier	**Accuracy** (%)
Neural Networks	59.28
Loglinear Regression	62.4
Decision Trees	75.33

Table 4.1: Best accuracy results from non-textual classification of 1525 negotiations

Classifier	**Acc** (%)	$\mathbf{R}_{pos}$	$\mathbf{R}_{neg}$
C5.0	76.4	95	20

Table 4.2: Best accuracy results from non-textual classification of 2004 negotiations

From the data description follows that the *Inspire* data originate from populations with different distribution. The large number of e-negotiations guarantees that the protocol data are approximated by normal distribution while the text data cannot be normally approximated (Manning and Schutze, 1999). The distribution of bargaining data depends on the problem at hand but can most closely be approximated either by binomial or by uniform distribution (Duda et al., 2000). The restricted range and fixed values of the bargaining data cannot guarantee normal approximation.

For example, consider bargaining data coming from the *Inspire* system where the negotiators issue standard formal offers using the mechanisms supplied by *Inspire*. They try to reach an agreement on price, delivery time, payment time and return conditions on the spoiled details. Each offer should contain all four issues. The system supplies a very limited choice, i.e. 3 – 4 fixed values, for the issue values that negotiators should use in their negotiations. For example, the price must be one of the following: *4.37, 4.12, 3.71, 3.47*. Obviously, a small number of pre-determined values suggests a small variety of choices for negotiators. In real life the negotiators operate within established and well-known bounds of numbers given either by market or government regulations.

The protocol data usually contain the history records of negotiations, the negotiators' personal information and the data of the operations of a system. All of the listed types are system-dependent and largely vary for different negotiation support types; see Section 1.4.

Note that we should not use in ML experiments the utility function values calculated by *Inspire*. *Inspire* calculates the utility function value using different parameters for each negotiation. Moreover, the parameters may change during the same negotiation. This means that the same value might mean different things in different negotiations and in different negotiation stages, and different values might mean the same thing in different negotiations and in different stages of the same negotiation. Thus, the values cannot be compared between negotiations or even within the same negotiation. Nor are any rules given to assign the pre-defined categories to the utility values. Hence, the main rule of Machine Learning – the data must be either ordered or categorical – is violated. We continue data analysis in the following sections of this chapter.

4.2.2 Examples of the Inspire text data

Examples of text messages extracted from the negotiation scripts (original spelling was preserved):

- Successful negotiation (all messages, chronological order).

  (Buyer) Dear Claudi, I'm anles, representing Cypress Cycles, a well-known manufacturer of bicycles based in Ford McMurry/Canada. We are seeking for new suppliers for rear wheel gear assemblies. As visiting your plant and examining your products we are interested in cooperating with you. Therefore we prepared a proposal with good perspectives for both sides. I hope that you will be able to accept it. cu anles

(Seller) Dear Anles thank you very much for your interest in our company. We are very looking forward to deal with you. Unfortunately I was not very happy about your recommended price, because I think you have to bear in mind our costs. My financial manager determinded that the price should not go below 3.75 so I actually hope very much that you also will be able to accept my suggestion of 3.98. All other items are very realistic and I think our company will not have problems to keep to them. Bye Claudi (By the way, which country are you from?)

(Seller) Hi Anles, I have just sent a counter-offer to you. It wasnt such easy, as I thougt cause it seemed I made my ratings wrong *g*. Well, now I already asked you, where you are from, cause I did not know that I would have the opportunity to contact you again. I am from Germany. Then, good luck with my offer, I am waiting for your answer. Bye Claudi

(Buyer) hi claudi, thank you very much for your offer. I think, the price is acceptable. I totally agree with you. Having informed at a trade fair in Frankfurt/Germany about metal components and comparing some prices and offers from other suppliers all around the world, I came to the conclusion that your offer is the best. It was a pleasure doing business with you. I'll give you a ring this week for more details. Best regards anles (wir wren jetzt wohl schon am ende unserer negotiation. leider war es nicht lang, da ich schon jetzt eine ziemlich hohe punktzahl erreicht habe. du vielleicht auch... ich komme brigens auch aus deutschland. lustig, oder? woher kommst du denn genau? wenn dir das geschreibe ber interneg zu langwierig ist, kannst du mich ja auch per mail erreichen: [...] ok. wrde mich freuen. cu anles (Anja)

- Unsuccessful negotiation (all messages, chronological order).

  (Buyer) Hi! We have just prepared an offer to purchase the rear wheel assemblies you produce in your plant. From our point of view, it is a good offer to accept.

  (Buyer) Please make a counter-offer, so that we can proceed a little bit further in our negotiation.

  (Seller) would this meet your expectations ?

  (Seller) sorry, but i had troubles to connect with inspire for several days....

  (Seller) its your turn now, dude :)

  (Buyer) Ok, this is my very last offer! This might be quite acceptable!

Issue	**Values**
Price	\$3.47, \$3.71, \$3.98, \$4.12, \$4.37
Delivery	20 days, 30 days, 45 days, 60 days
Payment	upon delivery, 30 days after delivery, 60 days after delivery
Returns	full price, 75 % with 5% spoilage allowed, 75 % with 10% spoilage allowed

Table 4.3: Values of negotiated issues in the *Inspire*-based negotiations

4.2.3 Examples of the Inspire numerical data

There are 180 different formal offers, in which the values of price, delivery, payment and conditions of returns values should be filled every time an offer is sent. Possible values are provided by the system and are shown in Table 4.3. At every log in to the *Inspire* system, a participant can send only one offer.

We present in Figure 4.1 typical buyer's and seller's offers and a possible final offer.

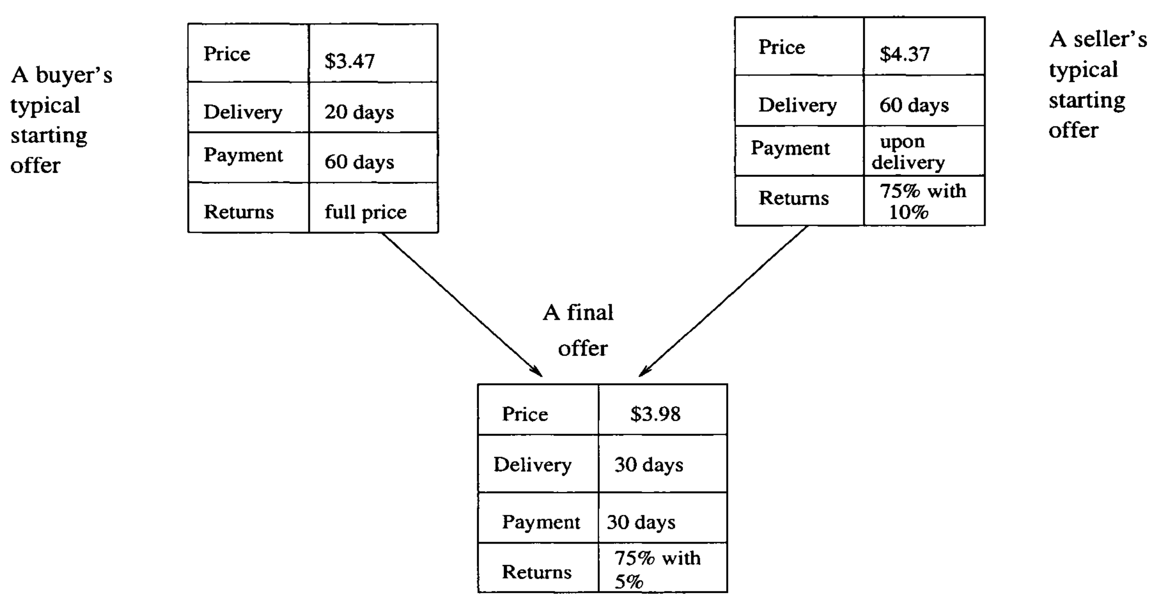

Figure 4.1: Examples of offers in *Inspire* negotiations

4.2.4 The SimpleNS system and its data

SimpleNS (Sim, 2002) represents *passive* e-negotiation systems. *SimpleNS* is a teaching tool used in university programs in a number of countries. *SimpleNS* negotiations were conducted as class assignments, thus allowing external supervision.

The system provides only a communication medium for the users. Neither negotiation support nor decision support are available; see Figure 1.1. As with *Inspire*, all negotiators should log in to the system to conduct negotiations. During a negotiation, negotiators exchange only freely written text messages . The negotiations conducted in *SimpleNS* have the following fixed parameters:

- the users negotiate one of following five topics:
  1. agreement among buyers on the purchase of crops;
  2. purchase of a car;
  3. signing a contract;
  4. signing a union agreement;
  5. purchase of bicycle parts;
- for each topic there are a few negotiable issues;
- for each issue there are several numerical values which the negotiators must use;
- each negotiation is bilateral;
- for each topic, the negotiators must assume fixed fake identities;
- the negotiators send to each other freely-form messages.

SimpleNS does not have the criteria of successful negotiations. Negotiations conducted through *SimpleNS* are short-term and thus do not provide enough grounds for developing negotiation strategies. This means that the *SimpleNS* data cannot be used to study the dependence between the language and negotiation outcomes. Nor can the data be used to study the negotiation strategies with respect to the negotiation outcomes.

However, the major disadvantage of *SimpleNS* data comes from the multiple participation of the system users in negotiations. Multiple participation reduces the number of data suppliers, thus restricting the generality of the data. In fact, the *SimpleNS* data available to us contain the transcripts of negotiations conducted by 20 students.

The obvious conclusion follows that the *SimpleNS* data is simpler and poorer than the *Inspire* data. Further in Chapters 5 and 6, where applicable, we use *SimpleNS* data to support hypotheses that we formulate about the *Inspire* data.

4.2.5 Example of SimpleNS data

The following example keeps the initial spelling and grammar.

Buyer1 Dear Dr. Sang, my name is Dr. Jam Rudec, and I am a representative of BioNet Inc. I would like to discuss with you the norocado crop in order to decide how to present a joint position in Magatawn. I am confident that we will find a solution which will satisfy both of us. Regards, Jam Rudec

Buyer2 Dear Dr. Rudec, I am glad that you have contacted me. As you know we are also interested in the norocado crop and feel that if we combine our effects, we can find a solution that meets are needs. Is there a specific target you were looking at purchasing? Regards, Dr. Oman Sang

Buyer1 Dear Dr. Sang, my company wants to buy the whole norocado crop. However, I am willing to elaborate an integrative solution with you. How many fruits, respectively which parts of them do you need? Regards, Jam Rudec

Buyer2 Dear Dr. Rudec I believe my organization needs the whole crop (just as you do) and I believe the total available supply is 20000. Unlike your company, UNFAA just needs one crop to provide sustainability to a region. We plan to use the crop to provide a better environment and stability in a less developed area of the world. Once this initial crop has been re-fertilized and is stable we are not looking to purchase any more norocado. As such I have provided two alternatives:

1) UNFAA purchases most of the supply of the norocado crop (75%) leaving the rest (25%) for your company and UNFAA agrees to sell the future crop production from this crop to your company at a set price (far lower then market value) for a set of number of years. In addition, UNFAA agrees to promote your organization for providing assistance to a less developed area of the world.

OR

2) We split the crop 50/50 with no future agreements on selling to your company in the future.

Please rely with a response as soon as possible. Regards, Oman Sang

Buyer1 Dear Dr. Sang, thank you for your response. I think you provided two fair alternatives, and I would be glad to work together with you, as you proposed in the first option. Yet I would like to clarify two issues: 1) Do you already have positive results in re-fertilizing the norocado fruit? As far as I am informed, any attempt to grow the plant elsewhere than in Magata has failed. 2) Do you need the whole fruit? For us, only the pulp is important; if you need the core, maybe we can come to an even better agreement. I am looking forward to your answer. Regards, Jam Rudec

Buyer2 Dear Jam Rudec, We have conducted experiments with the Norocado pits mixed with soil and they were very successful. We need the pulverized Norocado pits for plantation. Please revised the agreement as soon as possible. Regards, Dr. Sang

Buyer1 Dear Dr. Sang, it seems that we are getting close to an agreement. As you need the pits and we the pulp, I would propose the following agreement: We both buy the whole crop, splitting the costs 50/50. Then we separate the pits and the pulp from the fruits and allocate them to each of us. This way, we both get the maximum out of this years crop. Furthermore, I think that we can combine this with your offer to sell your future crop production to us. This cooperation surely would be profitable for both of us. I hope that you will find this offer attractive. Regards, Jam Rudec

Buyer2 Dear Rudec, I agree with your last offer as it provides the best alternative for both parties. It was going doing business with you Dr. Sang

The data presented in this section is used in the experiments throughout the remainder of the dissertation.

4.3 Profile of data contributors

In this section we show that the contributors of the data vary in their background. Unfortunately, the *SimpleNS* users, although coming from different backgrounds, do not represent a reliable population of data contributors: there are only 20 of them participating in multiple negotiations. Below we describe the background of the *Inspire* users.

In this work we use a record of 2557 negotiations. Each negotiator participates in only one negotiation; so there are more than 5000 negotiators. We have extracted background information from pre-negotiation questionnaires. Table 4.4 shows how much variety there is in the *Inspire* negotiators' background. 3125 negotiators identified their first language, 4276 their occupation.

First language	%	**Occupation**	%
English	28.1	students	82.8
German	22.8	professionals	13.1
Chinese languages	12.1	managers	1.8
Spanish	9.7	engineers	1.1
Hindi	4.6	teachers	0.6
Russian	3.8	professors	0.4
Finnish	3.4	executives	0.2
Others	15.5		

Table 4.4: Background of *Inspire* negotiators

While the negotiators' educational and cultural background differ, they have certain common characteristics: they are either university students or professionals, they all speak English (albeit many of them as a second language) and they were all given the same manuals and instructions about the negotiation process.

We will show in the following sections that despite this variety in mother tongue and occupation, the language used in *Inspire* messages is consistent across negotiations.

The messages sent by the *Inspire* users contain many misspelled words. We apply *ispell*, an off-the-shelf spell checker, in automatic correction of the most frequent negotiation-related words such as **negotiation, delivery, receive, agreement**. We correct only targeted words to avoid the introduction of additional errors in the data through false corrections. We use corrected data for experiments only when correction is necessary for obtaining reliable results. For example, part-of-speech tagging and text parsing were performed on the corrected data, but all statistical analysis was performed on uncorrected data.

The presented in the section information about the data contributors supports the observation by Schoop (2003) and Cellich and Jain (2004). They write that participants of electronic negotiations usually represent various professional, educational and cultural background. However, neither work provides any quantitative data to support the claims.

4.4 Learning tasks and data classes

As part of our work on supervised learning performed on the data, we consider different approaches to introducing learning tasks. In the dissertation we suggest tasks depending on the key elements of negotiations reported in section 3.4.4.

The most important element of negotiations is its goal. The ability to reach the goal or to fail it is shown through the negotiation outcomes. Thus, learning the negotiation outcomes becomes our major task. The only labelled data gathered through the *Inspire* system suggests two negotiation classes depending on the outcomes: successful and failed. We have found that such labels do not correspond to the outcomes of negotiations and renamed them as *successful* and *unsuccessful* .

We discuss the importance of roles in negotiations in Section 3.3. Learning of roles constitutes our next task. The topic of negotiations – purchase of goods – dictates that there be classes of buyers and sellers. The number of participants – two – suggests that there is one buyer and one seller in each negotiation and hence only these two roles are assigned to the negotiation participants. Note that if the number of participants is more than two then they can have roles different from "buyer" and "seller", such as a facilitator or the chairman of negotiations or the head of the seller group (Rubin, 2002).

From non-textual information we extract negotiation labels (*successful* and *unsuccessful*) that were given by the *Inspire* system when we analyze the text data with respect to the negotiation outcome. Note that some labels are questionable. For example, in one negotiations the last message of negotiations accepts an offer (Fig 4.2). Nevertheless, the negotiation is labelled by the system as unsuccessful. There are 3% of such untrustworthy labels in unsuccessful negotiations, that corrupt the learning process. Sometimes negotiators complain about *Inspire* but this type of noise is difficult to detect; see Figure 4.3. Among successful negotiations we find that some negotiators accept an offer because of external conditions, for example, lack of time; see Figure 4.4.

> *Hi Saba: I have decided to accept your offer because I have a good mood today. How about you? Anyway, thank for corporation and I hope both of us can learn much in this negotiation. Thanks a lot................ Ray*

Figure 4.2: False unsuccessful negotiations.

We use the role labels "buyer" and "seller" when we analyze the text data with respect to the negotiation roles. We also split buyers and sellers classes into two classes depending on the negotiation

Hello MB, I was very pleased with the first offer you sent. However, the inspire doesn't allow me to accept it. Please accept my offer. (It's the same one) Nice negotiating with you! Marwin

Figure 4.3: Complaints.

Hi Inge I have accepted your last offer, because I dont have any more time to spend on that game. good luck!

Figure 4.4: Forcing the end of negotiation.

outcome: successful and unsuccessful. The data classes listed above are used in machine learning and statistical experiments in the remainder of the dissertation.

4.5 Vocabulary analysis of the text data

In this section we explore vocabulary richness of the data. We employ corpus analysis methods. We report the results of a comparison with the standard data sets.

4.5.1 Vocabulary richness of the Inspire and SimpleNS data

We evaluate vocabulary richness using the standard corpus analysis measures; see Section 2.2.1. Concatenating messages exchanged in 2557 *Inspire* negotiation results in a collection with 1,514,623 word tokens and 25973 word types (Lyons, 1995). Concatenating messages exchanged in 83 *SimpleNS* negotiation results in a collection with 179,458 word tokens and 6532 word types. For both data sets, capitalization is unified to lower case. We do unification partially to eliminate the miscapitalization effect and partially to allow a fair comparison with other data which are often unified in lower case. We expect the vocabulary of *SimpleNS* to be more restricted than the vocabulary of *Inspire* due to the substantially smaller number of participants and the compulsory use of fictional personalities.

We attribute the *Inspire*'s lower type-token ratio to the fixed topic of discussion as opposed to the *SimpleNS* where there are five different topics of discussion. However, in both cases the type-token ratio is lower than in the standard data. For example, the Brown corpus has approximately 1,000,000 word tokens and 53,000 word types (Francis and Kucera, 1979). We explain the lower than standard type-token ratio by the origin of the data in a closed domain, namely negotiations.

Inspire and *SimpleNS* users discuss a fixed topic, negotiate a fixed problem, and choose among fixed options. We want to know how these conditions affect the growth of the vocabulary and its

N	40000	80000	96940	293152	364306	535244	614428	716176	809584	1139475	1514623
$TT(N)$	0.091	0.070	0.060	0.040	0.033	0.029	0.027	0.026	0.025	0.023	0.019
$P(N)$	0. 046	0.035	0.027	0.018	0.015	0.013	0.012	0.011	0.011	0.011	0.0084
$S(N)$	0.013	0.009	0.008	0.005	0.004	0.0037	0.0035	0.0033	0.0031	0.003	0.0023

Table 4.5: The vocabulary growth of the *Inspire* data.

convergence with respect to the sample size (Oakes, 1998). We need this information in order to choose a language model. Sometimes domain and purpose make the language restricted, e.g. the language of weather reports or the language of medical reports (Yangaber and Grishman, 2000). Such restricted languages, or *sublanguages*, can be characterized by a limited set of rules and relations between domain-specific semantic classes. Solsona et al. (2002) show that human-system interaction through a dialogue system produces restricted languages. Mast et al. (1996), Solsona et al. (2002) use the finite state grammar (FSG) as the language model of choice for human-system interaction. This model was applied to human dialogues through the VERBMOBILE system, although we could not find any results on whether the participants communicated through a restricted language or an unrestricted language. If the latter is the case, then FSG is not a good choice of the model.

We hypothesize that the *Inspire* vocabulary grows as the vocabulary of an unrestricted language. To prove it, we have investigated the growth of the vocabulary and its convergence with respect to the sample size. We calculate type-token ratio $TT(N)$ where N is the number of tokens, the vocabulary growth $P(N)$ and Sichel's characteristic $S(N)$; refer to Section 2.2.1 for formulae and explanations. As shown in Table 4.5 and Figure 4.5, new words are steadily being added at every stage. The vocabulary grows approximately at the same rate throughout the data, and Sichel's characteristic converges as expected (Tweedie and Baayen, 1998).

The results in Table 4.6 show that the same conclusions apply to the *SimpleNS* data. As predicted, the language used in the *SimpleNS* data is more restricted than the one in the *Inspire* data.

N	45387	72550	115092	165179
$TT(N)$	3155	3908	5359	6452
$P(N)$	0.024	0.019	0.018	0.014
$S(N)$	0.013	0.009	0.006	0.005

Table 4.6: The vocabulary growth of the *SimpleNS* data.

The unrestricted vocabulary growth suggests us to choose statistical modelling for the data. We

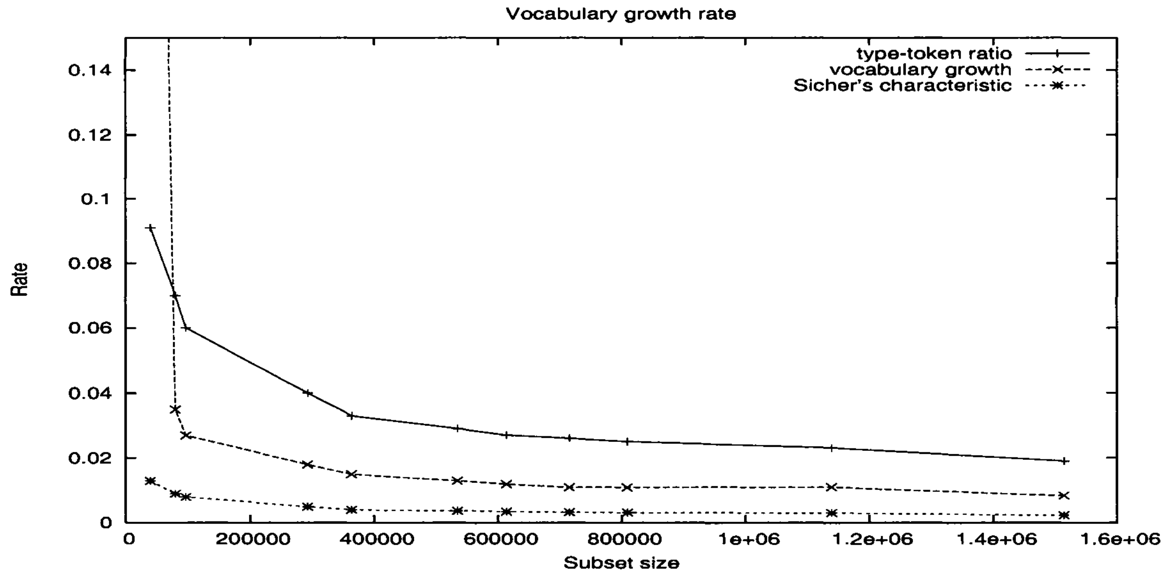

Figure 4.5: Characteristics of the vocabulary growth

build statistical models in Section 5.2.

4.5.2 Part-of-Speech tagging of the data

In this section we report results of the POS analysis of the *Inspire* data. We used the Brill tagger; see section 2.2.3 for its details. To make a more detailed comparison, we split each of the two sets of data into two subsets, according to the negotiation outcome (successful or unsuccessful negotiation) and delete negotiations where only one participant was active. This results in four classes: buyers-successful (BSN) with messages from 1424 negotiations and 544961 tokens; buyers-unsuccessful (BUN) with 689 entries and 209025 tokens; sellers-successful (SSN) with 1426 entries and 525049 tokens;, and sellers-unsuccessful (SUN) with 688 entries and 205524 tokens.

In Tables 4.7 and 4.8 we compare the percentage of text covered by the main POS: nouns (NN, including subcategories), verbs (VB, including all subcategories), adjectives (JJ), adverbs (RB), prepositions (IN), personal pronouns (PRP), and determiners (DT). We will employ ANOVA to analyze the numerical results obtained in this and the following sections and investigate how significant the difference between the corpora is. Until the analysis of the significance is done, we only categorize comparison: "+" means that the POS is used at least 10% more frequently in this class than in the opposite one, "-" means that the POS is used 10% less frequently, and "=" means that the frequencies

POC	BSN	SSN	BUN	SUN
NN	=	=	=	=
VB	=	=	=	=
JJ	=	=	=	=
RB	+	-	+	-
IN	+	-	=	=
PRP	=	=	=	=
DT	=	=	=	=

Table 4.7: Coverage of text by the POS, successful buyers/sellers, unsuccessful buyers/sellers

POC	BSN	BUN	SSN	SUN
NN	=	=	=	=
VB	=	=	=	=
JJ	=	=	=	=
RB	=	=	=	=
IN	=	=	-	+
PRP	=	=	=	=
DT	=	=	=	=

Table 4.8: Coverage of text by the POS, successful/unsuccessful buyers, successful/unsuccessful sellers

differ by less than 10% in both classes.

The POS are distributed evenly in all four classes, thus this information cannot give us insight into differences between four corpora.

In sections 3.6.1, 3.6.2, 3.6.3 we have shown that modals, verbs, comparative and superlative adjectives, and negations are of special interest in studying the language of negotiations because of their connections with the negotiation and influence strategies. We compare the use of Wh-determiners to investigate partially how conditional statements – implicit questions in written communications – are used in different classes. We also are interested in the distribution of cardinal numbers and coordinate conjunctions across classes because their use shows how different classes of negotiators approach the topic of negotiations – the purchase of bicycle parts.

As we said in Section 4.7 e-negotiation data are highly affected by the presence of the CMC characteristics (Herring, 2001; Murray, 2000). Among those, speech imitations can be traced through parsing. The use of speech imitations can be detected through interjections which appear in spontaneous speech as the markers of private thoughts and emotions (Schourup, 1985). These characteristics are also connected with "powerless"[1] speech and, thus, with the effectiveness of the influence strategies. Thus we look at their distribution to see if they are connected with the negotiation outcome or with the roles. We have looked at the distribution of verbs and personal pronouns in the previous

[1] Recall that we use the term "powerless" in the context of negotiations.

POS	BSN	SSN	BUN	SUN
MD	-	+	=	=
JJR	=	=	=	=
JJS	-	+	-	+
WDT	=	=	=	=
UH	=	=	+	-
CC	-	+	-	+
CD	-	+	-	+

Table 4.9: Coverage of text by the selected POS, successful buyers/sellers, unsuccessful buyers/sellers

POS	BSN	BUN	SSN	SUN
MD	-	+	=	=
JJR	=	=	=	=
JJS	=	=	=	=
WDT	=	=	=	=
UH	=	=	+	-
CC	=	=	=	=
CD	+	-	+	-

Table 4.10: Coverage of text by the selected POS, successful/unsuccessful buyers, successful/unsuccessful sellers

experiments. That is why we have selected for more detailed data analysis the following POS: modals, adjectives, wh-determiners, cardinal numbers and coordinate conjunctions, interjections.

We provide empirical results in Tables 4.9 and 4.10. MD are modals, JJR are comparative adjectives, JJS are superlative adjectives, WDT are Wh-determiners, UH are interjections, CC are coordinate conjunctions, CD are cardinal numbers.

The distribution of the selected POS shows that there is difference between four corpora although not statistically significant. The variances in POS over four classes are small, ranging from **0.27** (for nouns) to **0.003** (for comparative adjectives and interjections).

We continue corpora comparison in Chapter 6.

4.5.3 Comparison with non-negotiation data

We hypothesize that despite being a collection of written texts, e-negotiation text data are similar to spoken language data (Yates and Orlikowski, 1993), in particular to dialogues.

We compare our data with face-to-face conversations of upper-level college students, marked as Dialogues and borrowed from (Allen and Guy, 1974), with the Brown corpus (Francis and Kucera, 1979), the Lancaster corpus of British English (LOB) (LOB, 2004) and the Wall Street Journal (WSJ) corpus (Paul and Baker, 1992). We exemplify the commonalities and differences using the 10 most common types in these corpora; see Table 4.11.

Inspire	*SimpleNS*	Dialogues	Brown	LOB	WSJ
to	*the*	*i*	*the*	*the*	*the*
i	*to*	*you*	of	of	of
you	*i*	*and*	*and*	*and*	*to*
the	*you*	*the*	*to*	*to*	*a*
a	of	*to*	*a*	*a*	*and*
and	*we*	ah	in	in	in
your	*and*	*a*	that	that	that
offer	*a*	it	*is*	*is*	for
we	of	for	was	was	one
is	*is*	know	he	it	*is*

Table 4.11: Ten most common words in *Inspire*, *SimpleNS*, Dialogues, Brown, WSJ.

Communication via e-negotiation systems does not resemble regular dialogue. In e-negotiations people do not always communicate in turns. Some negotiators, considerably more active than their partners, send several messages before receiving a reply, while others rely on the system's tables of offers and choose not to send accompanying messages. Nevertheless, the e-negotiation data and face-to-face dialogues exhibit the same frequent use of first and second person singular pronouns **I**, **you**. This differs from the general (Brown, LOB) and business (WSJ) texts and is consistent with our view of *Inspire* negotiations as interpersonal communication. The high frequency of the negotiation-related word **offer** suggests that this communication has a specific topic – negotiating a purchase with an exchange of mandatory offers. Recall from section 2.2.1 that SRCC is a common measure for similarities between corpora with different sizes (Kilgarriff, 2001). We have calculated it on the intersections of the *Inspire* list with the other lists. The results indicate that the *Inspire* data are much more similar to the dialogue (SRCC = 0.7227) and the *SimpleNS* (SRCC = 0.76) data than to the Brown, the LOB (SRCC = 0.0681) or the WSJ (SRCC = 0.1409) corpus.

The results of this section are used in Chapter 5.

4.6 The numerical data in Inspire negotiations

In this section we look at the learning abilities of the *Inspire* bargaining data. The crucial difference between the *Inspire* bargaining data and the *Inspire* text data lies in the distribution from which they are obtained. While we have shown in Section 4.5 that vocabulary is unrestricted, the numerical offers have 180 values. We cannot combine these data together because the basic ML principle requires the data to be independent and identical distributed (i.i.d.) and thus produced from a population with the same distribution; see (Cherkassky and Muller, 1998) for detailed explanations. Nevertheless, the use of the numerical data for classification purposes assists us in understanding what kind of results we can expect from the text data.

For the analysis of the learning from the numerical *Inspire* data we consider several sets of numerical features. We look at the trends in classification accuracy when the number of numerical features increases.

We start with representing each negotiation by the exchanged values of one negotiation issue. We use the price values exchanged in the negotiation and the exchanged delivery values. The values are padded to the maximal length either by the last value or by 0. See Tables 4.12 and 4.13, where the classification based on the prices is presented. The price values were listed in chronological order, first extracted from the buyer offers and then from the seller offers. The baseline accuracy is 55.8 if we consider all examples as positive.

Classifier	Accuracy	Accuracy on positive examples	Accuracy on negative examples
SVM	74.4	95	53
NB	75.8	99.9	44.9
DT	87.7	94.8	78.7

Table 4.12: The accuracy on e-negotiations represented with the prices padded with the last price.

Note how the classifier performance changes when zeroes replace the last price repetition. For all three classifiers the true positive rate declined, for DT and NB the true negative rate declined as well. Especially, NB reversed its class accuracy performance when the padding of the price vectors was changed: from 99.9% to 42.4% on positive examples and from 44.9% to 96% on negative examples.

Classifier	Accuracy	Accuracy on positive examples	Accuracy on negative examples
SVM	79.98	91.4	65
NB	66.2	42.4	96
DT	85.36	92.9	75.3

Table 4.13: The accuracy on e-negotiations represented with the prices padded with zeroes.

To understand how a negotiation develops through time we represent it by the list of all prices, except the last one, padded with 0s. The overall classification accuracy becomes 79.6% for SVM, 65.5% for NB and 81.7% for DT. This means that the absence of the final price does not substantially change the learning ability of classifiers. Hence, the positive outcome is highly predetermined by initial and intermediate price choices.

From the literature review we know that price is the best predictor of the negotiation outcomes. That is why we keep price as a feature, gradually adding other numerical features. For the next feature set we represent a negotiation by the exchanged price and delivery values; see Table 4.14.

Classifier	Accuracy	Accuracy on positive examples	Accuracy on negative examples
SVM	77.4	88	65.8
NB	68.9	60.4	81.1
DT	83.8	92.1	75.2

Table 4.14: The accuracy on e-negotiations represented with the prices and delivery time padded with zeroes.

Then we add the payment values. Table 4.15 reports the classification results. Finally, we represent a negotiation by the complete exchanged offers which include prices, delivery, payment and return values; see Table 4.16 for results.

We see that the numerical data are highly predictable, especially when represented by prices only. Although 180 different offers construct a space of 2^{180} possibilities, the participants chose the paths easily recognized by ML methods.[2]

[2]NB correctly classified all but one positive examples when they were represented by prices padded with the last

Classifier	Accuracy	Accuracy on positive examples	Accuracy on negative examples
SVM	77.32	86.5	65.5
NB	68.29	57.8	81.7
DT	82.4	88.5	78.6

Table 4.15: The classification accuracy on the prices, delivery values and payment time padded with zeroes.

Classifier	Accuracy	Accuracy on positive examples	Accuracy on negative examples
SVM	81.7	92.6	69.8
NB	70.8	65.8	78.8
DT	82	88.5	75.7

Table 4.16: The accuracy on e-negotiations represented with the complete offers: price, delivery, payment, returns padded with zeroes.

We draw the following conclusions:

- price is the best predictor of the *Inspire* negotiation outcome; this corresponds to the situation of real-world negotiations;
- padding the price vectors with the last price allows correct classification of almost all successful negotiations;
- successful negotiations are easier to classify;
- with adding information about issues other than prices:
  - the classification accuracy steadily declines for DT and slightly fluctuates for SVM and NB;
  - the classification of the negative examples improves for SVM and DT and declines for NB;
  - the classification of the positive examples improves for NB and declines for DT.

price.

This means that adding more the negotiation issue values into a feature set reduces a gap between classification of classes and makes the performance of classifiers more balanced. We also see that the SVM performance is orthogonal to the DT performance on unsuccessful negotiations: when DT improves their classification SVM declines it and visa-verse. However, the restricted nature of the *Inspire* offer data makes it a poor material for learning negotiations. The same applies to the *SimpleNS* offer data. We analyze the performance of classifiers in more details in Chapters 7 and 8.

4.7 Computer-mediated communication characteristics of e-negotiation data

In this section we point out the stylistical peculiarities, arising from the use of Web-based NSS. These peculiarities are in fact due to the CMC characteristics present in the *Inspire* data. These characteristics correspond to variations from the standard English. We call them *noise.*

At the preliminary stage of the data investigation we found that the text data bear the non-standard characteristics of CMC text data. The characteristics are identified through analysis of the vocabulary, grammar and style. We started the exploration with a manual analysis of the original data. All *Inspire* negotiators should use English, but the use of other languages is not uncommon. We identify text portions in foreign languages and separate them from portions written in English that only sporadically include foreign words. This task is simpler than language recognition, which is essential for multi-lingual studies or for translation (McEnery and Wilson, 2001) and requires language recognition algorithms (Ludovik et al., 1999). An obvious way to identify segments written in foreign languages is to detect foreign function words but noise in data makes it highly unreliable. Instead, we search the list of types for long, "foreign-looking", words. We then find messages with such words and delete any message not written in English.

Through manual analysis we found that noise results from the following.

1. Messages with words containing non-letter characters.
2. Text segments in foreign languages, written in ASCII code.
3. Use of foreign words within English text.
4. Use of informal words, abbreviations, speech imitation and simplifications.

5. Spelling errors, missing punctuation and spaces between words and incorrect capitalization.

We consider five matching types of noise. *Noise-corrupted* words are those affected by noise. In Table 4.17 we give examples of noise-corrupted words in the e-negotiation data. Note that *SimpleNS* collection does not contain the noise type 2 because there is a high probability that the texts might be checked by a professor.

Data	2	3	4	5
Inspire	*offert* (38)	*niet*(11)	*tu* (32)	*monday* (32)
	ich (24)	*da*(8)	*yr* (24)	*deliverytime* (10)
SimpleNS	*pouvoir* (2)		*yr* (25)	30days (18)
			u (12)	monday(1)

Table 4.17: Noise examples in the *Inspire* and *SimpleNS* data.

Deliver and **negotiate** and their word forms are the most often misspelled words among negotiation-related words in both collections. In Table 4.18 we show some of their misspelled versions. We also show spelling versions of the two most often misspelled generic words, **Sincerely** and **Unfortunately**. Words written in **bold** are spelled correctly, words written in *italic* are noise-corrupted. A word's occurrence count in the *Inspire* corpus is shown in brackets, with *SimpleNS* count given after a slash.

Type	Misspelled versions
delivery(4859/695)	*delievery (21/11), delevery (11),delivey (8/2)*
negotiation (2201/228)	*negociation (152/3), negotation (24/6), negotitation (6)*
negotiate (570/97)	*negociate (64/6)*
receive (398/55)	*receive (51), recive (14)*
Sincerely (844/139)	*Sincerly (41/11), Sincerelly (7)*
Unfortunately (320/37)	*Unfortunatly (19/1), Unfortunatelly (5)*

Table 4.18: Examples of spelling mistakes in the *Inspire* and *SimpleNS* data.

We have observed that the placement of noise differs:

- noise of type 1 and 2 is concentrated in big chunks throughout the data,

- noise of type 3, 4, 5 is spread throughout the data.

The different types of noise also require different elimination approaches:

- noise of type 1 can be eliminated fully automatically from the data,
- removal of noise of type 2, 3, 4, 5 requires manual intervention.

We have compared (automatically) how the use of simplifications (e.g. *wanna, gonna, lemme, gotta, u, ur, thanx*, etc) and speech imitations (e.g. *hellloooooooooo, ooooops, huh, ahhh, atleaaast*, etc) is connected with the negotiation outcome. Simplifications and speech imitations are present in 15.5% of unsuccessful *Inspire* negotiations and in 18% of successful *Inspire* negotiations. The number of simplifications and speech imitation words vary from 1 to 11 in one negotiation, with the average 2.3 per negotiation. This holds for both successful and unsuccessful negotiations. These results show that computer-mediated characteristics do not have direct correlation with the negotiation outcome.

Due to the noise, we employ a straightforward cleaning procedure which removes negotiations conducted in languages other than English. In further studies we work with the cleaned data, although the word "cleaned" is omitted.

The results of this section are used in Section 5.4.

4.8 Text data in successful face-to-face and electronic negotiations

In this section we present a comparative analysis of the text data obtained from face-to-face negotiations (ftf-negotiations) and those coming from e-negotiations. The ftf data that we analyze here have had all non-textual elements removed. That is, the information about the body language, gestures, pauses and so on has been removed from the transcripts of these data. Also, the text data contains only transcripts of business meetings. Thus, we are not aware if there were any non-business interactions between participants, e.g. during coffee breaks.

4.8.1 The face-to-face negotiation data

Ftf data come from the *Cartoon* negotiations (Brett, 2001). It has been produced in the context of the training of negotiators, and finds use in research. We show the transcripts of ftf-negotiation business meetings resembles closely to the e-negotiation data.

The *Cartoon* negotiations are training negotiations held in US and Japan. It represents the class of *bilateral integrative business negotiations.* One participant is a buyer and the other is a seller. Both participants have the same cultural background – they are either Americans or Japanese. The goal of a negotiation is to make a virtual purchase of a TV program. A negotiation is successful if the virtual purchase has happened. The language of communication during these negotiations is English.

Although the core of the *Cartoon* and *Inspire* negotiations is the same, there are differences between them. The *Cartoon* negotiations are ftf-negotiations. The issue is the virtual purchase of a TV program. The whole negotiation process takes place during one uninterrupted meeting. The *Inspire* negotiations are e-negotiations. The issue is the virtual purchase of bicycle parts. The duration is up to three weeks, during which participants exchange freely written messages and formal numerical offers. Table 4.19 compares the various elements of *Cartoon* and *Inspire* negotiations models.

Elements	***Cartoon***	***Inspire***
Environment	Business	Business
Means	Face-to-face meeting	Electronic system
Topic	Purchase of a TV program	Purchase of goods
Goal	To make a purchase	To make a purchase
Synchronization	Synchronous	Asynchronous
Duration	one-timer	multi-timer, during three weeks
# of negotiators	two	two
Roles	buyer and seller	buyer and seller
Protocol	imposed by *Cartoon*	imposed by *Inspire*

Table 4.19: Parameters of *Cartoon* and *Inspire* negotiations

4.8.2 Example of Cartoon data

Samples from *Cartoon* negotiations follow.

I. Japanese-Japanese negotiations:

(*Buyer*) O.K. We are really interested in purchasing the program. I know that it was a popular program. I thought that because it was a big hit, it was going to take longer to be considered for a rerun. Youre ready to sell it for a rerun because its popularity wasnt as strong as its been believed, or whats the reason for making the program available for reruns so quickly?

(*Seller*) What do you mean?

(*Buyer*) I thought that popular programs would take longer to come into a re-run market. Do you have any information about why you think that the program should enter the re-run market?

(*Seller*) Well. While a program is being very popular, it makes more sense to put it in the rerun market. As you know, popular movies are made into video rental tapes fairly quickly nowadays. Its the same principle. Its a fast cycle.

(*Buyer*) I see. Youre right about popular movies. Titanic, that became a rental video right away.

(*Seller*) Thats true.

(*Buyer*) Yes, the video version of Titanic was very cheap.

(*Seller*) Thats right.

II. US-US negotiations:

(*Seller*) Good morning.

(*Buyer*) Uh, I appreciate the opportunity of calling us together, uh, to look at the possibility of procuring some of your film, and I'm just really pleased as the general manager WCHI I've got the chance to meet with you and investigate something possibly mutually beneficial.

(*Seller*) Well, and similar with me. I'm the midwest sales representative for Hollyfilms, and you couldn't be in that region of the country without looking at Chicago as a key market. So I'm certainly interested in putting together, um, hopefully a of package

(*Buyer*) Okay.

(*Seller*) Kind of as a ground rule, I'd like to establish that anything that, uh, nothing is agreed upon until all things are agreed upon. So, you know, if we–

(*Buyer*) Okay.

(*Seller*) If we come to some understanding about one aspect and then we talk about aspect four, we can go back and change aspect one without, you know, hurting each other's feelings.

(*Buyer*) I see.

(*Seller*) That will, I think, allow us the flexibility to try and come to an agreement that, you know, you need for your company and I'm looking for for mine. Uh, having said that, um, you know, I have, I have a couple films that we could talk about, but we're not limited to talking about those two films.

I'm interested in understanding as much about your company as I can, and I, uh, what do you need? I'm trying to sell you something to meet your needs.

(*Buyer*) And likewise. I want to, uh, you know, as, on the buying end of this business, uh, you know, I expect you to make a fair, uh, profit and, uh, you know, I know you you've got some sunk costs that you've already incurred in putting this film together.

4.8.3 Cartoon data analysis

Before proceeding to analyze the similarities and differences between the two data sets, we introduce them in detail and perform a vocabulary analysis.

We have access to the *tape record transcripts*[3] of 40 *Cartoon* negotiations, all of them successful. 20 negotiations were held between a Japanese buyer and a Japanese seller (the *Cartoon* Japanese data), and 20 negotiations between a US buyer and a US seller (the *Cartoon* US data). The text data contain 157,253 word tokens distributed among 4,244 word types, with 1686 *rare* word types. In this study we work with the text messages of 1,431 successful *Inspire* negotiations. The text data contain 1,073,398 word tokens distributed among 25,085 word types, with 12,897 rare word types.

This comparison shows that from the language perspective the *Cartoon* data and the successful *Inspire* negotiation data differ. The former is a small corpus obtained from recorded, transcribed and post-edited speech data, the latter is a relatively large corpus of unedited written texts. The *main question* that arises is this:

- will the common origin of the data, namely communication during a negotiation, and the common negotiation elements (for example, *business, bilateral, integrative, and successful*) bring similarities between the *Cartoon* data and the successful *Inspire* negotiations data?

To answer our question, we first employ corpus-linguistic methods to analyze and compare corpora; see Section 2.2.1 for details. We apply them partially to complete *Cartoon* data and partially to *Cartoon* Japanese and *Cartoon* US data sets.

We calculate the type-token ratio $TT(N)$ where N is the number of tokens. In both data $TT(N)$ is lower than in a benchmark Brown corpus. We explain the somehow restricted vocabulary by the fixed topic of discussion. We calculate the growth rate of the vocabulary $P(N)$ and Sichel's characteristic $S(N)$; see Section 4.5 for formulae. We present the results in Table 4.20.

[3]Some transcripts skip the introductory part of a negotiation.

Statistics	***Cartoon***	***Inspire-successful***
$TT(N)$	0.027	0.023
$P(N)$	0.4	0.5
$S(N)$	0.076	0.002

Table 4.20: The data statistics.

The empirical values of $P(N)$ show that the *Inspire* corpus exhibits traits of a richer vocabulary when compared to the *Cartoon* corpus. We explain this by the multi-cultural diversity of the *Inspire* data contributors and by the longer duration of *Inspire* negotiations. The latter allows more interaction and makes it possible to establish social contact between negotiators; the former introduces a variety of discussion topics, which may lead to a more extended vocabulary. Notice that other reported corpus statistics depend on size more than $P(N)$ (Tweedie and Baayen, 1998). $S(N)$, which is another measure of the vocabulary richness (Thomos et al., 2004), is in inverse proportion to the size of the corpus. $S(N)$ essentially tells us whether the corpus is homogeneous by converging, or not converging, to a limit; convergence is calculated on samples with the increasing sizes.

The most frequent words of e-negotiation data include nouns. This differs from the most frequent words of typical corpora, which contain only function words, for example, prepositions, determiners, and words such as pronouns or interjections. Table 4.21 lists 10 most frequent words from the *Inspire* data, all *Cartoon* data and from their Japanese and US subsets. We again compare our data with face-to-face conversations of upper-level college students (Dialogues data), the Brown corpus, and the Wall Street Journal (WSJ) corpus. We exemplify the commonalities and differences using the 10 most common types in these corpora; see Table 4.21. Spearman's coefficient, calculated on the most common words, has shown that the Japanese data are closer to the Dialogue data than to the *Cartoon* US data. Same applies to the US data: they are closer to the Dialogue data than to the *Cartoon* Japanese data; see Table 4.22. All other benchmark data (Brown, WSJ) are even more distant.

We use the results of this section throughout the dissertation to support our hypotheses on the negotiation process.

Cartoon J	*Cartoon* US	*Cartoon* All	Insp-succ	Dialogues	Brown	WSJ
i	the	the	to	i	the	the
you	to	to	i	you	of	of
to	that	you	you	and	and	to
the	you	i	the	the	to	a
million	i	that	we	to	a	and
runs	and	and	and	ah	in	in
a	we	we	a	a	that	that
for	a	a	your	it	is	for
we	of	of	offer	in	was	one
and	is	is	for	know	he	is

Table 4.21: Ten most frequent words in *Cartoon*, *Inspire*, Dialogues, Brown, WSJ

Data	*Cartoon* J	*Cartoon* US	*Insp*-succ	Dialogue
Cartoon J	-	0.45	0.53	0.42
Cartoon US	0.45	-	0.57	0.39

Table 4.22: Spearman's coefficient for the *Cartoon* samples

4.9 Conclusion

In this chapter we have analyzed e-negotiation data from multiple viewpoints,i.e. we considered different data representations gathered through different sources. We have shown that this provides the necessary insight into the data. We have shown that the e-negotiation data obtained from different sources exhibit different characteristics, hence accumulating different learning and predicting abilities.

We have obtained the following empirical results:

1. the vocabulary used in free texts exchanged in e-negotiations grows as the vocabulary of unrestricted languages in spite of the fixed topic of negotiations, the fixed issues of the topic, the restricted options, and the structured formal offers;

2. the type-token ratio of the e-negotiation data is lower than the standard type-token ratio; this observation, combined with the observation that top frequent types and rare types behave as usual, leads to the conclusion that the number of middle-ranked types is lower than in the standard corpora;

Negotiations	*SimpleNS*	*Inspire*		*Cartoon*
Types	e-negotiations	e-negotiations		ftf-negotiations
Data	***text messages***	**text messages** formal offers	questionnaires, history records, numerical data	**transcripts**

Table 4.23: Negotiations and corresponding data.

3. the most frequent words of the data are closer to the spoken dialogue data than to standard or business English; nouns and verbs have higher ranks usually occupied by function words.

4. the textual components of the e- and ftf-negotiations are considerably close statistically.

The numerical *Inspire* and *SimpleNS* data gathered in bargaining are highly predictable. For the remainder of the dissertation we concentrate on the text data; see Table 4.23. We put in **bold** the objects of our research. *SimpleNS* data plays a support role.

Chapter 5

Data Modelling

Data without generalization is just gossip.

Murphy's Law

5.1 Introduction

In this chapter we build statistical and semantic models of the textual data. We employ the results of Chapters 3 and 4.

For statistical modelling, we investigate the level of uncertainty in the e-negotiation data. We seek the model that better predicts the future behaviour of data; refer to Chapter 2. For the semantic modelling, we first propose how to identify semantic categories through the use of key elements of negotiations. After the semantic categories are identified we construct a corpus-based lexicon with general syntactic and (domain-oriented) semantic information. We have built a system containing a vocabulary extraction program, a lemmatizer and a syntactic and semantic information acquisition program. We use an off-the-shelf spell-checker and a general lexical resource with basic syntactic and semantic information. We prefer a general-purpose lexical resource because many negotiators have a limited English vocabulary which cannot be identified by a specialized business or computer dictionary.

While modelling the data, we kept in mind that Machine Learning methods will use the results in the next stage of learning, reported in chapters 7 and 8. We partially use the results of this chapter to support hypotheses in Section 9.3.

5.2 Statistical modelling of e-negotiation data

This section continues quantitative analysis of the data. We build a statistical model of the data. A model provides knowledge about word distribution, dependence and the level of data predictability. Statistical models are used in Data Generation, Machine Translation and building Human-Computer interfaces (Rosenfeld, 2000). The use of data models in dialogue systems is the area closest to our research. Statistical models, through data generation, can help improve NSSs by means of dialogue simulation. In such simulated "negotiation" the system could generate responses based on the received text messages. We suggest that responses be generated based on the statistical model which uses role-dependent parameters.

5.2.1 Modelling results

In this section we provide the results of the statistical modelling of *Inspire* data. As shown in section 4.5, the *Inspire* data have many interesting characteristics quite different than the characteristics of such widely used NLP corpora as Brown or the WSJ. Unlike those corpora, ours has a lower type-token ratio and high percentage of most frequent words, although regular percentage of rare words which is approximately equal to half of the number of all word types in the corpus. To expand our understanding of the data we look for a statistical model for the *Inspire* data. There was no previous work done on the statistical modelling of negotiation data. There also seem to exist no references to work on modelling text data from bilateral CMC.

We look at the performance of the models on the Switchboard data, the only available models of the dialogue data. We consider the complexity model of implementation (the number of adjustable parameters, the necessity of the validation set, the complexity of the smoothing formula). We chose to use trigram models with the modified Kneser-Ney smoothing method (KN model) and the Katz variant of the Good-Turing smoothing method. The KN model is the following:

$$P_{KN}(w_1...w_3) = \max\{c(w_1w_2) - D, 0\} + DN_{1+}(w_1w_2)P_{KN}(w_3|w_2) \tag{5.1}$$

where $c(w_1w_2)$ is the number of occurrences of N-gram w_1w_2, $D = \frac{n_1}{n_1+n_2}$, where n_i is the number of N-grams with occurrence i, $N_{1+}(w_1w_2)$ is the number of different $N-1$-grams preceding w_3. The GTK model is the following:

$$P_{GTK}(w_1...w_3) = \frac{r^*}{N} \tag{5.2}$$

where N is the number of tokens and r^* is the estimate of the N-gram occurrence r. Equations 2.7 – 2.9 give details on calculation the GTK probabilities. We built GTK models for $k = 5 \ldots 20$. The results show that cross-entropy remains within 0.15 difference for $k = 5 \ldots 18$. This indicates that the models fit with the same goodness (Chen and Goodman, 1998). Cross-entropy increases when k is set to 19.

Here is why we use KN and GTK models: their performance is high compared with other models, they require only training and test data, but not validation set, and are easy to implement and apply, for example, they have no adjustable parameters.

We employ five-fold cross-validation to estimate the goodness of the model. The cross-entropy results are reported in Table 5.1. Recall that the smaller the cross-entropy is, the better the model represents the data.

Model	*Inspire*	BSN	BUN	SSN	SUN
GTK model	9.94	11.09	11.25	10.81	11.17
KN model	5.69	6.42	6.61	6.57	6.57

Table 5.1: Cross-entropy, *Inspire*

The results show that the KN model fits e-negotiation data better than the GTK model. The increase of cross-entropy on four data sets is due to their smaller sizes. The four data sets' size variations affect less the KN model. We attribute this to the recursive estimation of the unseen events in KN model. Recall that the KN probability of a word depends on the number of preceding word types. If a previously unseen trigram contains an established collocation – which is often happens in *Inspire* because of the restricted topic – then KN treats such trigram differently from the one that does not contain established collocations. On contrary, GTK calculates the frequency of frequencies of N-grams on the training set, thus heavily depending on the training and testing set sizes.

The following natural question arises: how well can one data set predict a word from another? Put differently, how close can the buyer and seller data approximate each other with respect to models? In order to answer, we have built cross-models of the data for successful buyers and sellers and unsuccessful buyers and sellers. For cross-modelling, we use a training set from one data class and a testing set from another data class. Table 5.2 presents the results of cross-modelling. X vs Y means that X is a training set and Y is a testing set. To make a fair comparison, we use the same training and testing sets as in the previous modelling experiments.

Model	BSN vs SSN	BUN vs SUN	SSN vs BSN	SUN vs BUN
GTK model	10.94	11.26	11.26	11.35
KN model	6.33	6.64	6.48	6.65

Table 5.2: Cross-modelling results for the buyer and seller data

Predictably, the KN model fits the data better than the GTK model. For both models, fitness declines and cross-entropy increases while cross-modelling unsuccessful buyer and seller data. Interesting results appear in the BSN versus SSN case: both models improve their fitness! This phenomenon requires additional attention, possibly in future work.

With respect to the language of e-negotiations, the cross-entropy shows that its complexity and predictability lie within the range usual for English texts. The cross-entropy of English texts ranges from around 5.64 to 9.70, depending on the type of text and models (Chen and Goodman, 1998).

5.2.2 Comparison of N-gram distribution within data classes

As we said in Section 4.4, we separate the text data into classes with respect to the negotiation outcome and the roles of participants.

For four classes, we build lists of unigrams, bigrams and trigrams, and rank the N-grams by frequency. We will use these frequencies to find language patterns (see section 6.2). Briefly, we compare the ranks of the same N-gram in successful and unsuccessful negotiation data, and find N-grams frequently present in one set and rarely present or absent in the other. These N-grams are easily detected when we plot N-grams from successful and unsuccessful negotiations; see Fig. 5.1 for the plot of unigrams. The N-grams with a large difference in ranks are depicted as the outliers. Note that the graph for successful negotiations lies above the graph for unsuccessful negotiations because the data of successful negotiations are more numerous than the data of unsuccessful negotiations.

Figure 5.2 shows the similarity in the distributions of 300 most frequent unigrams in buyer and seller data sets. We have compared the occurrences of these unigrams in our data when they are used by all negotiators, only by buyers, and only by sellers. These unigrams are mostly stop words, negotiation-related words such as **offer, price, delivery, agree, accept**, and process-related words such as **send, receive**. Further analysis of word categories appears in section 5.4.1.

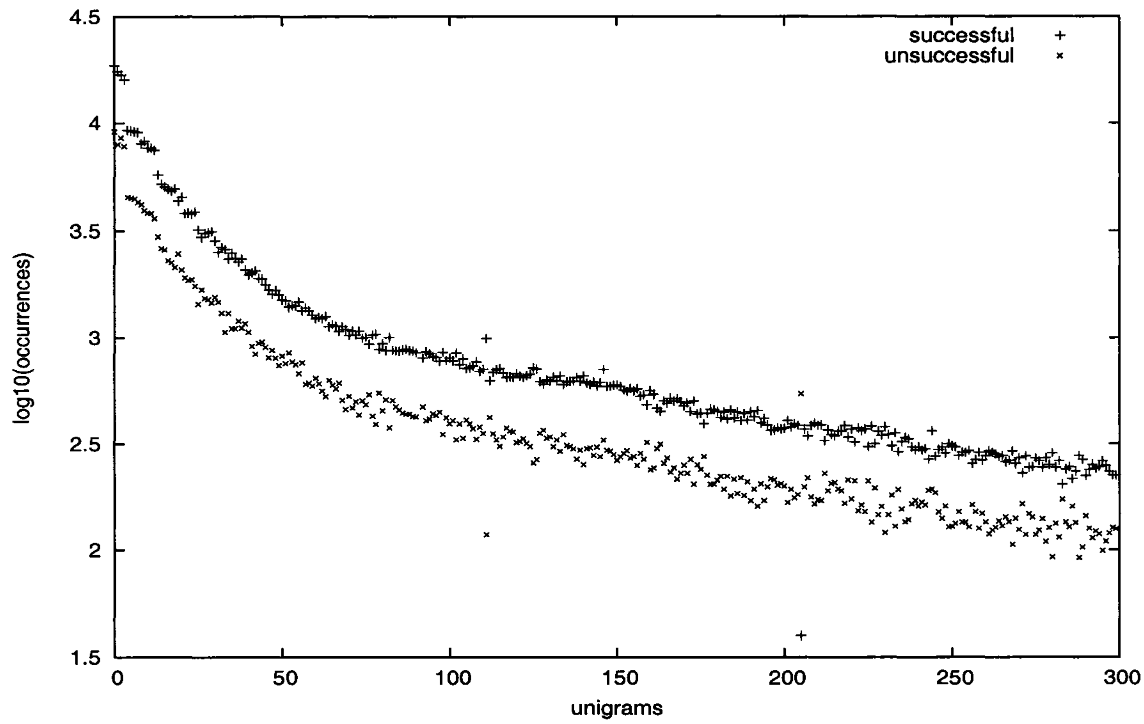

Figure 5.1: Distribution of unigrams, successful/unsuccessful negotiations

We employ the statistical modelling results in Section 5.3, throughout Chapter 6, and in Section 7.2.

5.3 Statistical modelling of face-to-face-negotiation data

In this section we compare statistical models of the *Cartoon* and the *Inspire* data. Corpus analysis shows that the *Cartoon* data indeed consist of two easily distinguishable subset: the *Cartoon* Japanese data and the *Cartoon* US data. We model these data separately. We use samples of successful *Inspire* negotiation data to get comparable results since the statistical modelling results depend on the size of the data set.

5.3.1 Sampling

We consider the size in tokens and the number of negotiations as the main parameters for choosing samples. The average length of a successful *Inspire* negotiations is 750 tokens per negotiation. We have found a subclass of *Inspire* negotiations where the average negotiation length is 1,160 tokens. Both

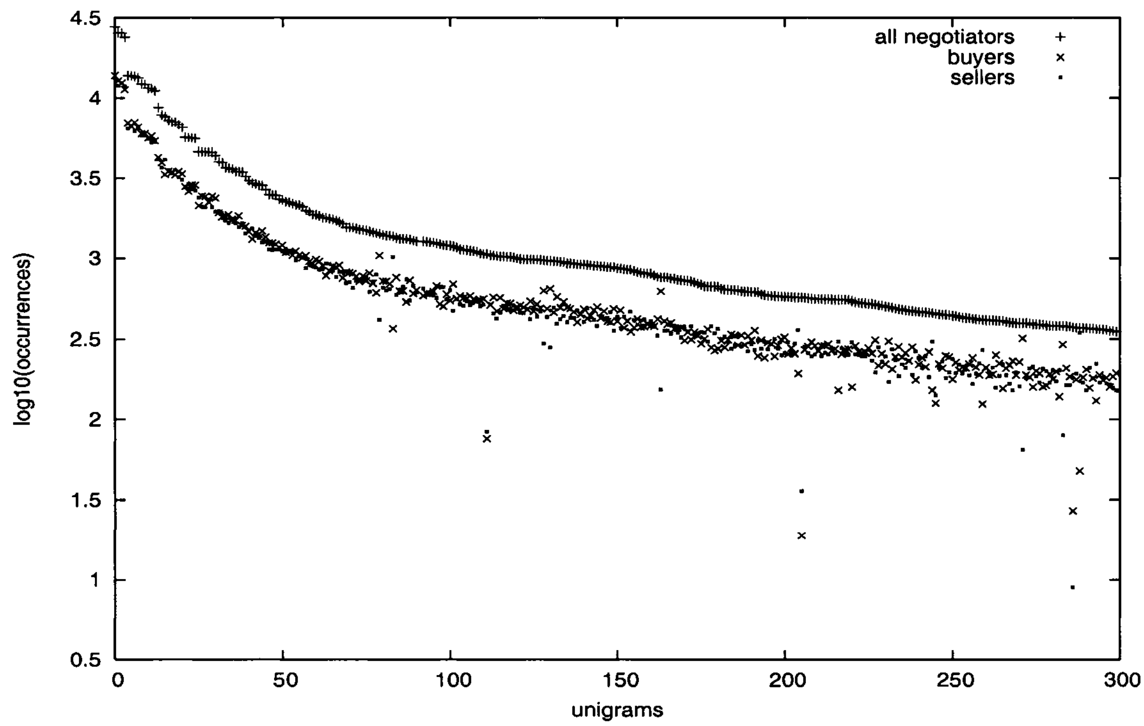

Figure 5.2: Distribution of unigrams, buyers/sellers

average lengths are closer to the *Cartoon* Japanese negotiations in which the average length is 1,340 than to the *Cartoon* US negotiations with 6,250 tokens on average. We construct two samples from the successful *Inspire* negotiation data. *InspGen1* is the sample constructed from successful *Inspire* negotiations and *InspEmail* is constructed from the more lengthy successful *Inspire* negotiations. For both samples we chose negotiations distributed uniformly in the original population.

The results show that the number of tokens and the number of negotiations of the samples correspond to those of the *Cartoon* Japanese data while the word distribution preserves the distribution of the successful *Inspire* negotiation data; see Tables 5.3 and 5.4. The results prove the sample characteristics are similar and the modelling results will be comparable.

We calculate the growth rate of the vocabulary $P(N)$ and Sichel's characteristic $S(N)$; refer to Section 2.2.1. We also calculate the ratio of bigrams $bi(N)$.

$$bi(N) = \frac{\#\ of\ bigrams}{N} \tag{5.3}$$

and the ratio of trigrams $tri(N)$

$$tri(N) = \frac{\#\ of\ trigrams}{N} \tag{5.4}$$

Characteristics	*Cartoon* **J**	**InspEmail**	**InspGen1**	*Cartoon* **US**
# of negotiations	20	25	33	20
N	26795	26987	27218	130458
T	1412	2764	2539	3702
V(1,N)	534	1333	1180	1357
V(2,N)	194	390	361	519

Table 5.3: Word distribution; comparison of the samples.

rank	*Cartoon* **J**	***InspEmail***	***InspGen1***	rank	*Cartoon* **J**	***InspEmail***	***InspGen1***
1	i	i	you	6	runs	a	and
2	you	the	to	7	a	we	we
3	to	to	i	8	for	is	for
4	the	you	the	9	we	for	your
5	million	and	a	10	and	your	offer

Table 5.4: Ten most common words in the samples.

Table 5.5 presents the empirical results.

Statistics	*Cartoon* **J**	***InspEmail***	***InspGen1***	*Cartoon* **US**
$TT(N)$	0.053	0.102	0.093	0.028
$P(N)$	0.020	0.049	0.043	0.011
$S(N)$	0.007	0.014	0.013	0.003
$bi(N)$	0.377	0.510	0.337	0.303
$tri(N)$	0.714	0.827	0.529	0.699

Table 5.5: The sample statistics.

The results show that $bi(N)$ are $tri(N)$ higher for *InspEmail.* Recall that this sample contains messages where negotiators discuss personal matters. This leads to richer vocabulary because the negotiators discuss different topics. We also suggest that $bi(N)$ and $tri(N)$ depend on the language fluency of the data producers.

5.3.2 Models

In our study we employ the *Kneser-Ney* and *Good-Turing* trigram models; for details and formulae see Section 2.2.2. We employ five-fold cross-validation to train and test statistical models. Table 5.6 reports the average five-fold cross-validation results for all experiments. GTK stands for the Good-Turing model with Katz smoothing, KN for the Kneser-Ney model.

Model	*Cartoon* J	*InspEmail*	*InspGen1*	*Cartoon* US
GTK	10.06	11.8	11.04	12.21
KN	6.67	7.79	7.15	7.27

Table 5.6: Cross-entropy, samples

For all four data sets the KN model outputs lesser cross-entropy than the GTK model. This corresponds to the KN and GTK performance results reported by (Chen and Goodman, 1998) for the Switchboard corpus. We cite the results here because the Switchboard corpus contains transcripts of spoken dialogues between a telephone operator and a client. Thus, while it is not a business negotiation, Chu-Carroll and Carberry (2000) refer to such conversation as collaborative negotiation.

The empirical results show that the predictability of *Cartoon* Japanese data is closer to the predictability of *InspireGen1* samples than to the predictability of *Cartoon* US data. The vocabulary richness of *Cartoon US* is closer to that of *InspireEmail*; this supports the previous claim about vocabulary being richer in *Cartoon US* and *InspireEmail* than in the other two samples.

We continue with cross-modelling of the samples. Table 5.7 presents the results of cross-modelling. X vs Y means that X is a training set and Y is a testing set. Again, we use the same training and testing sets as in the previous modelling experiments. Prediction of a word becomes harder when a

Model	*Cartoon* J vs US	*Insp Email* vs *Gen1*	*Insp Gen1* vs *Email*	*Cartoon* US vs J
GTK	9.84	11.33	11.49	12.30
KN	7.35	7.14	7.39	7.62

Table 5.7: Cross-modelling results for the sample data.

model is trained on *InspGen1* and *Cartoon* US data and tested on *InspEmail* and *Cartoon* Japanese data respectively. The same holds for KN model trained on *Cartoon* Japanese and tested on *Cartoon* US data. However, GTK model improves its fitness for *Cartoon* US data when trained on *Cartoon*

Japanese data. This is an unexpected result.

Other improvements of the model fitness are due to the sample data. Cross-entropy reported in *Insp Email* vs *Gen1* is smaller than in *InspEmail.* Note that the same training data was used in both cases. Difference in cross-entropy results signals that the *InspGen1* testing set has fewer unseen data than the *InspEmail* testing set. This supports the previously stated hypothesis that people use diversified vocabulary when exchange personal information.

We use the results of ftf-negotiation modelling in Section 5.5.

5.4 Semantic Model

In this section we build a semantic model of the data. We introduce semantic categories related to e-negotiation data and use them in modelling the data. We define semantic categories based on the properties of negotiations reported in Chapter 3. Specifically, we find the semantic categories that capture the characteristics of interpersonal communication in our data. A list of word types with semantic tags is the resulting model.

5.4.1 Semantic categories for e-negotiation data

We consider a word type (Lyons, 1995), or a word, to be the basic feature of the text data. Every word can be used in different situations and have different meanings depending on context (Matthews, 1995). For example, in our domain the word "message" denotes a neutral statement about the message received by the negotiator while the word "messages" denotes a complaint about either the counterpart or the *Inspire* system. However, both words indicate a communication process supported by Inspire.

We investigate how the meanings of the words used in e-negotiations can be generalized in groups. We call these groups *semantic categories.*

We consider a semantic category to be a characteristic of the data. We want to know which groups of words are essential to e-negotiations, how these groups are relevant to learning from e-negotiation data and how the search process for these groups can be generalized to other data. To establish which semantic categories are necessary and sufficient for the data to be grouped, we look at the key elements of CM and how they apply to e-negotiations (Argyle et al., 1981; Hargie and Dickson, 2004).

In our case the data come from e-negotiations. The topic of negotiations and the issues discussed are the reason for these data to appear. Words related to negotiations are unusually frequent as compared to the standard corpora (see Section 4.5). Hence, the most important category is negotiation-related. This category also includes words appearing in the so-called business expressions, e.g. "thanks for you offer", "I hope this will work", etc. The data are gathered through an exchange of electronic messages supported by the NSS. We assign a category which relates to the electronic negotiation support and a category which tags such features as CMC characteristics, e.g. informal words, speech imitations, abbreviations. The latter category does not include misspelling, mistyping, miscapitalization and wrong spacing.

The data are gathered through a rather lengthy dyadic communications. This is why a possibility exists that partners exchange information not related to the main topic. This information can be divided into two subgroups: one for contact information which includes names, place names and email, and another for extra information which can be roughly subdivided into personal and professional. In general, the groups of the words describing personal and professional information depend on the contingent of data contributors. As we showed in section 4.3, 82.8% of the participants of e-negotiations are students with other well defined groups covering less than 2% each. Thus student life style is dominant and brings forth general hobbies as personal and studies as professional categories.

Obviously, these categories cannot include or predict every word because the vocabulary grows as the vocabulary of unrestricted languages, as shown in section 4.4. This means that the unpredicted words can appear in the data. We add one category "Others" to collect all the words that are not function words and could not be mapped to other categories.

5.4.2 Building a semantic lexicon

As we showed in section 4.7, the data of e-negotiations are highly affected by noise. Thus, the preliminary phase consists of cleaning the data. We build a lexicon out of the cleaned data. The cleaning procedure was described in section 4.7. We now describe a lexicon-building procedure as follows.

1. Identify semantic categories.
2. Build the list of word types.
3. Identify and separate "lexicon-ready" data.

4. Process the remaining data and increase the "lexicon-ready" portion.

5. Build a lexicon from the identified data.

We begin with identifying semantic categories (see Section 5.4.1) and obtaining a list of word types from the unigram model of the data (see Section 5.2). In the next step we divide the obtained list in two: words which should be tagged with personal semantic categories (*Personal names, Email addresses, Place addresses*), such as proper names (Australia, Fumiko, Micheal, Sahraj), and non-personal types. The latter words are corrected for spelling and then automatically lemmatized. Note that multiculturalism of the negotiators brings legitimately written words that are not recognized by a standard spell checker, and wrongly marked for correction. This false correction introduces new noise. To avoid false correction, we perform partial correction using `ispell`. We build a list of most commonly misspelled negotiation-related words; see Appendix B. A misspelled word is corrected if its corrected version belongs to the list. We employ isolated-word error correction (Jurafsky and Martin, 2000), because the high volume of noise makes content-dependent error correction unreliable. If the spell-checker suggests several corrections, then we check the number of occurrences of each suggested substitution. We select a substitution with the most occurrences in the data.

After correction and lemmatization the non-personal types are divided into words present in the lexical resource LDOCE (*dictionary types*), and words not found there (*non-restorable types*). Non-restorable types include foreign words (*bien*, *niet*, *Zdravstvuj*), misspelled words (*effictive*, *goodby*, *offert*), words with non-letter characters(*Fant^{im}mas*, *won''t*), informal (*Helllooooooooo*, *thnax*, *Thnks*) and slang words (*gotta*,*u*), abbreviations (*pls, btw*) and unrecognized proper names. The CMC category words are filtered out from the non-restorable types.

The fifth step equips a dictionary word with syntactic and semantic information. Syntactic information consists of part-of-speech tags. Semantic information consists of word senses and semantic tags. Semantic information, which can be tuned to the domain, defines *semantic categories*. Categories classify words into several general topics. The e-negotiation data fall into ten categories (the last of them comprises closed-category words).

{Negotiation-related, e-negotiation process, Studies, Hobbies, CMC, Personal names, Place addresses, Email addresses, Others, Function words}.

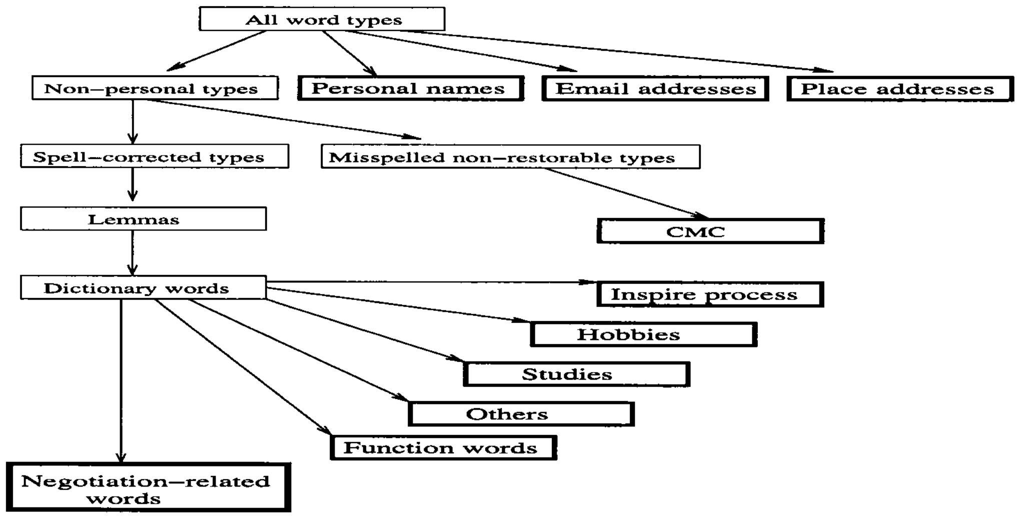

Figure 5.3: Application of the lexicon-building procedure on e-negotiation data.

All categories but Others and Function words are listed according to the percentage of text they cover. The first four categories vary by the domain of the text collection.

As the final step, the procedure gets dictionary words tagged by semantic field tags from the lexical resource, e.g. LDOCE. We present the procedure outcome in Figure 5.3.

We use LDOCE (Summers, 2003) as the lexical resource. The semantic field tags provided by LDOCE are organized in a hierarchical structure (see Appendix A). For each category tag we find which set of LDOCE tags gives the same semantic information. We seek an *automatic* mapping between data-dependent category tags and sets of semantic field tags.

We say that a mapping classifies a word correctly if one of its fields corresponds to a semantic category, within which the word appears in the corpus. To improve semantic tagging the most frequent classified words are manually checked using frequencies of bigrams (and trigrams, if necessary). Among all available mappings we choose the mapping that gives the smallest number of misclassified words. Now we can describe the step of the procedure that tags dictionary words with syntactic and semantic information. To find syntactic and word-sense information we extract POS and word-sense values from LDOCE and tag dictionary words with those values. Now, we automatically build a semantic lexicon and tag the remaining words with semantic tags.

Once a semantic lexicon has been built, we calculate the percentage of occurrence among the 100 most frequent unigrams of the words of each semantic category except function words. We work with

100 most frequent unigrams because they coincide for the data of the successful negotiations and the data of the unsuccessful negotiations.The largest percentage comes from the words of the negotiation-related category – as expected (see Table 5.8). For instance, the word *offer* appears among the 10 most frequent words in our corpus; only function words appear among the top 10 frequent words in the three widely used corpora Brown, WSJ and LOB. In Table 5.9 we compare the ranks of most frequent negotiation-related words common in the *Inspire* and *SimpleNS* data with their ranks in the *Brown* corpus. 5000+ means that the word's rank > 5000.

Category	***Inspire***	***SimpleNS***
Negotiation-related	57.9	62
e-negotiation process	5.4	0
Others	36.7	38

Table 5.8: Distribution of categories in 100 most frequent unigrams

Type	***Inspire* rank**	***SimpleNS* rank**	**Brown rank**
offer	8	18	1320
price	20	26	5000+
delivery	27	37	4993
accept	35	92	5000+
days	40	42	229
payment	43	66	2053

Table 5.9: Ranks of common negotiation-related words

We now consider only the negotiation-related words to further build our data set and to perform learning and classification. We give further experimental details in Section 5.4.3.

5.4.3 Mapping to the semantic fields of Longman Dictionary of Contemporary English

The semantic fields in LDOCE have a hierarchical structure. For example, BUSINESS (B) includes BUSINESS BASICS (BB), which in turn includes ADVERTISING (BBA), COMPANIES (BBC), BUSINESS

MANAGEMENT (BBB), MARKETING (REM), OFFICES (BBO), TRADE; see the electronic version of (Summers, 2003) for more details. For five categories defined in section 5.4.1, we constructed a mapping into fields through search of the correct LDOCE tagging of most frequent words in *negotiation-related, hobbies, studies* categories. We present the mapping in Table 5.10. CMC words, personal and place names, e-mail addresses do not appear in LDOCE, so no mapping is shown for these categories. Function words are a closed class, in this case no semantic analysis is required for them.

We provide examples when LDOCE tags correspond to data-dependent semantic categories:
athlete: POS: n , FIELD: DS (Sport) (*Hobbies*)
accountant: POS: n , FIELD: BF (Finance) , FIELD: BO (Occupation) (*Negotiation-related*)
partnership: POS: n , FIELD: BB (Business Basics) (*Negotiation-related*)

The following examples of words tagged only with medicine or publishing illustrate that exhaustive search is necessary, and justify manual intervention into automatic semantic tagging.
analyse: POS: v , FIELD: MP (Psychology&Psychiatry) (*Negotiation-related)*
growth*: POS: n , FIELD: MI (Medicine) (*Negotiation-related)
LDOCE tags the word "margin" only as PUBLISHING, but the bigrams "profit margin", "low margin", "gross margin" show that for *Inspire* data the word should be negotiation-related. In some cases bigrams do not provide enough information about a word and we have to use trigrams. For example, the LDOCE field tag of "delivery" is BIRTH. This tag does not correspond to the e-negotiation domain. The most frequent bigrams do not provide enough information to tag the word ("the delivery", "upon delivery", "delivery time"). Only the 10^{th} most frequent bigram "delivery payment" and the 15^{th} bigram "price delivery" relate the word to the negotiation-related category. On the other hand, the 2^{nd} trigram is "payment upon delivery", the 4^{th} one is "delivery and payment". So, we put the word "delivery" in the negotiation-related category.

Table 5.10 reports the mapping with the smallest number of misclassified words. From some LDOCE fields we first remove words in subfields. For example, we remove biological terms (they appear as nicknames in the *Inspire* data) from the education field, and we mark the remaining words Studies. In the LDOCE column, "A without B" means that we extracted all words tagged by A but not by B, because words tagged by both do not belong to the corresponding category. In most cases, for example when A = BUSINESS and B = DAILY LIFE, these words were used as nicknames. Recall from section 4.2.1 that the purchase of bicycle parts is the topic of negotiation.

Data-dependent tag	**LDOCE field tags**
Negotiation-related	Business, Crime & Law (both without daily life) birth, death, publishing, bicycles, cars (all of them without biology)
Inspire process	Data processing & computing
Hobbies	General sports, leisure
Studies	Education, general economics (both without biology)
Others	Daily life without sports, general society, politics, religion general transport, general engineering, general industry, (all of them without biology)

Table 5.10: Correspondence between the *Inspire* semantic categories and LDOCE fields.

We use the results discussed in this Section for feature selection and data representation; see Chapters 7 and 8.

5.5 Conclusion

In this chapter we have reported the following results:

1. Statistical modelling:

    - although the vocabulary grows as the vocabulary of unrestricted languages, the predictability of the data is higher compared to the range of predictability of typical English texts;
    - Kneser-Ney model better fits negotiation data than Good-Turing model. The former finds collocations established in a training data and uses them to predict a word, the latter uses frequency of the frequencies of N-grams in a training data;

  - the predictability values are almost the same for the e-negotiation and ftf-negotiation text data of similar sizes ;

2. Semantic modelling:

  - there is a set of semantic categories that capture the key elements of communication during negotiations;
  - the semantic categories are defined based on the person-situation context of CM presented in Section 3.2.

We have presented a lexicon-building procedure. Though the data exploration and lexicon construction procedure was designed for specific data, it is adjustable to data gathered from sources other than the *Inspire* system. In order to use the procedure on other data, the researchers should consider the types of noise and how to clean them from the data and define data-dependent semantic categories.

The procedure builds a lexical language model for use in e-negotiations. The lexical model is used for feature selection and data representation in further studies of e-negotiation data. As an application, the model can be used in a study of the influence of cultural, educational and sociolinguistic background on the process and the results of negotiations; see Section 9.5.

Many possible extensions result from this chapter. The statistical models can prove quite useful in dialogue simulation. The semantic models can in future assist in building lexical resources adjusted for Web-based communication data.

Chapter 6

Learning Language Patterns

Each unacceptable offer has an equal and opposite unreasonable demand.

Murphy's Law

6.1 Introduction

In this chapter we aim to find the significant characteristics of the "business-oriented" negotiation process described in Chapter 3. From negotiation theory we employ a well-established term *language patterns* (Dupont and Faure, 2002; Simons, 1993) which includes discourse style, timing of messages, question-answer sequence, etc. In our study we define a **language pattern** as a sequence $w_1 ... w_n$ found in the corpus. For each w_i two values are given: the POS and the sense with respect to the negotiation and influence strategies. For example, *"the best deal I can make"* and *"the best price we can give"* belong to the same language pattern "**the best** *noun PersPron1* **can** *verb*". Our search for language patterns is based on the results on relations between the negotiation and influence strategies and language, discussed in Section 3.6.

Due to the absence of nonverbal communication in e-negotiations, messages that negotiators exchange are the main available source of behaviour disclosure (Hargie and Dickson, 2004; Perloff, 2003). We look for language patterns that signal the use of common negotiation tactics that, in the verbal mode, include substantiation, argument, persuasion and bargaining. We look for expressions of appeal which usually mark different negotiation strategies (Brett, 2001). Our main hypotheses are that the language used in such messages varies according to the roles assigned to the negotiators and according to the outcome of negotiations. We disregard the negotiations where only one negotiator activated the system because these negotiations cannot contribute to the studies of different roles. This data reduction leaves us with a collection of messages of more than 1,000,000 words, exchanged among more than 4000 negotiators. We look for general language trends and learn the language patterns from them. We observe that role-dependent language patterns do exist, despite a wide variation in the negotiators' cultural and educational background, current occupation and fluency in English. We support this observation by statistical information from *Inspire, SimpleNS* and *Cartoon* negotiations whenever the data information is applicable.

Along with seeking the language patterns we compare the performance and results of two corpus-based learning approaches to the identification of these patterns: one based on calculating the log-likelihood statistics and another based on calculating the rank difference. The chapter concludes with hypotheses on the role-dependent language patterns and language patterns used in successful and unsuccessful negotiations.

We use the extracted information in Chapters 7 and 8 to find features relevant to the negotiation studies.

6.2 The indicative words in buyer and seller data

We have hypothesized that the textual data in negotiations allow us to distinguish buyers and sellers. In section 2.2.1 we have described the significant relative frequency technique. We use this technique to find words on which buyer and seller data differ. We compute the log-likelihood statistics LL for each word w given by Equation 2.4. The words with higher LL are better indicators of a corpus because they are frequent in our corpus and are rare in the other. Thus, we call the found words the *indicative* words.

The procedure of finding the indicative words is straightforward. We calculate $LL(w)$ in directional and non-directional cases. A directional case, e.g. sellers vs buyers, finds words which are frequent in seller data and rare in buyer data. Table 6.1 reports 10 (directional) indicative words with highest LL. These words are useful when we seek for differences between buyer and seller data. The non-directional case combines two directional cases. The (non-directional) indicative words are used as features in Machine Learning, when the unified data representation is required. We list all the indicative words in Appendix B.

Direction	Indicative words
Buyers vs Sellers	cypress, cycles, parts, pay, bikes, suppliers, supplier, purchase, need, bike
Sellers vs Buyers	itex, manufacturing, our, department, below, low, reduce, deliver, return, lower

Table 6.1: Ten indicative words with the highest *LL*.

The company names – Cypress Cycles for buyers and Itex Manufacturing for sellers – head both lists. Other indicative words differ for buyers and sellers. The buyer indicative words signal their purchase interests. The seller indicative words signal bargaining.

The interesting fact is that the words corresponding to the negotiation and influence strategies are among the indicative words, e.g. *need, low, best, can, must*, with *need, low* being among ten most indicative words. We use these words as pattern seeds in Section 6.3. We use the indicative words as features to classify buyers and sellers in Section 7.7.

6.3 Language patterns in buyer and seller roles

In this section we compare the language of buyers and sellers within the same outcome class. The superlative adjective **best** is used primarily as the attribute of nouns and pronouns, as in **the best** *offer* or *It is* **the best**. They indicate "the sole feature of the referent" (Warren, 1984) for someone who uses them; this indication is, or should be, "evident to one's interlocutor" (ibid.). While it appears only marginally more often in successful than unsuccessful negotiations, the adjective **best** is used in different contexts in the four classes. In successful and unsuccessful negotiations, sellers use it more often than buyers in the negotiation-related context, in the patterns **the best** *offer/price/deal PersPron* **can** *make/do/give* or *offer/price/time* **is the best** *that/for*. In the literature on language use in negotiations (Brett, 2001), these patterns are considered indicative of substantiation and persuasion. Referring to the identifying function of the superlative adjective **best**, we say that sellers tend more than buyers to emphasize that their offer cannot be improved and that this should be obvious to their partners.

An interesting difference comes from the use of **can** and **will, must** in sellers' and buyers' data. While both BSN and BUN use **will** and **must** slightly more often than **can**, for SSN and SUN the reverse is true: **can** is used slightly more often than **will** and **must**. In their most common use, **can** expresses the possibility and ability of doing something, **must** – obligation, requirement or logical necessity of some actions; **will** means prediction or volition.

The pattern **you can** *Verb* is used more by sellers than by buyers in successful negotiations. The reverse is true in unsuccessful negotiations: the buyers use **you can** *Verb* more often than sellers. We do not specify the verbs in these patterns because of their variety except that *Verb* in patterns **you can/must/will** *Verb* does not include the verb **be**. As we said before, we are interested in business-oriented patterns. In the *Inspire* data **you can/must/will** *be* mostly correspond to the assumption about the counterpart's personality, e.g. *you must be a nice girl.*

As to the use of the patterns *I/we* **can** *Verb* and **you can** *Verb* within the same class, the ratio is 2.9 for SSN, 2.6 for BSN, 3.3 for SUN, and 2.4 for BUN. Although the use of the impersonal pattern **it can** is infrequent in both successful and unsuccessful negotiations, it is worth noting that it appears 3.5 times more often in successful than unsuccessful negotiations. This suggests that the use of impersonal patterns is positively correlated with successful negotiations.

As it was shown before, buyers use the modal **must**, that is, express requirements and logical necessity in their arguments, more often than sellers. However, different trends in its use appear when we compare successful and unsuccessful negotiations. Buyers and sellers in successful negotiations apply **must** to themselves relatively more often than in unsuccessful negotiations. This is true when the main verb is a mental verb or a volition verb. BSN use **must** in the patterns *I/we* **must** *say/tell/inform* (mental verbs) and *I/we* **must** *insist/change/make* (volition verbs) 4 times more than in the patterns **you must** *consider/understand/think* and **you must** *send/pay/change.* For BUN the ratio is only 2. SSN use the patterns *I/we* **must** *say/tell/inform* (*change/make*) 2.5 times more than the patterns **you must** *consider/understand/think* (*send/pay/change*). SUN use the former patterns only 1.3 times more than the latter ones. This shows that in successful and unsuccessful negotiations participants exhibit self-obligation and self-control more than they assume authority over counter-parts. However, the comparative decisiveness is higher in successful negotiations than in unsuccessful negotiations.

The modal **will** appears most often in the same patterns as **must**, which corresponds to its most common use. Its use varies in successful and unsuccessful negotiations. In Table 6.2 we compare the BSN and SSN collections, and BUN and SUN collections. *Noun* means offer/price/deal/time/delivery and their spelling versions. *Verb* in patterns **you ...** does not include the verb **be**. Brackets indicate that a word is optional. The pattern distribution in the *SimpleNS* data belongs neither to successful nor to unsuccessful *Inspire* data distribution. We explain this by the uncertainty of the *SimpleNS* negotiation outcomes and by the restricted number of *SimpleNS* contributors who participate in several negotiations; see Section 4.2.4. *MI* was calculated for each word for two cases: BSN and SSN, BUN and SUN. Its value ranges from **0.19** (for **best** in BUN and SUN) to **0.02** (for **will** in BSN and SSN). The small MI values correspond to small variations in the data (see the end of section 4.5). *MI* is a reliable characteristics of the common words; see Section 2.2.1. For numerical values refer to Table 8.2 in Appendix C.

Pattern	*Inspire*				*SimpleNS*	
	BSN	SSN	BUN	SUN	B	S
the best (*Noun*) *PersPron* **can** *Verb*	-	+	-	+	=	=
Noun **is the best** *Prep*	-	+	-	+	=	=
I/we **can** *Verb*	-	+	-	+	-	+
you can *Verb*	-	+	+	-	-	+
it can *Verb*	=	=	+	-	+	-
I/we **must** *Verb*	+	-	+	-	+	-
you must *Verb*	=	=	-	+	-	+
it must *Verb*	=	=	+	-	=	=
I/we **will** *Verb*	+	-	+	-	+	-
you will *Verb*	=	=	-	+	-	+
it will *Verb*	=	=	+	-	+	-

Table 6.2: Distribution of patterns with modal verbs, buyers and sellers

Table 6.2 shows that in successful negotiations buyers and sellers request counterparts' concessions with the same frequency. However, in unsuccessful negotiations sellers request concessions from buyers more often than buyers request from sellers. The impersonal requirement is used with the same frequency by buyers and sellers in successful negotiations, when in unsuccessful negotiations buyers use it more often than sellers. A summary of the results follows:

- patterns indicating requirement and obligation are more frequent in unsuccessful negotiations than in successful negotiations;

- patterns indicating persuasion are used more often by sellers than by buyers;
- patterns indicating possibility are used with self-referring by sellers more often than by buyers;
- patterns indicating self-obligation are used more often by buyers than by sellers;
- patterns indicating impersonal requirements are used more often in successful than in unsuccessful negotiations;
- patterns clarifying negotiation moves are more frequent in successful than in unsuccessful negotiations;
- participants refer to their own necessary actions more often than they require actions from counter-parts; however, the ratio self-requirement/request-to-others is higher in successful negotiations than in unsuccessful negotiations.

Results presented in this section support our hypotheses about participants of unsuccessful negotiations being more demanding than participants of successful negotiations (see Section 8.2), about participants of successful negotiations using more implicit language than participants in unsuccessful negotiations (see Sections 8.2 and 8.3). The results also suggest that sellers concentrate their strategies on the counterparts and buyers concentrate their strategies on themselves. The latter corresponds to conclusions that sellers may be more dependent on buyers than buyers on sellers (Drake, 2001).

6.4 Language patterns and the intensity of negotiations

Kersten and Zhang (2003) studied the dependence of the negotiation outcome on the intensity of offer exchange. Their results have stimulated us to test the use of language patterns with the superlative adjective **latest**, for example, **the latest** *offer/price/delivery PersPron* and *PossPron* **latest** *offer/price/delivery*. Such patterns correspond to the negotiator's reaction either to their own or the partner's move. Buyers and sellers in successful negotiations use them more often than buyers and sellers in unsuccessful negotiations as it is shown in Table 6.3. We conclude that in successful negotiations buyers and sellers react more often. This supports the results – obtained from the non-textual *Inspire* data – that positively correlate the frequency of offers and the negotiation outcome.

The procedure of finding characteristic language patterns is not unique. We demonstrate another way of finding language patterns, which is based on the N-gram model. A formal description of the

Pattern	BSN	BUN	SSN	SUN
the latest *Noun PersPron*	+	-	+	-
PossPron **latest** *Noun*	+	-	+	-

Table 6.3: Distribution of patterns with *latest*, successful/unsuccessful buyers, successful/unsuccessful sellers

procedure is presented in section 6.7. We employ the bigram model to find more language patterns indicative of the intensity of negotiations. The pattern seeds are negotiation-related words. We locate the bigrams, containing the most frequent negotiation-related words, in the lists of 100 most frequent bigrams from successful negotiations, and present them and their ranks in Table 6.4. The smaller rank means that the bigrams appears more often in successful negotiations than in unsuccessful ones.

Inspire			***SimpleNS***	
bigram	rank_s	rank_u	bigram	rank
your *offer*	2	2	*number* of	8
the *price*	5	10	30 *days*	18
offer I	24	30	signing *bonus*	24
offer and	32	44	your *offer*	37
new *offer*	37	65	the *price*	38
my *offer*	39	31	*contract duration*	46
this *offer*	41	60	of *songs*	47
last *offer*	45	48	*promotional concerts*	52
accept your	48	69	*shipment* is	58
to *accept*	57	70	*days* after	59

Table 6.4: Bigrams of the most frequent negotiation-related words.

Table 6.4 shows that *new offer, this offer, last offer* are more frequent in successful negotiations than in unsuccessful. This means that in successful negotiations negotiators are more precise about negotiating moves. In unsuccessful negotiations *my offer* is ranked higher than in successful. This supports our previous conclusion about negotiators being more self-centered in unsuccessful negotiations than in successful.

The higher ranks of the bigrams *the price, the delivery, the payment, offer is, price is*, which deal with specific issues of negotiations, support our claim that the language used in successful negotiations is more direct than the language used in unsuccessful negotiations. This result supports the positive correlation of "powerful" speech to the influence strategies (see Section 3.6.1). The higher rank in unsuccessful negotiations of the general bigram *the offer* also supports the claim.

The absence of *new, this, last* in negotiation-related bigrams in the *SimpleNS* data correspond to the shorter negotiation time. The participants exchange fewer offers than in *Inspire* negotiations, thus do not need to emphasize which offer they speak about.

A summary of the results follows:

- patterns indicating the intensity of interactions are more frequent in successful negotiations than in unsuccessful negotiations;
- patterns indicating specific issues of negotiation are more frequent in successful negotiations than in unsuccessful negotiations;
- patterns clarifying negotiation moves are more frequent in successful than in unsuccessful negotiations.
- patterns indicating the intensity of interactions are more frequent in long-term negotiations than in short-term negotiations;

The results support our hypothesis that participants in successful negotiations show interest in continuing negotiations; for further analysis see Sections 8.2 and 8.4. The results are generalized in the chapter's conclusion.

6.5 Language patterns and argumentation, persuasion, substantiation

We have shown in section 3.6.1 that argumentation, persuasion and substantiation partially correspond to the use of modals. We also showed that negotiators from different classes vary in the use of modals (see Section 8.2). A more detailed comparison of the patterns shows that buyers and sellers in successful negotiations use **must** and **will** more with self-obligation and self-intention (Leech, 1987) than to express insistence or authority over their partners. As for the modal **can**, self-referring *I can* happens more often than *you/it can* in all four classes. In successful negotiations the impersonal pattern *it can/must/will* appears more often than in unsuccessful negotiations.

Our next goal is to compare N-grams in which modal verbs and the verb *have* appear in successful and unsuccessful negotiations. We look for patterns most frequent in their classes. As expected, the verbs "have" and "will" appear more frequently in unsuccessful than in successful negotiations, while the verb "can" is more frequent in successful than in unsuccessful negotiations.

We also compare the *Inspire* pattern ranks with the *SimpleNS* pattern ranks. The *SimpleNS* pattern ranks are substantially lower. We attribute this to much shorter duration of a *SimpleNS* negotiation. This shortness does not provide enough time for persuasion, argumentation and substantiation, thus simplifying the negotiation strategies of the participants. See Table 6.5 for results.

word	***Inspire***				***SimpleNS***	
	trigram	rank_s	trigram	rank_u	trigram	rank
have	I *have* to	49	I *have* to	36	I *have* to	59
	that we *have*	55	we *have* to	43	we *have* to	75
	we *have* to	66	that you *have*	75	that we *have*	102
	that you *have*	92	that we *have*	92	we *have* a	151
	to *have* a	123	will *have* to	162	you *have* to	189
can	that we *can*	16	that we *can*	16	that we *can*	22
	hope we *can*	62	hope we *can*	74	if you *can*	153
	hope you *can*	73	hope you *can*	99	hope you *can*	202
	you *can* accept	89	if you *can*	117	hope we *can*	205
	that you *can*	118	that you *can*	147	we *can* find	239
will	that you *will*	36	that you *will*	22	that we *will*	50
	hope you *will*	53	hope you *will*	30	*will* be able	93
	I *will* be	69	you *will* find	44	hope you *will*	104
	that we *will*	83	I *will* be	52	we *will* be	107
	will be able	85	we *willl* be	93	that you *will*	127

Table 6.5: The most frequent trigrams with verbs *have, can, will.*

We look for trigrams frequent in one class and rare in another class of negotiations and calculate the distance between ranks. We report the results in Table 6.6. A negative distance means a trigram more frequent in successful negotiations, a positive distance – a trigram more frequent in unsuccessful negotiations.

word	**trigram**	d^3	**trigram**	d^3
have	we will have	-213	you have any	-163
	if you have	-119	offer I have	96
can	we can come	166	that I can	-106
	can come to	95	you can accept	-94
cannot	I cannot accept	-97	we can not	127
	can not accept	51		
will	and I will	125	will not be	67
	you will be	-52		
would	we would like	52	I would be	-28

Table 6.6: Trigrams with *have, can, cannot, will, would*, the largest differences in ranks

Consider the distances of trigrams with the pronoun "I": we posit that the negotiators' assertiveness is higher in successful than in unsuccessful negotiations. Trigrams that suggest uncertainty and rejection are indicative of unsuccessful negotiations. This supports our claim stated in section 6.4 about the intensity of moves and clarity of the language in successful and unsuccessful negotiations. The results are generalized in the chapter's conclusion.

6.6 Language patterns and politeness

Negotiators employ different strategies to implement politeness in language. In Section 3.6 we have listed such politeness strategies as the use of mental verbs, hedging opinions, indirect expressions and shown their many relations with the negotiation and influence strategies. In this section we discuss how negotiators use formal polite words in standard e-negotiation procedures; refer to Section 3.6.1 for the brief overview.

One of the business interaction rules is a well-established introduction and closure of each interaction, which vary across cultures. In e-negotiation one interaction is equal to a message. That is why the opening of a message denotes introduction and the ending of a message denotes closure. CMC brings in additional variety to the language used in introductions and closures. Thus, to avoid investigating several small threads and to make our research feasible, we consider one aspect of the closure, namely politeness.

We note the fact that the highest-ranked bigram *thank you* for the unigram *thank* is the tenth most frequent in successful negotiations and fourteenth in unsuccessful negotiations and the highes-ranked bigram *thanks for* for the unigram *thanks* is ranked 94 in successful negotiations and 104 in unsuccessful negotiations. The smaller rank means that the bigrams appears more often in successful negotiations than in unsuccessful ones. The hypothesis that naturally arises from this observation is that politeness is a characteristic feature of successful negotiations. We have calculated the combined percentage of the orthographic variations of the words *thank*, *thanks*. It is twice higher in successful than in unsuccessful negotiations! In Table 6.7 we report the ranks of most frequent trigrams with the politeness indicator words. The results show that the *SimpleNS* data predictably contain fewer politeness indicator words than the *Inspire data*. We explain this difference by the fact that the politeness indicator words appear either at the beginning or at the end of messages. The lack of politeness indicators corresponds to a lower number of messages sent in one negotiation.

Inspire				***SimpleNS***	
trigram	rank_s	trigram	rank_u	trigram	rank
thank you for	1	*thank* you for	2	*thank* you for	2
thanks for your	20	*thanks* for your	19	*thanks* for your	41
thank you very	99	*thank* you very	162	*thank* you very	136
to *thank* you	330	*thanks* for the	580	*thanks* a lot	521
thanks for the	448	offer *thank* you	583	i *thank* you	910

Table 6.7: Five most frequent trigrams with *thank* and *thanks*, *Inspire*, *SimpleNS*

Predictably, negation is present more often in unsuccessful than in successful negotiations: *not* ($\text{rank}_s = 30$, $\text{rank}_u = 26$). We compare the ranks of the most frequent bigram (both in successful and unsuccessful negotiations) with negation – *is not* ($\text{rank}_s = 95$, $\text{rank}_u = 65$). As expected, the "negative" bigram is more frequent in unsuccessful negotiations.

Another indicator of politeness is the word **best** when it appears in **best regards**, **all the best** and **best wishes**. These expressions are typical of the closing part of a polite written conversation. For buyers and sellers, these contexts appear more often in successful than in unsuccessful negotiations; see Table 6.8. This supports our hypothesis about the use of polite words in successful negotiations.

Expression	BSN	BUN	SSN	SUN
best regards	+	-	+	-
all the best	+	-	+	-
best wishes	+	-	+	-

Table 6.8: Distribution of expressions with *best*, successful/unsuccessful buyers, successful/unsuccessful sellers

However, the hypothesis about polite words in successful and unsuccessful negotiations has not received a strong support when we divided data into the final messages and the rest of negotiations. The final messages include the message in which negotiators announce their decision on reaching or rejecting a deal. If there are other messages sent later then they are also included in the final part. Without the final messages successful negotiations do not have more polite words than unsuccessful negotiations. That is why we have to reject the hypothesis about polite words in successful and unsuccessful negotiations.

6.7 Learning language patterns from N-gram ranks

We are looking for trigrams that show the negotiators' goal (win by any means, reach a compromise, do away with the assignment), their attitude to partners (friendliness, aggressiveness, indifference), and behaviour in the negotiation process (flexibility, stubbornness). The same trigrams must be noticeably present in either successful or unsuccessful negotiations. As we said before, these trigrams have a large difference between their ranks in successful and unsuccessful negotiations. Hence two major elements affect the N-gram selection: words contained in an N-gram and its rank.

The idea behind finding of N-grams representative for each corpora is quite simple. It is a bootstrapping procedure which learns from a small number of words corresponding to the basic negotiation moves, such as agreement, refusal, negotiating issues. These words are called seeds. The procedure allows us to learn differences in corpora of successful and unsuccessful negotiations.

A bootstrapping procedure of building the lists of representative N-grams for successful and unsuccessful negotiations

Input: text data of all negotiations, text data of successful negotiations, text data of unsuccessful negotiations, seeds.

1. Build the list L of unigrams for all negotiations
2. Build the lists of N-grams ($N = 1, 2, 3$) for successful negotiations (NS).
3. Build the lists of N-grams ($N = 1, 2, 3$) for unsuccessful negotiations (NU).
4. In L find unigrams of seeds among k most frequent unigrams (k is a predefined cut-off point). Build the list W of such seeds.
5. For each $w \in W$:
   - Find its rank r_s^1 in the list of the unigrams of NS.
   - Find its rank r_u^1 in the list of the unigrams of NU.
   - Calculate $d_w^1 = r_s^1 - r_u^1$.
6. Delete from W all w such that $d_w^1 < d$ (d is a predefined distance).
7. For each $w \in W$:
   - Find its bigrams among m most frequent bigrams on the list of bigrams of NS (m is a predefined cut-off point).
   - Find its bigrams among m most frequent bigrams on the list of bigrams of NU.
   - - Find the rank r_s^2 of the $i - th$ bigram on the list of the bigrams of NS.
     - Find the rank r_u^2 of the $i - th$ bigram on the list of the bigrams of NU.
     - Calculate $d_i^2 = r_s^2 - r_u^2$.
   - Calculate $d_w^2 = \sum_{i=1}^{m} d_i^2$
8. Delete from W all w such that $d_w^2 < d$.
9. Find most frequent trigrams containing unigrams from W: repeat steps 7-8 for trigrams instead of bigrams.
10. Build the list L_R of trigrams, containing $w \in W$, with their ranks.

Output: L_R.

Although it is possible to investigate N-grams with $N > 3$, the procedure stops at trigrams because of the data characteristics listed earlier: simplified grammar, dense and short sentences, restricted lexicon. The only adjustable parameters are the distance d and the cut-off points k, m. In order not to overload the procedure, we do not use weights to tune distances between N-grams, though it seems a natural thing to do. We have tested the procedure with $d = min(100, 2 * rank_s)$, $k = 100$, $m = 700$. To find k, m we have chosen the values that guarantee that the procedure works with N-grams covering the same percentage of texts in both successful and unsuccessful negotiations and eliminate low-frequency N-grams, thus keeping representative N-grams in negotiation data. Needless to say, the choice of distance depends on the cut-off points. For our cut-off points, the distance ensures that the ranks used to calculate it correspond to different N-gram frequencies. Examples from the resulting list are presented in Table 6.9.

word	**trigram**	rank_s	**trigram**	rank_u
have	we *have* to	66	that you *have*	75
	that you *have*	92	that we *have*	92
accept	to *accept* your	55	to *accept* your	103
	you can *accept*	90	you will *accept*	132
agree	*agree* with your	395	you will *agree*	533
	I *agree* with	426	*agree* with you	565
will	I *will* be	69	you *will* find	44
	that we *will*	83	I *will* be	52

Table 6.9: Examples of representative trigrams.

We notice that in the trigrams from the unsuccessful negotiations there is a trace of aggressive behaviour (**you will accept, you will agree**), which is absent from the corresponding trigrams in the successful negotiations (**you can accept, agree with your**). Tracing the trigrams with "you", we found that in successful negotiations they correspond to politeness, in unsuccessful negotiations – to aggressiveness. The results of this section are generalized in Section 6.9.

6.8 Language patterns in face-to-face-negotiation data

In this section we compare the language patterns in ftf-negotiation data (*Cartoon*) and e-negotiation data (*Inspire*). Recall that we only have access to successful *Cartoon* negotiations. Thus the language patterns in ftf-negotiation data are compared with those in the successful e-negotiation data (*Inspire*-success).

6.8.1 Language patterns in successful negotiations

Table 6.10 shows that lacking the separate formal offers makes the ftf texts overloaded with specific information such as the negotiation issues and exact numbers, e.g. ultra rangers, 5 million, 8 runs, etc. The same table shows that in *Inspire* - success negotiations, where the formal offers exist, the texts refer to the whole offer more often than to the separate issues.

Inspire-success				*Cartoon*			
bigram	rank	bigram	rank	bigram	rank	bigram	rank
your *offer*	2	my *offer*	39	per *episode*	3	in *terms*	35
the price	5	this *offer*	41	*ultra rangers*	17	*down payment*	36
offer I	24	last *offer*	45	*up front*	26	8 *runs*	46
offer and	32	*accept* your	48	*terms* of	28	*number* of	50
new *offer*	37	to *accept*	57	5 *million*	29	*okay*	53

Table 6.10: Bigrams of the most frequent negotiation-related words, *Inspire*-success, *Cartoon*.

In Tables 6.11 and 6.12 we compare bigrams of the verbs *can* and *have* in *Cartoon* and *Inspire*-success data. We omit bigrams with *will* because its use in *Cartoon* data is negligible; for the verb distribution see Table 8.2 in Appendix C.

The bigrams show that in e-negotiations participants apply to, or imply about, the counterpart plans more often than it is done in ftf-negotiations. We attribute this to the fact that in ftf-negotiations participants can deduct, or guess, the counterpart opinion on some aspect of negotiations from non-verbal behaviour and body language. In e-negotiations participants have to ask questions, either explicitly or implicitly, to gain necessary information. However, the presence of the official offer in *Inspire*-success data reduces the number of direct questions about the negotiation issues. This

Inspire-success		*Cartoon*	
trigram	rank	trigram	rank
that we *can*	16	if we *can*	32
hope we *can*	62	that we *can*	50
hope you *can*	73	we *can* do	64
you *can* accept	89	think we *can*	113
that you *can*	118	we *can* work	128

Table 6.11: The most frequent trigrams with *can*, *Inspire*-success, *Cartoon*.

Inspire-success		*Cartoon*	
trigram	rank	trigram	rank
I *have* to	49	do you *have*	31
that we *have*	55	that we *have*	35
we *have* to	66	we *have* a	49
that you *have*	92	we *have* to	65
to *have* a	123	going to *have*	104

Table 6.12: The most frequent trigrams with *have*, *Inspire*-success, *Cartoon*.

conclusion is supported by high ranking of the direct question *do you have* in the *Cartoon* data.

In Table 6.13 we compare the use of formal polite words in both data. To simplify comparison we make capitalization uniform. The results show that ftf-negotiation data contain the formal polite words (substantially) less than e-negotiation data. Partly this corresponds to the use of less formal *okay* in the *Cartoon* data instead of formal *accept* in the *Inspire*-success data; see Table 6.10. Partly we attribute the difference to incompleteness of the *Cartoon* data where transcripts often omit the introductory and post-settlement phases. The most striking difference appears in the ranks of *thank you for* which is the most frequent trigram in the *Inspire*-success data and is negligibly used in the *Cartoon* data. From studies on the *Inspire*-success data we know that the use of *thank* and its word forms is heavily concentrated at the end of successful negotiations.

Inspire-success		*Cartoon*	
trigram	rank	trigram	rank
thank you for	1	(pause) *thank* you	900
thanks for your	20	*thank* you very	1586
thank you very	99	*thank* you (pause)	2940
to *thank* you	330	*thank* you for	3326
thanks for the	448	*thank* you i	4289

Table 6.13: Five most frequent trigrams with *thank* and *thanks*, *Inspire*-success, *Cartoon*.

The provided comparison support the conclusion that the language used in e-negotiations differs from the language used in ftf-negotiations. Such differences are partly due to the form of communication – written versus personal – and partly due to the form of negotiation – with negotiation support or without it.

6.8.2 Language patterns in buyer and seller roles

We find the (non-directional) indicative words for buyers and sellers in the *Cartoon* negotiations; see Table 6.14 and the list of all the indicative words in Appendix B, page 229. As expected, the indicative words in *Cartoon* do not contain modals except for *might* and *would.* Nor do they contain as many adjectives as the indicative words for buyers and sellers in the *Inspire*-success data. However, there is a group of words present among the *Cartoon* indicative words and absent among the *Inspire*-success indicative words. The verbs in the progressive form such as *paying, talking, buying,* etc, make up 7% of the *Cartoon* indicative words and 0% of the *Inspire* indicative words.

Buyers and Sellers	skip, paying, you, got, my, we've, certainly, sell, might, also

Table 6.14: Ten indicative words with the highest *LL*.

Based on this observation we select the language patterns representative for buyers and sellers in *Cartoon* negotiations. Table 6.15 shows which patterns represent buyers and which patterns represent sellers in the *Cartoon* data. *BE* denotes all forms of the verb *be*, (...) denotes an optional word. The distribution of patterns clearly shows that buyers use patterns with active verbs more often than sellers whereas the opposite is true for the patterns with modals.

Pattern	Buyer	Seller
PersPron **might** *Verb*	-	+
that (*PersPron*) **might** *Verb*	-	+
PersPron **would** *Verb*	-	+
that (*PersPron*) **would** *Verb*	-	+
PersPron (*BE*) **paying**	+	-
talking/thinking about	+	-
PersPron **talking/thinking**	+	-
thinking ***Prep/Adv***	+	-

Table 6.15: Five most frequent trigrams with *thank* and *thanks*, buyers/sellers

Comparing *Cartoon* buyer and seller data we have confirmed the previously stated hypothesis that the language used by the buyers and sellers reflects their differences within negotiations.

The results of this section are analyzed in Section 6.9.

6.9 Conclusion

In this chapter we used statistical methods to extract information characteristic for different data classes.

We aim to find the relations between the negotiation and influence strategies and the language presented in Chapter 3. We learn how these strategies are implemented through the language patterns in different data classes. The patterns discovered support our hypotheses on the difference in negotiators' behaviour in successful and unsuccessful negotiations. In addition, we have stated a hypothesis about the difference in sellers' and buyers' interests in negotiations. This hypothesis corresponds to hypotheses about the imbalance of sellers' and buyers' relations in face-to-face negotiations. We also have supported the last hypothesis by providing statistical information from the face-to-face *Cartoon* negotiations. We have shown that language patterns used in ftf-negotiations differ from language patterns used in e-negotiations. As an additional result, we have shown that language patterns indicating the negotiation strategies are easier to find in long-term negotiations than in short-term negotiations.

We have presented two automated methods for learning from corpora. The methods are used to recognize and identify language patterns corresponding to the recorded outcomes and/or well-established roles in e-negotiations. One method is based on using the indicative words as pattern seeds, the other on N-gram models. By applying two methods to the data classes we have shown that different patterns exist for different role bearers and for different negotiation outcomes. The patterns found highlight the differences between negotiators in successful and unsuccessful negotiations, between buyers and sellers, between buyers in successful and unsuccessful negotiations, and between sellers in successful and unsuccessful negotiations. The consistency of the results obtained by the methods is well seen in the language patterns they find. Both methods, *learning entirely from the text data*, find the language patterns corresponding to the strategies described in Chapter 3.

We have compared language patterns obtained by both methods. The comparison shows that the sets of patterns complement each other. We conclude this chapter with the following summary of the obtained patterns:

- patterns indicating formal politeness are substantially more frequent in e-negotiation than in ftf-negotiations;
- patterns indicating the intensity of interactions are more frequent in successful negotiations than in unsuccessful negotiations;
- patterns indicating specific issues of negotiation are more frequent in successful negotiations than in unsuccessful negotiations; however, the general use of patterns depends on the available negotiation support;
- patterns indicating requirement and obligation are more frequent in unsuccessful negotiations than in successful negotiations;
- patterns indicating persuasion are used more often by sellers than by buyers;
- patterns indicating possibility of their own actions are used more often by sellers than by buyers;
- patterns indicating self-obligation are used more often by buyers than by sellers;
- patterns indicating impersonal requirements are used more often in successful than in unsuccessful negotiations;
- patterns clarifying negotiation moves are more frequent in successful than in unsuccessful negotiations;
- participants refer to their own necessary actions more often than they require actions from counterparts; however, the ratio of self-requirement to request-to-others is higher in successful negotiations than in unsuccessful negotiations.

The learned patterns can be employed by NSSs in the negotiation phase. Although it seems beneficial to provide the language patterns used in successful negotiations to the future participants in the pre-negotiation phase, we think that the language patterns used separately, without the actual exchange of offers, may not help in preparation for negotiations. However, spotting language patterns in on-going negotiation can prove to be useful as a "predictor" for the probable outcome.

Chapter 7

Negotiation-related Features and E-negotiations

By definition, when you are investigating the unknown, you do not know what you will find.

Murphy's Law

7.1 Introduction

In this chapter we continue to work with semantic categories started in Chapter 5. We use a feature selection method based on the previously introduced semantic categories. We hypothesize that the data can be classified accurately if represented by a negotiation-related data subset. We use a semi-automatically built set of features representing the e-negotiation text data through words tagged as negotiation-related (negotiation-related features). Using such features, we apply supervised ML methods in analyzing and classifying e-negotiation texts according to the success and failure of a negotiation or buyer and seller roles. In separate set of experiments we represent data by words tagged with Hobbies, Studies and some other, not negotiation-related, semantic categories. These features are built semi-automatically.

We run experiments on the data of different sizes to see how the size of the data affects the classification results. We also solve a multi-class classification problem in which successful buyers, successful sellers, unsuccessful buyers and unsuccessful sellers are the class labels. In all experimental settings we compare the classification results for semi-automatically built features and for automatically built features.

We explore the classification performance of ML methods and their learning ability on e-negotiation data. We also show that negotiation-related features can be used for reliable representation of ftf-negotiation data.

Our working hypothesis is as follows:

- the subject of discussion during negotiations provides a reliable ground for learning success of negotiations.

We continue to work with feature selection and compare negotiation-related features with another set of knowledge-based features in Chapter 8. Partially the chapter results are used to support or reject hypotheses listed in Section 9.3 and to develop new tasks of the e-negotiation studies in Section 9.5.

7.2 Data representation by negotiation-related features

In this section we experimentally verify our claim that the e-negotiation data can be represented by a subset of negotiation-related word types. To do this, we form a bag of *negotiation-related* words that we identified by the procedure described in section 5.4.2.

As the first step of building such a bag of negotiation-related words, we remove all unigrams with fewer than 6 occurrences. The cut-off used is suggested by the Katz smoothing model which gave relatively low cross-entropy on the data; see section 5.2. The removal of unigrams with fewer than 6 occurrences also serves to remove the personal information from the data, as well as the rare words which might not be statistically representative of the data. We concentrate on the negotiation-related information and do not investigate in depth the distribution of personal information. However, the general study of the data suggests that the presence of a personal email address is a trustworthy indicator of the personal nature of (part of) an *Inspire* message. Email addresses are usually exchanged when the partners perform self-disclosure. We extracted 512 negotiations that contained personal email addresses and tested the distribution of the occurrences of unigrams corresponding to personal information. 90% of such unigrams had fewer than 6 occurrences. Katz model gives our next occurrence cut-off which is 18. Recall that the model increases its cross-entropy, i.e. declines its fitness, when threshold is more than 18.

This gives us a data set of negotiation-related words of dimensionality 123; see Figure 7.1.

offer price delivery accept hope days time payment think company know quality negotiation returns terms business agreement return Cypress Itex agree products deal policy full product parts pay spoilage response reply day refund thanks answer order relationship counter conditions thank negotiations term production negotiate interested received market issue gear cost point assemblies deliver offers line issues bikes contract receive companies supplier compromise part suppliers negotiating proposal needs costs respond counteroffer customers decided interest period package accepted problems firm option deadline increase contact delay offered consideration money concessions condition reduce unacceptable options offering beneficial benefit prepared rear mutually reconsider bike manufacturing recent system considering continue prices partner decision purchase shipment profit considered satisfied proposed solution concern schedule concerns concerning negociation bicycles risk requirements accepting

Figure 7.1: Negotiation-related words in Inspire.

We call the negotiation-related words *negotiation-related features.* We add to this a count of the number of unigrams in each example that do not belong to the *negotiation-related* semantic category. Hence, for each example we have a bag of words with 124 attributes. Each of the first 123 attributes in each example represents the number of occurrences of the corresponding unigram from the negotiation-related category while the last attribute gives the total number of other unigrams present in the example. Each example is labelled positive if the corresponding negotiation resulted in success and negative otherwise.

7.3 Classification results obtained on the negotiation-related features

In this section we report data parameters for ML experiments and show that the negotiation-related features give a reliable classification accuracy when the negotiation outcomes are classified.

7.3.1 Data parameters

We have a total of 2557 negotiations in our data set, of which 1427 are successful and 1130 unsuccessful. The average length of a successful negotiation is 752 tokens and the average length of an unsuccessful negotiation is 443 tokens. The reported sizes are much smaller than the usual size of texts (Shawe-Taylor and Cristianini, 2004).

At least 3% of unsuccessful data are questionably labelled by the system, and the classification results are corrupted by this noise; see Section 4.1 for details. Recall also that on the pre-processed data the negotiation outcomes are classified with 75%-76% accuracy when are represented by the history records; see Section 3.4.

We consider the whole negotiation to be a data unit. Our reasons are the following: a negotiation is formally defined, independent and self-complete. We prefer a negotiation to lesser units, e.g. sentences or phrases, to avoid controversy inherited from speech; refer to Section 3.5 for explanations. We decide in favour of a negotiation towards a message partially because a system most commonly labels the whole negotiation. Another reason lies in a connection of a message with other messages within a negotiation. Such connection prevents us to consider message as a separate unit.

We concatenate all messages belonging to the same negotiation, unifying capitalization and labelling the negotiation with the label given by *Inspire*. Each negotiation becomes an example, either positive (successful) or negative (unsuccessful).

7.3.2 Classification results

We report the tenfold cross-validation results for all the experiments. These results were obtained over the set of parameters for each classifier that yielded the highest classification accuracy. The parameters were selected using an exhaustive search over the space of possible parameters. We have employed several classifiers freely available in the Weka suite (Witten and Frank, 2000): Instance-based using 20-nearest neighbor (IBK), Naive Bayes classifier (NB), Decision Stumps (DS), Decision Tables (DTb) and linear SVM. For decision trees we have used C5.0 (Quinlan, 1993). We used the Decision List Machine (DLM) as a data-based classifier (Sokolova et al., 2003). BL indicates when all the examples are labelled as positives.

The accuracy results appear in Table 7.1. We present the best accuracy achieved by each classifier after we have performed exhaustive search on adjustable parameters. Precision, recall and F-measure are calculated with respect to the successful negotiations. To maintain readability of this and following sections we report statistical significance in a separate Section 7.6.

Classifier	Accuracy	Precision	Recall	F-measure
BL	55.8	55.8	100	71.6
NB	66.1	75	59	66
DTb	73.9	71.2	88.3	78.6
IBK	70.6	**76**	69	72
SVM	71.72	75.8	72.5	74
C5.0	72.39	71.2	87.2	78.4
DS	73.13	71.4	85.6	77.8
DLM	**75**	69.4	**93.9**	**79.6**

Table 7.1: The classification accuracy, successful/unsuccessful negotiations; representation by the negotiation-related words

The results show that the overall accuracy is similar for most of the employed classifiers and to those obtained on the history records (refer to Section 4.2.1).

7.4 Comparison with other feature selection methods

In this section we compare negotiation-related features with other features of e-negotiation data. With a set of classification experiments involving various features we show that the negotiation-related features are necessary and sufficient. Note that from now on we will use the following four classifiers: NB, C5.0, SVM, and DLM. By eliminating DS and DTb we avoid redundancy. Their learning paradigms repeat those of C5.0 and DLM respectively. We will use IBK when applicable..

7.4.1 Automatically selected features

Now, in order to verify our claim that the data representation using only the negotiation-related words is sufficient, we perform the following three sets of experiments on automatically selected features. We compare classification results on features capturing similarities in two data sets and indicating differences between them.

In the first set of experiments, we represent the data using the ***500 most frequent unigrams.*** That is, we form bags of words for each negotiation using the number of occurrences of the 500 most frequent unigrams over the whole set of negotiations. We use this set of features because most of 500 words are similarly present in the successful and unsuccessful negotiation data. We perform tenfold for NB and C5.0 and fivefold for SVM cross validation and report the corresponding results in Table 7.2.

Classifier	Accuracy	Precision	Recall	F-measure
BL	55.8	55.8	100	71.6
NB	63.4	71.2	46.4	55.8
SVM	71.7	76.8	70.6	73.6
C5.0	74.3	73.2	85.2	78.6

Table 7.2: The classification accuracy, successful/unsuccessful negotiations; representation by the 500 most frequent words

Next we represent the data using unigrams selected by the ***Best First*** feature filter. The filter selects features based on hill climbing with forward search. Hill climbing conducts a local search for features (Witten and Frank, 2000). Forward search means that a feature set begins as an empty set and grows one feature at a time. On each iteration the best among unselected features is found and added to the feature set. The evaluation of features is done by calculating correlation between a feature and a class. These features show local differences between two data. We use the filter to select features from 1500 most frequent ones. 66 words selected by the Best First filter show that the filter benefits high-frequency function words such as *this, on, but, at,* etc. We form bags of words for each negotiation using the number of occurrences of features selected by the filter. We perform tenfold cross-validation and report the corresponding results for the whole set in Table 7.3.

For the third set of experiments we represent the data by the unigrams selected through ***the Information Gain*** filter. This filter ranks features based on their information gain. These features show overall differences between two data. We use the filter to select features from 1500 most frequent ones. The words ranked highest by the filter show that this filter also benefits from high-frequency function words such as *this, on, but, at,* etc. We further work with the top 123 features so as to have a fair comparison with the negotiation-related words; see Section 7.2.

Classifier	Accuracy	Precision	Recall	F-measure
BL	55.8	55.8	100	71.6
NB	62.4	79.4	49.8	59.6
C5.0	73.5	71.2	88.32	79
SVM	70.6	74.8	71.5	73.8
DLM	71.4	69.8	85.9	77.2

Table 7.3: The classification accuracy, successful/unsuccessful negotiations; representation by the *Best-First* features.

We form bags of words for each negotiation using the number of occurrences of features selected by the filter. We perform tenfold cross-validation and report the corresponding results for the *Inspire* dataset in Table 7.4.

Classifier	Accuracy	Precision	Recall	F-measure
BL	55.8	55.8	100	71.6
NB	65.3	74.7	57.4	64.9
DT	71.8	70.4	85.2	77.1
SVM	71.73	75.1	73.9	74.5
DLM	70.8	70.3	82.5	75.9

Table 7.4: The classification accuracy, successful/unsuccessful negotiations; representation by the *Information Gain* features.

Statistical generalization of the results follows in Section 7.6.

7.4.2 Generic features

We consider two types of features found by knowledge-based feature selection methods; see Section 2.4 for other references. The negotiation-related words capture the specifics of the negotiation process. They represent what we call *process-specific* features. At the same time, there are features[1] that can be found in data gathered from other communication. We call such features *generic*. In our data generic features are the words tagged with semantic categories that are not negotiation-related e.g.,

[1]Except function words

Hobbies, Studies, Place names. We call them "casual talk" words because they appear in casual discussion during negotiations. An example of casual talk in e-negotiation messages:

Hallo Miriam, thanks for your offer, as I wasn't quite happy with it I'm sending you another one. *I am 23 year old girl, currently living in Australia. Prior arriving here I have lived in Germany for 7 1/2 years, in Heilbronn - near Stuttgart. What town do you live in?* Hoffentlich magst Du dies Angebot besser. Bis bald Nesta

Hello Nesta, i'm glad about your quick message. *First I want to tell you some more facts about me. My birthplace is Freiberg in saxony- east Germany. Today I live in Chemnitz- also in Saxony. Please send me if you like this your Email address for a private contact.* Some words to your offer: I think we have to fight for a better result before we come to a solution. May be it won't be so long to this. Have a nice week, Ciao Miriam PS: Ich hoffe, du findest das Angebot nicht so schlimm.

Casual talk is usually marked by email and place addresses. It represents the life style and interests of negotiators, hence for the Inspire data can be detected by names of schools, e.g. *the State University of NY at Buffalo, Carleton University in Ottawa, UTE of Ecuador, Hong Kong Baptist University* and programs, i.e. *business admistration, international management, Financial Ingenier, Criminology, Computer Science*, and hobbies, e.g. *I like travelling, I like playing basketball, Are you collecting stamps?*, or personal details *I'm actually home with the flu, When I was driving my rental car down to Toronto from Ottawa on Friday night, I had an minar but serious accident!!, Sometimes I had a home sick, I had a wonderful weekend in Berlin, I am a male student.*

We show that "casual talk" appears very often and in different negotiations with the distribution of a few words which correspond to socializing and personal information. Table 7.5 shows the number of occurrences and ranks of the words. Note that each of the words usually is used only once by a negotiator, hence the number of occurrences shows how many negotiators have revealed this type of personal information during negotiation process. The rank of the word shows its relative use among other words. The number of types without rare words is 13890, hence the listed ranks are high. In fact, the ranks of "university" and "student" place them among function and common words and negotiation-related words.

In the next run of experiments we represent the data using a set of casual words. There are 135 words. We want to see how well the casual talk words allow reliable classification of the negotiation outcomes. The classification results are reported in Table 7.6.

The accuracy results show that a relatively comparable (in fact marginally better) accuracy can be obtained when only negotiation-related knowledge is used to represent the data. We have shown

Word	Occurrence	Rank	Word	Occurrence	Rank
university	435	380	program	121	927
student	418	393	male	77	1201
live	378	435	female	59	1436
school	223	623	travelling	47	1684

Table 7.5: Indicators of casual talk and their occurrences

Classifier	Accuracy	Precision	Recall	F-measure
BL	55.8	55.8	100	71.6
NB	65.3	74	58.2	65.2
C5.0	72.6	70	89	78.6
SVM	66.5	72.4	64.5	68.2
DLM	71.1	68.8	87.8	77.4

Table 7.6: The classification accuracy, successful/unsuccessful negotiations; representation by the casual talk words

that casual talk words do not provide the same reliability of the outcome classification as negotiation-related words. We have shown that the representation with negotiation-related words is necessary for the accurate classification. To support our claim we represented the data using semantic categories other than negotiation-related. This representation did not allow the same classification accuracy as negotiation-related word. Statistical generalization of the results follows in Section 7.6.

7.5 Classification of the data sets of different sizes

In this section we show that negotiation-related words are a reliable data representation for the data sets of different sizes. We have compared the accuracy obtained by the classifiers on the data sets of the following sizes: 50, 100, 500, and 1000 examples. The data sets are constructed from the *Inspire* data. The results presented in Figures 7.2 – 7.9 report the average accuracy obtained by each classifier on randomly chosen data set of corresponding sizes. The highest accuracy (almost) always is obtained for the negotiation-related words. This is natural: the feature selection does not depend on the data set size.

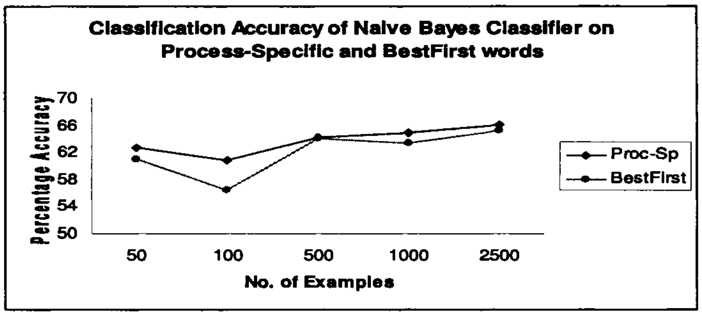

Figure 7.2: Accuracy of NB on data sets with different sizes; negotiation-related and Best-First representations

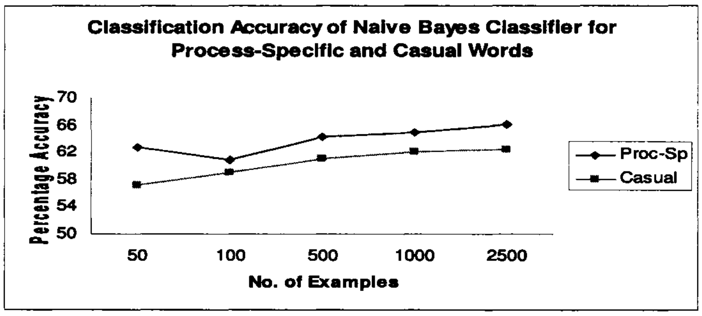

Figure 7.3: Accuracy of NB on data sets with different sizes; negotiation-related and casual representations

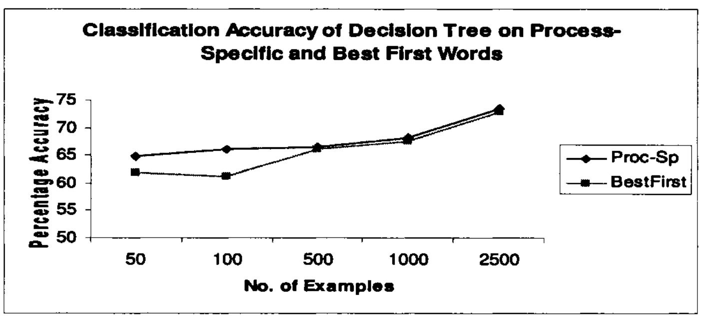

Figure 7.4: Accuracy of C5.0 on data sets with different sizes; negotiation-related and Best-First representations

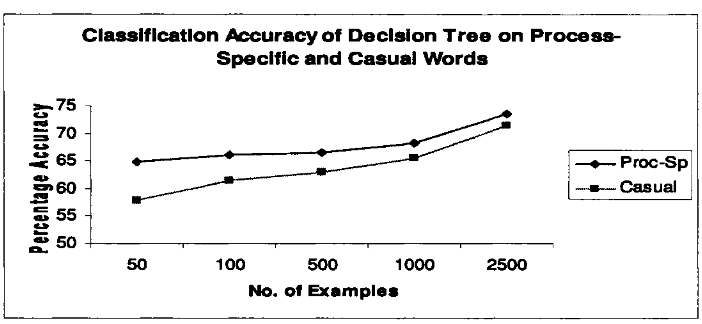

Figure 7.5: Accuracy of C5.0 on data sets with different sizes; negotiation-related and casual representations

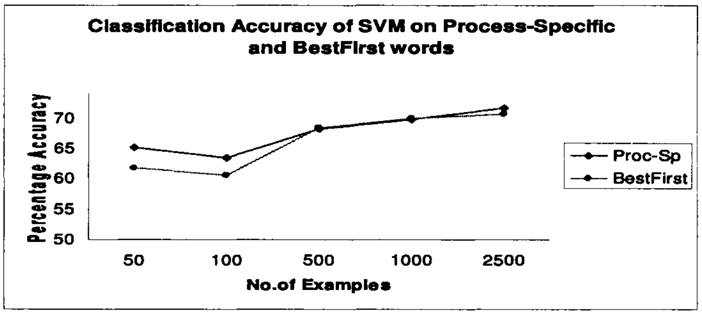

Figure 7.6: Accuracy of SVM on data sets with different sizes; negotiation-related and Best-First representations

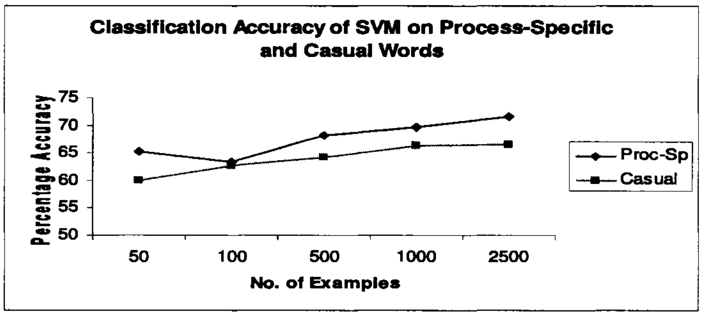

Figure 7.7: Accuracy of SVM on data sets with different sizes; negotiation-related and casual representations

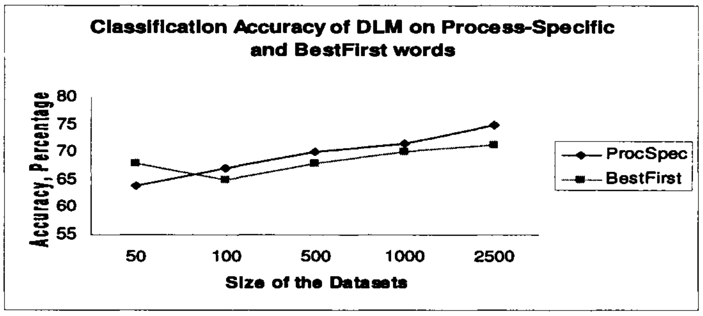

Figure 7.8: Accuracy of DLM on data sets with different sizes; negotiation-related and Best-First representations

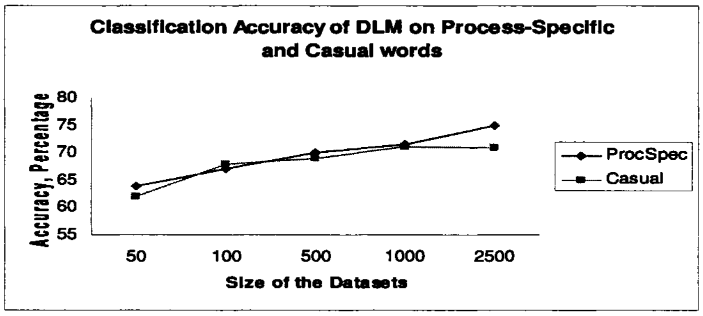

Figure 7.9: Accuracy of DLM on data sets with different sizes; negotiation-related and casual representations

7.6 Statistical evaluation of the results

We have also investigated the statistical significance of our classification accuracy results. This was done in three ways. The accuracy results that we have presented are with respect to tenfold cross-validation testing. Even though the results are averaged over 10 different runs of learning algorithm on randomly obtained subsets of data, a random choice of training and testing subsets might incur an imbalanced distribution of training instances. Hence, the standard deviation of accuracy over various folds gives a better estimate of our evaluation. In order to investigate this, we first provide a comparison of standard deviation measure computed for all classifiers for different feature selection methods. See Figure 7.10 for a comparison of standard deviations over the classification accuracies with respect to various feature selection methods. Second, Figure 7.11 details the comparison of standard deviations of classification accuracies with respect to various classifiers. As can be observed in plots of Figures 7.10 and 7.11, the results are relatively stable (lower variance) in case of classifier accuracies when negotiation-related representation of the e-negotiation text is applied. It should be noted that this stability is also seen for generic features. However, the accuracy results of the classifiers on data represented using these features is quite poor.

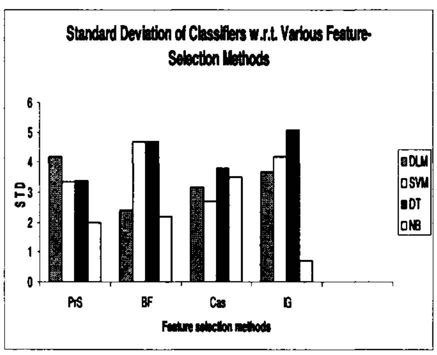

Figure 7.10: Standard deviation of accuracy for feature selectors

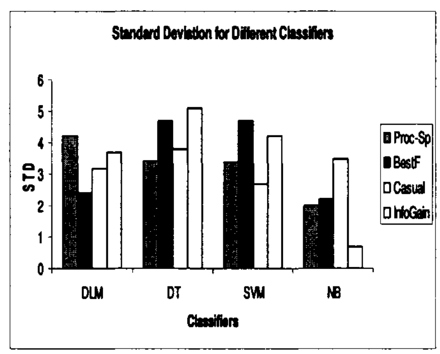

Figure 7.11: Standard deviation of accuracy for four classifiers

Moreover, as we can see in plots of Figures 7.2 – 7.9, the performance of the classifiers in terms of the classification accuracy is generally better when the data is represented using negotiation-related features (esp. for large and more practical dataset sizes). In the case of DLMs, the results are better by a wide margin than the other results while for other classifiers the gain is sometimes marginal. In order to investigate this, we also perform t-test on the accuracy results obtained by different classifiers on the classification of negotiation outcomes. For all classifiers on all data sets when compared with respect to the negotiation-related representation as opposed to the "casual talk" representation, the t-score was found to be 2.43. This goes to show that the difference in accuracy results is statistically significant with $p < 0.05$. Similarly, a t-score of 1.20 and 1.27 was found when the results with the negotiation-related representation were compared to the ones selected by Best First feature selection and Information Gain based feature selection respectively.

7.7 Classification of buyer and seller data

To understand if there is a difference in the use of negotiation-related features by different types of negotiators, we separate buyer data and seller data and classify them. For each negotiator we concatenate the messages sent and label them with the role assigned by the system: buyer (positive) or seller (negative). We assign a positive label to a buyer because through available choices a buyer has more power than a seller in modern negotiations held in developed countries (Brett, 2001). The ML tools we use are NB, SVM, and C5.0.[2] Again, we employ tenfold cross-validation and exhaustive search among the adjustable classification parameters.

We run experiments on classifying buyer and seller data for three subsets of the *Inspire* data: all negotiations (B and S), successful negotiations (BSN and SSN), and unsuccessful negotiations (BUN and SUN). For each classification problem, we represent data by bags of negotiation-related words. As a baseline (BL) we consider a classifier that classifies everything as buyer data. Table 7.7 shows the highest average classification accuracy obtained by each classifier. We put in ***bold*** the highest accuracy achieved for each class.

Data	**BL**	**C5.0**	**SVM**	**NB**
B and S	49.98	72.7	***73.1***	68.3
BSN and SSN	49.96	70.8	***72.9***	67.7
BUN and SUN	50.03	***70.2***	68.7	65.2

Table 7.7: Classification accuracy on negotiation-related features, buyers/sellers

The precision, recall and F-measure corresponding to the reported accuracy are presented in Tables

[2]We exclude DLM because of its slowness on buyers and sellers data.

7.8 and 7.9. B and S, BSN and SSN, BUN and SUN mean the same as in Table 7.7. We put in ***bold*** the highest precision, recall and F-measure achieved for each class.

Data	C5.0			SVM			NB		
	P	**R**	**F**	**P**	**R**	**F**	**P**	**R**	**F**
B and S	78.3	58.1	66.5	***79.9***	***61.7***	***69.7***	73.3	57.5	65.6
BSN and SSN	76.5	58.2	65.8	***76.9***	***65.4***	***70.5***	73.2	55.9	62.7
BUN and SUN	63.6	***65.7***	63.5	***76.8***	53.8	63.6	65.2	65.4	***65.3***

Table 7.8: Classification of buyers on the negotiation-related words .

Data	C5.0			SVM			NB		
	P	**R**	**F**	**P**	**R**	**F**	**P**	**R**	**F**
B and S	66.7	83.9	73.5	***68.8***	***84.6***	***75.9***	65.1	79	71.4
BSN and SSN	66.3	***82.1***	73.1	***69.9***	80.4	***74.8***	64.4	79.5	71.1
BUN and SUN	64.5	62.3	63.0	64.4	***83.7***	***72.8***	***65.3***	65.1	65.2

Table 7.9: Classification of sellers on the negotiation-related words.

The classification results show that the ML systems detect differences between messages written by buyers and sellers and represented by negotiation-related words. This supports our hypothesis that the language exhibits the negotiator's role. To obtain more insight into the buyer and seller roles we compare the representativeness of the negotiation-related features with the indicative words. Table 7.10 shows the highest classification accuracy obtained from tenfold cross-validation on the indicative words. The precision, recall and F-score are reported in Table 7.11 for buyers and in Table 7.12 for sellers. BL, B and S, BSN and SSN, BUN and SUN mean the same as in table 7.10. We put in ***bold*** the highest accuracy, precision, recall, and F-measure achieved for each class.

Data	**BL**	**C5.0**	**SVM**	**NB**
B and S	49.98	76.5	***77.5***	72.8
BSN and SSN	49.96	73.7	***77.5***	66.7
BUN and SUN	50.03	72.5	***75.1***	68

Table 7.10: Buyers and sellers, classification accuracy on indicative words.

Data	C5.0			SVM			NB		
	P	**R**	**F**	**P**	**R**	**F**	**P**	**R**	**F**
B and S	76.9	75.8	76.4	78.3	***76***	***77.1***	***82.4***	58.1	68.4
BSN and SSN	72.8	***75.8***	74.2	***78.5***	75.7	***76.2***	78.3	46.1	58.2
BUN and SUN	72.3	***73.0***	72.6	***80.7***	66.1	***72.7***	67.2	70.3	68.7

Table 7.11: Classification of buyers on the indicative words.

Data	C5.0			SVM			NB		
	P	**R**	**F**	**P**	**R**	**F**	**P**	**R**	**F**
B and S	67.9	80.4	73.5	***76.7***	78.9	***77.8***	67.7	***87.6***	76.3
BSN and SSN	69.7	77.4	72.8	***76.6***	***79.3***	***77.9***	68.9	65.6	67.2
BUN and SUN	67.4	65.2	66.0	***71.3***	84.1	***77.2***	61.9	***87.2***	72.4

Table 7.12: Classification of sellers on the indicative words.

The classification results show that the representation by the indicative words allows ML tools to differentiate buyer and seller classes reasonably well, with 22–25 percent of accuracy improvement over baseline. Comparison with the classification results in Section 7.7 shows that representation by the indicative words slightly improves the overall accuracy and reduces the difference in accuracy of classification of buyers and sellers for all three data subsets. In all cases buyers are classified better when represented by the indicative features than by the negotiation-related features. In unsuccessful negotiations, sellers were classified more accurately.

For buyer and seller classification SVM performs better in terms of the overall accuracy than the other two classifiers. However, all conclusions about buyer and seller classification are consistent across three classifiers.

7.8 Multi-class classification of e-negotiation data

In this section we continue multi-class research started in Chapter 6. Recall that we have compared the language patterns employed by four classes: buyers in successful negotiations (BSN), buyers in unsuccessful negotiations (BUN), sellers in successful negotiations (SSN) and sellers in unsuccessful negotiations (SUN). We have claimed that all four classes have distinctive language patterns.

The consequence following from this claim is to proceed with multi-class classification and compare the classification results for four classes. A multi-class classification problem, or choosing one of M classes (1-of-M), differs from a binary classification problem and requires different methods of feature selection and classification. Our current results only sketch some of the problems encountered when the binary classification problem transfers to the multi-class classification problem.

We also show that the knowledge-based feature selection is task- and problem-dependent whereas the statistical feature selection is not. We compare the classification results obtained on the indicative words (corpus-based feature selection) with the results obtained on the negotiation-related features (knowledge-based feature selection). In these experiments the indicative words represent independently scoring features. Examples of selection methods that evaluate features independently are Information Gain, Mutual Information, etc; refer to (Forman, 2004) for the overview of multi-class feature selection.

The multi-class results are better for the indicative words than those for the negotiation-related features. The latter was designed for the reliable separation of successful and unsuccessful negotiations, which is a binary classification problem.

7.8.1 Multi-class classification based on negotiation-related features

Previously we have shown that negotiation-related features give a good accuracy for successful negotiations. However, they do not provide the same good accuracy for unsuccessful negotiations. In this section we investigate how the negotiation-related words affect multi-class classification. We solve the same multi-class problems as in the previous subsection. The results are reported in Tables 7.13 – 7.15.

Table 7.13 shows the results of classification achieved by C5.0.

Data	BL	Accuracy	Precision	Recall	F-measure
BSN vs others	33.7	65.1	48.1	42.5	45.1
SSN vs others	33.7	66.9	51	45.5	48.1
BUN vs others	16.3	79.2	26.5	15.6	19.6
SUN vs others	16.3	78.5	26	17.5	20.9

Table 7.13: Multi-class classification, C5.0; representation by the negotiation-related features.

Table 7.14 shows the classification results of SVM. [3]

Data	BL	Accuracy	Precision	Recall	F-measure
BSN vs others	33.7	70.8	68.2	94.2	78.9
SSN vs others	33.7	70.5	68.6	23.1	34.6
BUN vs others	16.3	N/A	N/A	N/A	N/A
SUN vs others	16.3	N/A	N/A	N/A	N/A

Table 7.14: Multi-class classification, SVM; representation by the negotiation-related features.

Table 7.15 shows the classification results of NB.

Data	BL	Accuracy	Precision	Recall	F-measure
BSN vs others	33.7	66.9	51.7	28.6	36.8
SSN vs others	33.7	68.6	56.6	30.5	39.6
BUN vs others	16.3	42.2	5.5	9.8	1.4
SUN vs others	16.3	15.4	0.6	1.1	0.8

Table 7.15: Multi-class classification, NB; representation by the negotiation-related features.

The results show that in the multi-class classification the negotiation-related features provide better classification of classes related to successful negotiations than classes related to unsuccessful negotiations. The classification accuracies for successful buyers and successful sellers are reliable and consistent within different class splits. However, in multi-class classification the negotiation-related features do not represent unsuccessful buyers and unsuccessful sellers reliably for classification. The reported results confirm the earlier obtained results that decision-based classifiers, e.g., C5.0, benefit from the negotiation-related features if compared with other classifiers.

7.8.2 Multi-class classification based on indicative words

We investigate how the indicative words separate each of the four classes. To do this we represent buyer and seller data by the indicative words.We address the four-class classification problem in the following manner: each class is compared with the other three classes. Three classifiers – C5.0, SVM, and NB – are applied. For each class, we put in ***bold*** the highest accuracy, precision, recall, and

[3] N/A means that everything was classified as a majority class.

F-measure obtained across the classifiers.

Table 7.16 reports the results obtained by C5.0 : accuracy (*Acc*) and corresponding precision (P), recall (R) and F-measure (F). BL denotes accuracy if all examples are classified as the class under investigation.

Data	BL	Accuracy	Precision	Recall	F-measure
BSN vs others	33.7	78.0	67.2	***68.1***	67.6
SSN vs others	33.7	77.0	65.5	***67.3***	66.3
BUN vs others	16.3	88.8	64	66.1	***65***
SUN vs others	16.3	***89.3***	67.1	67	***67***

Table 7.16: Multi-class classification, C5.0; representation by the indicative words.

Table 7.17 reports the results obtained by SVM : accuracy (*Acc*) and corresponding precision (P), recall (R) and F-measure (F). BL denotes accuracy if all examples are classified as the class under investigation.

Data	BL	Accuracy	Precision	Recall	F-measure
BSN vs others	33.7	***82.3***	***83.2***	59.7	***69.5***
SSN vs others	33.7	***81.2***	***81.7***	57.1	***67.2***
BUN vs others	16.3	***90.0***	***85.7***	46.2	60.1
SUN vs others	16.3	89.0	***86.1***	38.9	53.6

Table 7.17: Multi-class classification, SVM; representation by the indicative words.

Table 7.18 reports the results obtained by NB : accuracy (*Acc*) and corresponding precision (P), recall (R) and F-measure (F). BL denotes accuracy if all examples are classified as the class under investigation.

The reported results show that all four classes exhibit different behaviour.

Successful sellers are easily identified within their roles - recall for them is relatively high. However, for successful buyers precision and recall are almost the same. This means that (almost) the same number of successful buyer messages can be mistaken for other class messages as the other class messages can be mistaken for successful buyer messages. We conclude that the indicative words are not suitable for separation of successful buyers from the other classes.

Data	BL	Accuracy	Precision	Recall	F-measure
BSN vs others	33.7	69.6	67.8	54.2	60.3
SSN vs others	33.7	68.3	52.1	75.9	61.8
BUN vs others	16.3	79.3	43.9	***97.4***	60.6
SUN vs others	16.3	79.6	44.2	***97.8***	60.9

Table 7.18: Multi-class classification, NB; representation by the indicative words.

For successful sellers precision and recall differ: recall is noticeably higher. This means that mistaking successful sellers for other classes is relatively harder than the other classes for successful sellers.

Unsuccessful buyers have the lowest precision among all classes. Consequently, their recall is the highest. This means that unsuccessful buyers can hardly be mistaken for other classes but other classes can be easily mistaken for unsuccessful buyers.

The empirical results show that the indicative words favour SVM performance more than C5.0's and NB's performance. The difference is especially noticeable for the reported accuracy and F-measure. For all the three classifiers we have ran t-test on F-measure values obtained on the negotiation-related and the indicative words. The results – 4.3167 for C5.0 and 3.8441 for NB – are significantly different with 0.95 confidence. t-test for SVM gives 1.7960, a significant difference with 0.80 confidence.

In this section we have shown that the features selected because of their good and reliable performance in the binary problem do not provide the same reliable performance for multiple classification. The multi-classification empirical results confirm that SVM performs better on statistically selected features than on knowledge-based selected features. The opposite is true for C5.0: it performs better on knowledge-based selected features than on statistically selected features. Feature selection and the applicability of various learning settings for the multi-class classification requires further investigation that we address in Chapter 8.

7.9 Analysis of the classification results

We have shown that the negotiation-related features are succinct and use relevant knowledge. We have compared the informative power of the negotiation-related features with other features. The results

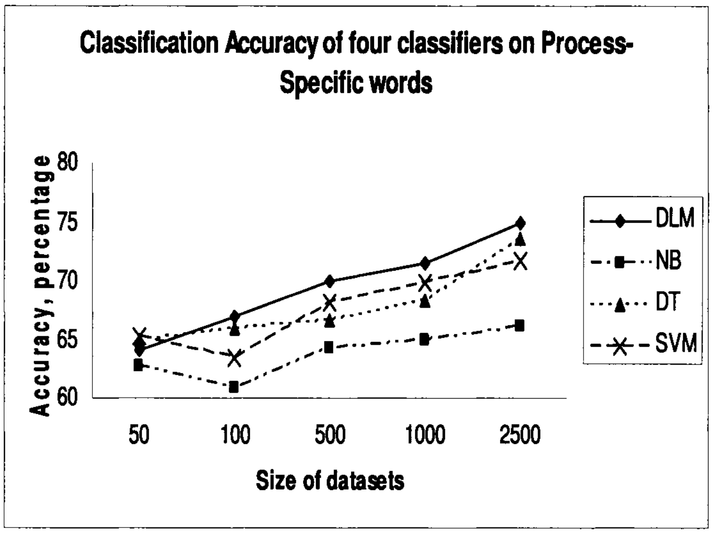

Figure 7.12: Accuracy results of four classifiers

of the statistical analysis support our claim.

The obtained results show that the accuracy of different classifiers varies less when the data are represented through negotiation-related features; see Figure 7.12.

The negotiation-related representation is more stable than other textual representations with respect to classifier performances. The same conclusion applies when we compare the results with the ones obtained on the history records of negotiations. The results show that the correct classification of the successful negotiations benefits from the representation with negotiation-related features. The accuracy of classification of the unsuccessful negotiations is much lower with the exception of Naive Bayes classification. We have looked for reasons that lead to the difference in accuracy results. Analyzing the performance results of Naive Bayes we conclude that the assumption of conditional independence of features, i.e. negotiation-related features, **is not met** in successful negotiations and **is met** in unsuccessful negotiations. We conclude that the negotiation-related features are correlated among themselves in successful negotiations and are not correlated among themselves in unsuccessful negotiations.

We suggest that the difference in accuracy partially comes from inaccurate labelling. The labels given by the *Inspire* system do not always correspond to the real outcome of negotiations; see section 4.4 for examples. Negotiation is labelled as successful if the box "Accept" is checked and as unsuccessful otherwise. Another, and very important, aspect is that semantic categories show **what** people discuss in a negotiation but omit **how** they do it. We will approach the negotiator attitude towards the counter-part and perception of the situation in Chapter 8.

Another interesting observation comes from the classification of buyers and sellers in successful and unsuccessful negotiations. The difference in classification accuracy in successful negotiations is higher than the difference in accuracy of the classification in unsuccessful negotiations with sellers classified more accurately in both cases. This means that in successful negotiations buyers and sellers use the negotiation-related features in more distinct ways than in unsuccessful negotiations.

The multi-class classification results support our conclusions obtained from the studies on language patterns used by different classes; see Section 6.9. The crosswise validation of results makes the obtained results more reliable.

The results of Sections 7.8.2 and 7.8.1 show that the class of unsuccessful negotiations should be addressed in a manner different from the one used for the class of the successful negotiations. We suggest a solution to this problem in Chapter 8.

7.10 Negotiation-related features of higher linguistic order

After empirical results have shown that negotiation-related features are essential for learning the negotiation outcome, we aim to find more about discussion of the negotiation-related issues. To investigate the ways participants use the negotiation-related words we have looked into relations between negotiation-related words and other word types. When applicable, we continue the analysis of the multi-class problem.

We use the opportunities given by the XIP parser; refer to Section 2.2 for a brief description. First, we find which dependencies can have negotiation-related words as their parameters. The candidate dependencies are built for nouns, adjectives and verbs. They are:
– ADJMOD, an intra-chunk dependency between an adjectival head and an adverb or superlative, also works with complex wh-adverbs like "how nice";
– NMOD, a syntactic dependency noun modifier,
– VMOD, a syntactic dependency that links a verb and the head of a nominal chunk which is either a complement or a modifier of this verb.
Each of these dependencies has a few modes:

ADJMOD: POST and PRE modes, where POST mode tags complex dependencies, e.g. "more beautiful than convenient" and PRE mode tags comparative forms and phrases where adjectives are used as nouns, e.g. "East German";

NMOD[4]: POST, PRE, RELATIV, DEDUCED, INFINIT and APPOS modes, where POST mode tags the noun phrase with a related prepositional phrase or the following noun phrase, PRE mode tags when an adjective phrase is before the noun phrase and when the noun phrase contains two nouns, RELATIV mode tags when two noun phrases are not on the same sentence level, i.e. one is in a clause and another is not, DEDUCED mode tags coordinated adjectives with the first noun on their right, INFINIT mode tags an infinitive clause and a noun and APPOS mode is applied between two noun phrases;

VMOD: POST, PRE, NEG, TEMP, COMPLTHAT, SENTENCE, UNSAFE, LOC, MANNER, DURATION and DEDUCED modes, where, for example, POST and PRE modes tag the position of the verb modifier, NEG mode tags the verb negation, TEMP mode tags verbs with temporal features, e.g. day or year, COMPLTHAT mode tags verbs from indirect speech and the main sentence, UNSAFE mode tags a verb in the subclause with the first main verb on its left and MANNER mode tags a verb with a structure "how + adjective".

Next, we apply the parser to the *Inspire* data and extract these dependencies. We calculate the distribution of the negotiation-related words in the dependency parameters; the calculation is done separately for each mode. For this we calculate how many parameters used in a dependency belong to each of the semantic categories. Finally we find modes whose parameters have the negotiation-related words as a major parameter category; see Table 7.19 for the distribution of modes among the data classes. The results show that three modes – ADJMOD-POST, NMOD-PRE, NMOD-RELATIV – are uniformly distributed throughout the classes. By this we consider the ratio of the mode occurrence to the data size in tokens. NMOD-POST and NMOD-DEDUCED, are present more in SS – especially for NMOD-DEDUCED – and SU than in BU and BS data.

We want to see the diversity of dependency pairs across four classes. Table 7.20 reports the number of pair types appearing in each dependency. The type – the number of dependency pairs ratio is size-dependent. It has a negative correlation with the number of dependency pairs. However, for each

[4]XIP makes an error tagging "hope" as noun in the standard business phrase "I hope". This error mostly affects NMOD dependency.

Data	ADJMOD-POST	NMOD-POST	NMOD-PRE	NMOD-RELATIV	NMOD-DEDUCED
BSN	4580	29901	39248	1627	*987*
SSN	4331	29910	38024	1429	***1200***
BUN	1683	10966	15016	661	*350*
SUN	1653	10978	14363	641	*401*

Table 7.19: Distribution of modes among four classes, negotiation-related words.

Data	ADJMOD-POST	NMOD-POST	NMOD-PRE	NMOD-RELATIV	NMOD-DEDUCED
BSN	2379	15544	14611	1096	731
SSN	2257	15444	14541	987	885
BUN	1108	6884	6953	520	290
SUN	1030	6874	6789	509	345

Table 7.20: Distribution of the dependency pair types among four classes, negotiation-related words.

dependency mode the number of types increases with more pairs added. This supports the claim of the unrestrictive growth of the language used in electronic negotiations.

Tables D.1 – D.4 show ten most frequent dependency pairs. Note that we emphasize *hope* which has been mistakenly tagged as a noun. We draw conclusion about the general trends in relations of the negotiation-related words with other words in a sentence. Ten most frequent pairs demonstrate that negotiation-related words are used in the same dependencies across four classes for all the dependencies but one, thus showing that most frequent syntactic relations are the same in most considered cases. The exception – NMOD-RELATIV – puts successful sellers aside from other classes. The NMOD-RELATIV ten most frequent pairs emphasize specifics with which successful sellers address the negotiation issues. To support our claim we report 11th most frequent pair in the corresponding table slot; set in *italics* in Table D.3. This confirms the previously acquired results on the language patterns; see Sections 6.3 and 6.4.

The next question that we consider is whether the dependency results will contribute to feature selection process. By this we mean finding new sets of features missing in previous consideration. In the case of dependencies built on negotiation-related features the answer is negative. The words participating in those dependencies are either function and stop words or negotiation-related words.

Section 7.4.1 reports automatic feature selection methods which find the sets of features combining function and stop words and negotiation-related words. The classification results reported in the section show that little or no accuracy improvement was made when function and stop words were added to negotiation-related words.

We continue to work on the word dependencies in Section 8.12.

7.11 Negotiation-related features in face-to-face-negotiation data

In this section we compare ftf- and e-negotiation data. A representation of e-negotiation data through the negotiation-related features allows reliable classification by Machine Learning algorithms. The classification accuracy was especially high when successful negotiations were classified. This means that negotiation-related features capture the essential and sufficient information from successful negotiations.

We have found the distribution of the negotiation-related features in the *Cartoon* data using the same procedure as for *Inspire* data. The obtained results correlate with those obtained on the on the subsets of *Inspire* data. Refer to Section 5.3 for more explanations. see Table 7.21. Table 7.22 reports the most frequent bigrams used in the negotiation-related context.

Cartoon J	*Cartoon* US	*Cartoon* All	*Inspire*-success
64.8	51.1	56.6	57.9

Table 7.21: Percentage of negotiation-related features in 100 most frequent unigrams (excluding function words)

Although the absence of unsuccessful *Cartoon* negotiations prevents the use of Machine Learning methods to perform classification, the ability to represent these data effectively using negotiation-related features makes a case for this representation scheme. Empirical corpus analysis results show that there are similarities between the *Cartoon* and the *Inspire* data.

7.12 Conclusion

We show that in spite of being short the e-negotiation texts allow reliable classification.

We propose a new procedure for the representation of the e-negotiation text data. We show that

CartoonJ	*Cartoon* US	*Cartoon* All	*InspEmail*	*InspGen1*	*Inspire*-s
5 **million**	you know	you know	i am	i am	i am
want to	i **think**	i **think**	the price	your offer	we are
ultra ranger	going to	per **episode**	i **think**	for you	your **offer**
i see	per **episode**	going to	if you	the **price**	the **price**
down payment	of the	we can	for you	we are	i **hope**
8 **runs**	okay...	of the	in the	i have	for your
the **program**	we can	okay...	we are	thank you	we can
o k	if we	want to	and i	if you	that we
if you	i mean	if we	i **hope**	and i	if you
how about	would be	in the	that we	i **hope**	i have

Table 7.22: Ten most common bigrams, *Cartoon*, *Inspire*-successful.

the negotiation-related knowledge is necessary and sufficient for reliable classification when classes are introduced with respect to the process. The procedure does not depend on the origin of the data (the specific NSS, for example). It only depends on the availability of verifiable negotiation-related knowledge and its language representation. This representation captures the relevant characteristics of such data while leaving out the adverse CMC traits. The empirical results show that such a representation of the e-negotiations provides stable outcomes for different classifiers. The negotiation-related features are suitable for representation of textual ftf-negotiation data.

The negotiation-related features represent the specifics of the negotiation process, so we refer to them as process-specific. We have compared them with a few automatically built features and features generic for the negotiation process, e.g., "casual talk" words. For a binary classification problem the negotiation-related features provide better accuracy than other features. The accuracy increase is statistically significant.

Analysis of classification results shows that there is a difference in the use of negotiation-related features in texts of successful and unsuccessful negotiations. The higher classification accuracy of successful negotiations suggests that the texts of successful negotiations are better identified through the negotiation-related features than the texts of unsuccessful negotiations. The classification of unsuccessful negotiations may also give worse results because the negative class is a disjunctive concept.

We confirm this hypothesis in the next chapter.

The higher classification accuracy obtained for sellers than for buyers suggests that sellers are better identified through the negotiation-related features than buyers, although in unsuccessful negotiations the difference in identification is smaller than in successful negotiations. We report the results of role classification when the data are represented by indicative words. Classification results show that the accuracy of role bearers improves if compared with the representation through the negotiation-related features. The improvement is statistically significant. Alongside, we have shown that SVM benefits, by the standard classification measures, from the statistical feature selection whereas C5.0 benefits from the knowledge-based feature selection.

Further work on the use of negotiation-related features and corresponding results will be presented in Chapter 8.

Chapter 8

Strategy-related Features and E-negotiations

Say no, then negotiate.

Murphy's Law

8.1 Introduction

In this chapter our aim is to find features that make possible the accurate prediction of the unsuccessful outcome. We point out that this is a challenging task because of the multiple reasons the participants have for not reaching an agreement in negotiations. In Chapter 7 we have found that negotiation-related features help predict the successful outcome with high accuracy. However, the features do not guarantee a high accuracy in predicting the unsuccessful negotiation outcome.

In Chapter 3 we wrote that negotiators use the negotiation strategies and the influence strategies. These strategies are more persuasive when delivered in a *powerful* style and less assertive if delivered in a *powerless* style, where "powerful" and "powerless" are standard terms defining the language in negotiations, for example (Burrell and Koper, 1998). We want to emphasize that although sentiment and emotion analysis is a well-developed area of DM, ML and NLP research, we could not find any related research on the strategies in negotiations. The only results that we found come from the planning dialogues in human-computer interactions (Chu-Carroll and Carberry, 2000).

We hypothesize that the negotiation and influence strategies are exhibited in language through the use or the avoidance of mental verbs, modals, adjectives, personal pronouns and negatives. A word of caution: although we cite some specific results of (Brown and Levinson, 1978), we do not support their main assumption that all the participants of interpersonal communications are rational agents, i.e. always choose actions that will satisfy their goals. Such assumption later was relaxed in studies on negotiations; see Bazerman et al. (2000).

Our working hypotheses are as follows:

1. successful and unsuccessful negotiations employ different strategies;buyers and sellers differ in the use of influence strategies;

2. language shows the difference between strategies;

3. how negotiators discuss the negotiation issues provide a reliable ground for learning failure of negotiations.

We use the results of this chapter to support or reject hypotheses listed in Section 9.3 and propose new tasks for the e-negotiation studies in Section 9.5.

8.2 Necessity, possibility, permission and volition in e-negotiations

In this section we search for common trends in the negotiation strategies used by different classes of negotiators. Modals found in the top 100 frequent unigrams in both successful and unsuccessful negotiations are *can*, *will*, and *would*. The modal *could* is found in the 100 most frequent words from unsuccessful negotiations and absent in the top 100 unigrams from successful negotiations. The negotiators who fail to reach an agreement use the modals marginally more frequently than the negotiators who succeed in reaching a compromise. To support this claim, we report the ranks for the modals in Table 8.1. Note that the *have to* rank is reported according to its bigram occurrence. We report statistical significance of the results in Section 8.6.

We compare the distribution of modals in all four classes; see Table 8.2. The ratio of occurrences of *can* and *may* corresponds to that in dialogues, thus supporting previous results about similarities of the e-negotiation data and dialogues; refer to Chapter 4 for more details.

In Tables 8.3, 8.4 and 8.5 we compare the distribution of text covered by patterns with modals. "+" means that the pattern is used more frequently in this class than in the opposite one, "-" means

Modal	$rank_s$	$rank_u$	**Modal**	$rank_s$	$rank_u$
have to	40	37	could	103	97
can	23	24	should	152	154
will	24	22	may	178	175
would	37	34	must	184	177

Table 8.1: Ranks of the most frequent modal verbs.

Modals	*Inspire*				*Cartoon*				*SimpleNS*	
	BSN	SSN	BUN	SUN	BJ	SJ	BUS	SUS	B	S
	544961	525049	209025	205524	8053	8757	56753	63225	51253	47493
can	3647	4015	1536	1596	75	84	284	360	325	288
must	422	370	195	165	5	5	3	3	42	33
will	3952	3929	1544	1582	44	52	58	87	464	356
have to	675	296	639	245	25	20	53	85	62	113
may	349	149	379	169	4	3	41	29	22	18
could	935	362	918	333	12	11	200	189	77	101
would	2841	1095	2548	1025	50	68	392	423	249	268
should	663	255	572	190	9	9	17	21	58	59

Table 8.2: Distribution of modals in buyer and seller data.

that the pattern is used less frequently, and "=" means that the frequencies differ less than 10% in both classes. Recall that the BSN collection is the largest: ||BSN||=1.03*||SSN||=2.6*||BUN||=2.65*||SUN||. Refer to Table C.1 for exact numbers.

Generally, participants in unsuccessful negotiations use primary modals in expressions of obligation and necessity more often than participants in successful negotiations. The average deviations of the specific Mutual Information (Oakes, 1998) of the patterns in Table 8.3 are small, ranging from **0.516** (for **you will**) to **0.403** (for **you must**).

Participants in unsuccessful negotiations use primary modals in permission and possibility settings more often than participants in successful negotiations:

- supposedly permissive **you can** is used more by buyers in unsuccessful negotiations than by

Pattern	BSN	BUN	SSN	SUN
I/we must/will/have to	-	+	-	+
you must/will/have to	-	+	-	+
I/we must	+	-	+	-
you must	+	-	+	-
I/we will	=	=	-	+
you will	-	+	-	+
I/we have to	-	+	=	=
you have to	-	+	+	-

Table 8.3: Distribution of patterns with the primary modals (obligation/necessity), successful/unsuccessful buyers, successful/unsuccessful sellers.

buyers in successful negotiations and is used more by sellers in unsuccessful negotiations than by sellers in successful negotiations;

- obligational **you will** is used more by buyers in unsuccessful negotiations than by buyers in successful negotiations and is used more by sellers in unsuccessful negotiations than by sellers in successful negotiations.

Both buyers and sellers show their authority over counterparts less in successful negotiations than in unsuccessful negotiations. Also, buyers in successful negotiations use democratic imperative *we can/may* less often than buyers in unsuccessful negotiations. The same applies to sellers.

If we compare the numerical values, then sellers in successful negotiations use the patterns **you can/may** and **I/we could/would/should** less than negotiators from other classes. For all four classes of negotiators the pattern denoting self-ability **I/we can/may** is the most frequent pattern used. **You could/would/should** is the least frequent pattern for all buyers and sellers in unsuccessful negotiations. **You can/may** is the least frequent pattern for sellers in successful negotiations.

From these results we conclude that participants in unsuccessful negotiations use secondary modals in exercising permission, possibility and necessity over counterparts more often than participants in successful negotiations. In exercising self-obligation there is a difference between sellers and buyers: buyers in successful negotiations assume that they have power over themselves more often than buyers

Pattern	BSN	BUN	SSN	SUN
I/we can/may	-	+	-	+
you can/may	-	+	-	+
I/we can	-	+	-	+
you can	-	+	-	+
I/we may	-	+	=	=
you may	-	+	-	+

Table 8.4: Distribution of patterns with primary modals (permission/possibility), successful and unsuccessful buyers and successful and unsuccessful sellers.

in unsuccessful negotiations. The opposite is true for sellers: sellers in successful negotiations assume that they have power over themselves less often than sellers in unsuccessful negotiations.

We look for patterns which accommodate the most used mental verbs *think, know, understand* and *consider*. We compare how the mental verbs used in different classes of negotiators (see Table C.2 in Appendix C for details). Buyers and sellers in successful negotiations use *think* (which signals the statement of opinion) and *consider* (which signals further work on agreement) more often than buyers and sellers in unsuccessful negotiations. The opposite is true for the verbs *know* and *understand*, which implicitly signal finding and seeking for excuses respectively. The verb *think* is mostly used in statements *PersPron (ModalVerb) think*: 80% of all occurrences of the verb appear in such statements, with high preference for the pronoun *I*. The verb *know* (almost) equally appears in statements *I/we know*, *you ModalVerb know* and suggestions *to know*. The verb *consider* appears most often in suggestion *please consider* and equally often in statements *I/we (can) consider*, *you (can) consider* and suggestions *to consider*.

We separate statements identifying possibilities and permissions, *I/we (can/may/would/could) think*, from the ones identifying necessities and obligations, *I/we have to/should/must think*. We also distinguish the self-referred statements *I/we (can/may/would/could) know* from the statements applied to the counter-parts *you (can/may/would/could) know*. We notice that for all four data classes the verbs *think* and *know* mostly seed simple statements *PersPron MentalVerb* such as *I/we/you think*. Statements *FirstPron ModalVerb MentalVerb* account for less than 1% of the simple statements. Statements *SeconPron ModalVerb MentalVerb* account for 2-4% of the simple statements. However,

Pattern	BSN	BUN	SSN	SUN
I/we could/would/should	+	-	-	+
you could/would/should	-	+	-	+
I/we could	+	-	+	-
you could	-	+	=	=
I/we would	+	-	-	+
you would	+	-	-	+
I/we should	=	=	+	-
you should	-	+	-	+

Table 8.5: Distribution of patterns with secondary modals, successful/unsuccessful buyers, successful/unsuccessful sellers.

statements *PersPron ModalVerb understand* include the modal verb *can* with high frequency: 15 – 20% of all statements *PersPron understand.* In Table 8.6 for each verb we report patterns in the order of their occurrences. The verbs are listed in order to their occurrences. We report statistical significance in Section 8.6.

Obviously, the directive statements *you have to/should/must MentalVerb* are more frequent in unsuccessful negotiations than in successful negotiations. The implicit suggestions for a feedback, e.g. *I/we think/know/understand* are more frequent in successful negotiations than in unsuccessful negotiations.

Table 8.7 shows how different groups of volition verbs are distributed through different classes; see Table C.3 for more information. The positive volition verbs are used more often in successful negotiations; the negative volition verbs, the verbs that imply additional effort and the verbs that communicate no intention to realize the deal are used more often in unsuccessful negotiations. Note that the total number of the negative volition verbs is much smaller than the total number of the positive volition verbs. We report statistical significance in Section 8.6.

Based on the analysis of the use of modals, mental verbs and volition verbs in four data classes, we state the following working hypotheses:

- in unsuccessful negotiations participants demand more than in successful negotiations, with the direct requests to the counterparts more frequent in unsuccessful negotiations than in successful

Mental Verb	Pattern	BSN	BUN	SSN	SUN
think	*I/we think*	+	-	+	-
	you (can/may/would/could) think	=	=	-	+
	to think	+	-	+	-
	you have to/should/must think	-	+	-	+
know	*to know*	-	+	-	+
	I/we know	+	-	+	-
	you know	-	+	-	+
	you have to/should/must know	-	+	+	-
understand	*I/we (can) understand*	-	+	+	-
	you (can) understand	+	-	+	-
	to understand	+	-	-	+
	you have to/must understand	-	+	-	+
consider	*please consider*	+	-	=	=
	to consider	-	+	-	+
	you (can) consider	+	-	+	-
	I/we (can) consider	+	-	+	-

Table 8.6: Distribution of patterns with the *mental verbs*, successful/unsuccessful buyers, successful/unsuccessful sellers.

negotiations; they express a negative attitude towards discussed issues more often than in successful negotiations; in unsuccessful negotiations the negotiators look for excuses more than and implicitly state their opinion less than in successful negotiations;

- in successful negotiations positive attitude towards the offer and implicit suggestions for continuing negotiation are more frequent than in unsuccessful negotiations;

- sellers in successful negotiations behave differently from other classes: they are less demanding than other classes and avoid the explicit use of authority or explicit mentions of authority.

Volition Verbs	BSN	BUN	SSN	SUN
hope, want, wish, like, prefer	+	-	+	-
agree, promise	+	-	+	-
ask	=	=	+	-
afford, aim, choose, decide, intend, look, plan, propose	=	=	-	+
make, manage, move, proceed, try	-	+	-	+
decline, refuse, reject disagree	-	+	-	+
delay, hesitate	-	+	-	+

Table 8.7: Distribution of *volition verbs* and their wordforms, successful/unsuccessful buyers, successful/unsuccessful sellers.

8.3 Rejection and explicit and implicit denial in e-negotiations

In this section we investigate how different classes of negotiators express rejection and denial; for explanations of terms and theoretical background refer to sections 3.6.1 and 3.6.2. In Table 8.8 we compare texts covered by the negatives in all four classes of negotiators.

Table 8.8 shows that in total, for buyers as well as sellers, negatives are used more in unsuccessful negotiations than in successful negotiations. When we consider different forms of negatives then the exceptions appear in the use of fuzzy and no-negatives by buyers. These negations are more often used by buyers in successful negotiations than by buyers in unsuccessful negotiations. If we compare the use of no-negations and not-negations and the use of affixal negatives within classes, then the results support the claim that the e-negotiation data combines properties of both written and spoken language. In all classes not-negatives are used more often than no-negatives just as it happens in spoken communications; see Table 8.9.

The obtained results show that both buyers and sellers use more explicit negations in unsuccessful

Negatives	BSN	BUN	SSN	SUN
Total	-	+	-	+
Fuzzy	+	-	=	=
No-negations	+	-	=	=
Not-negations	-	+	-	+
Affixal	=	=	-	+
any, ever	-	+	-	+

Table 8.8: Distribution of *negatives*, successful/unsuccessful buyers, successful/unsuccessful sellers.

Negatives	**Inspire**				**Cartoon**				**SimpleNS**	
	BSN	SSN	BUN	SUN	BJ	SJ	BUS	SUS	B	S
Fuzzy	7	6	6	5.5	1.2	1.3	1.3	1.8	4	2
No-negations	24	23	23	22.5	4.8	5.5	6.4	6.2	20.2	25.2
Not-negations	54.6	53.7	57.3	55.2	84.3	90.3	77.9	78.4	60	62.4
Affixal	5	6.4	3.9	6	-	-	3.7	1.8	2	1.4
any, ever	9.4	10.9	9.8	10.8	9.6	2.8	12.7	11.8	13.7	9

Table 8.9: Percentage covered by *negatives* in buyer and seller data.

negotiations than in successful negotiations. As to implicit and fuzzy negations, sellers use them with almost the same frequency in successful and unsuccessful negotiations, but buyers use them more often in successful negotiations than in unsuccessful negotiations.

We look at more details for the use of not-negatives, the most frequent group of negatives. Negotiators rarely use the "face-saving" patterns which are intended for continuing of communication patterns and contain the mental verbs. Such patterns are indicative of powerless speech and make persuasion and argumentation weaker . The tendency to avoid "face-saving" patterns goes along with the low percentage of no-negatives with respect to all negatives, because both express implicit negations. We report results in Table 8.10. Every pattern includes formal and informal spell variations "not" and "-n't" and first-, second- and third-person forms; *cannot* also includes frequently found incorrect spelling *can not* which accounts for 0.185 – 0.199 of all negations of *can*. MV denotes mental verbs. **DO not** does not include patterns with mental verbs.

Not-negatives	*Inspire*				*Cartoon*				*SimpleNS*	
	BSN	SSN	BUN	SUN	BJ	SJ	BUS	SUS	B	S
BE not	26.6	19.8	28.6	42.4	10	3	11.1	9.8	34.1	28.7
cannot	26.8	31	28	28	27	20	6.7	3.6	16.9	17.6
DO not	22	21	21.9	20.2	10	18.5	11.7	16.5	17.9	24.3
HAVE not	6.7	5.8	6.1	6	1.4	3.1	2.9	3.2	2.9	1.5
DO not MV	3.8	3.4	4.1	3.4	2.9	7.6	9.8	5.3	3.1	4.4

Table 8.10: Percentage covered by groups of *not-negatives*, buyers and sellers.

The results show that sellers in unsuccessful negotiations have different trends in the use of BE-negatives than sellers in successful negotiations. The difference between sellers is even bigger than the difference between buyers and sellers within the same class. This prompts us to look at the use of different subgroups of BE-negatives with respect to all Not-negatives; see Table C.4 in Appendix C. Table C.4 shows that sellers in successful negotiations use the simple explicit and specific denial **is not** far less than sellers in unsuccessful negotiations: 7.9% vs. 30.8% with respect to all not-negatives. It also shows that for sellers in unsuccessful negotiations this is the most frequent type of denial. For other classes the most typically used denial is **I/we cannot/can't** with 19.3% for BSN, 19.3% for BUN, and 22.2% for SSN; see Table C.5 in Appendix C. This means that all buyers and sellers in successful negotiations, while reasoning about refusal or denial in negotiations, mostly refer to their inability or not having permission of doing something.

We state a hypothesis that denial varies between different roles in successful and unsuccessful negotiations and has to be studied for all four classes. We should first separate buyers from sellers, and then, buyers in successful negotiations from buyers in unsuccessful negotiations, and sellers in successful negotiations from sellers in unsuccessful negotiations. We support the hypothesis with the following:

- sellers in successful negotiations differ from sellers in unsuccessful negotiations when they deny the buyer's offer: the explicit and specific denial is used more often in unsuccessful negotiations and rarely in successful negotiations;
- buyers in successful negotiations use more fuzzy negations than buyers in unsuccessful negotiations;

- participants in successful negotiations and buyers in unsuccessful negotiations soften their denial by referring to the personal inability or lack of permission to do something.

8.4 Comparison and classification of events in e-negotiations

We look for the distribution of the most frequent adjectives in the data of all four data classes. 20 most frequent adjectives and their relative occurrences (ranks) almost coincide. The common adjectives are: *good, best, new, more, last, high, full, important, first, acceptable, possible, long, better, most* and *lower.* The top adjectives different for the classes are: for BSN - *reasonable, fair, higher* and *next*, for BUN – *little, higher, final* and *reasonable*, for SSN – *low, fair, little* and *next*, for SUN – *low, reasonable, fair* and *little.* These lists show that there is no significant difference among adjectives in different data classes.

In all four classes the intensifiers *very, most* and *best* are commonly used in general settings *most important, very important, most recent, very good* and *best possible.* The classifying adjectives *new, last, first, acceptable* are commonly used with the noun *offer.* The frequent *new line* indicates the introductory phase of negotiations when negotiators discuss the reasons for purchasing bicycle parts. The frequent *long term* indicates that even in simulated e-negotiations participants use appeal to long term co-operation, which is considered to be a positive signal for continuing negotiations. The comparative adjectives *more, better* and *lower* are used in different patterns that involve sophisticated structures.

We discovered that sellers in successful negotiations avoid the use of *reasonable*; buyers and sellers in successful negotiations often discuss *next offer* which is rather specific and implies the continuity of discussion; buyers in unsuccessful negotiations emphasize sending *final offer*, which signals the end of negotiations, and avoid the use of *fair.*

These results support the working hypothesis stated in section 8.2 that implicit suggestions for continuing negotiation are more frequent in successful negotiations than in unsuccessful negotiations. The results add to the hypothesis that explicit signals of the end of negotiations are more frequent in unsuccessful negotiations than in successful negotiations. However, most frequent identifications of events are similar in all four classes

8.5 Implicit questions, reasoning and explanations

The important aspects of negotiation strategies include questioning of the counterpart, explanations and reasoning. In this section we concentrate on identifying language expressions of "cause – effect" relations reflected in phrases like "if – then". In written communication such phrases indicate (simple) questions (Brown and Levinson, 1978; Leech, 1987). Phrases " as/but PersPron" indicate explanations. We parse the data with the aim of finding implicit questions and explanations.

Our system consists of the parser XIP (Chanod and others, 2001) and the in-house scripts to extract the parsing results. We exract XIP's COMPOUND relation that links together single tokens that belong to a complex linguistic expression. We build the unigram models of the data of the extracted phrases. For each extracted phrase we analyze the distribution of the unigrams to identify whether the phrase relates to negotiations or to personal information. In Tables 8.11, 8.12 and 8.13 we report different types of clauses found in each data class. We also compare them with clauses found in the *SimpleNS* data. Recall that the number of tokens in the *SimpleNS* data is approximately 10 times smaller than in the *Inspire* data. However, the results on the *SimpleNS* data mostly support those obtained on the *Inspire* data. Table 8.11 shows that

Data	*If you...*	*As you...*	*But I...*	*As I ...*	*If we ...*	*So I ...*	*As we ...*	*But we ...*
BSN	502	262	112	97	80	132	47	48
SSN	525	273	172	103	66	144	37	56
BUN	232	121	57	43	39	38	32	28
SUN	235	86	54	40	36	45	20	16
B	89	42	35	30	20	41	42	23
S	95	49	59	29	28	43	29	21

Table 8.11: Distribution of most frequent beginnings of conditional phrases.

- in successful and unsuccessful negotiations sellers ask buyers more than buyers ask sellers,
- in successful negotiations sellers explain more often than buyers; the reverse is true in unsuccessful negotiations;
- in successful and unsuccessful negotiations buyers use more questions and explanations referring to mutual actions than sellers; this might mean that buyers condition some actions on preliminary

steps taken by both negotiators.

To investigate in more details how negotiation strategies are used, we work with the most frequent conditional statement "If you". In Table 8.12 the following abbreviations are used:

... accept = If you accept/agree/take

... would = If you would/could/can

... need = If you need/want/wish/like

... offer = If you offer/make

Table 8.12 shows that, as in the case of negations, sellers in successful negotiations use phrases

Data	*If ...*	*If you ...*	*... accept*	*... would*	*... need*	*... have*	*... offer*	*... do*
BSN	941	502	40	64	63	50	2	8
SSN	875	525	36	50	93	66	2	3
BUN	387	232	7	27	22	18	2	4
SUN	391	225	11	25	32	17	1	1
B	242	89	8	25	14	10	2	2
S	263	95	15	20	9	5	1	4

Table 8.12: Distribution of the positive *If you* sentences, buyers/sellers, *Inspire*, *SimpleNS.*

differently from other classes of negotiators: they ask about the wishes of their counterparts more often than other negotiators.

In Table 8.13 the following abbreviations are used:

... don't = If you don't/do not/donnot/dont/doesn't

... cannot = If you cannot/can't/can not

... haven't = If you haven't/have not/havent

Table 8.13 shows that in both successful and unsuccessful negotiations buyers use negative "if – then" phrases more often than sellers, although the frequencies of the phrases are low for every data class.

We showed that questioning differs for buyers and sellers, with sellers asking more questions than buyers. The explanation differs for successful and unsuccessful negotiations, being the prerogative of sellers in successful negotiations and the prerogative of buyers in unsuccessful negotiations.

The results support a previously stated hypothesis about the difference between sellers in successful negotiations and other classes. The difference is vividly seen in the use of implicit questions, reasoning

Data	*If ...*	*If you ...*	*... don't*	*... cannot*	*... haven't*	*... not*	*... won't*	*... reject*
BSN	941	502	41	21	0	3	1	0
SSN	875	525	29	13	2	0	1	0
BUN	387	232	22	19	1	0	0	0
SUN	391	225	19	8	2	0	1	1
B	242	89	1	1	0	0	0	0
S	263	95	0	2	0	0	0	0

Table 8.13: Distribution of the negative *If you* sentences, buyers/sellers, *Inspire*, *SimpleNS*.

and explanations. The sellers use such phrases more often than other classes of negotiators.

8.6 Statistical significance results

We have run two-tailed/non-directional t-test on the relative frequencies of these POS in the data. The null hypothesis is the assumption that the difference between two samples is due to chance. The null hypothesis was rejected with 95 per cent significance level for not-negations of the verb *be* and with 80 per cent significance level for primary modals and collocations *PersPronoun PrimModal.* Table 8.14 presents the t-test results. Statistically significant difference between the data from the "successful" and "unsuccessful" corpora is shown in bold, followed by the confidence level with which the null hypothesis was rejected. *Be not, do/have/can not* correspond to negations of all inflections of the verbs *be, do, have, can* and their spelling versions found in the data, e.g., "can not".

8.7 Classification of the negotiation outcomes and discussion

We want to support experimentally our claim that the language implementation of strategies differs in successful and unsuccessful negotiations. We represent data through the words corresponding to the strategies (strategy-related features) and run classification experiments.

We have a total of 2557 examples in our data set, of which 1427 are positive (successful negotiation) and 1130 negative (unsuccessful negotiation).Recall that the class labels are noisy, partially due to the *Inspire* system's flaws. Analysis of the data has shown that 3 – 5 % of the negotiations that the

Sample	**Degrees of freedom**	t value	**Significance level**
Primary Modals	10	**1.435**	0.8
Secondary Modals	6	0.180	insignificant
You PrimModal	10	**1.7328**	0.8
I/we PrimModal	10	**1.738**	0.8
Positive Volition Verbs	40	0.051	insignificant
Negative Volition Verbs	12	0.452	insignificant
Mental Verbs	10	0.021	insignificant
Negations	8	0.150	insignificant
Be Not	7	**2.524**	0.95
PersPronoun do/have/can not	12	0.602	insignificant
the latest & is the best	7	0.716	insignificant

Table 8.14: Statistical difference between samples.

system records as unsuccessful ended with the participants agreeing verbally to accept an offer.

The data are represented by bags of "strategic" words, and each bag corresponds to one negotiation. In each bag, attributes have numerical values equal to the number of occurrences of a strategy-related features in negotiation, and an additional attribute whose value is equal to the number of other words in the negotiation.

We compare the performance of kernel, decision-based and probabilistic classifiers on negotiation data. The first class is represented by Support Vector Machines (SVM), the second by decision trees (C5.0) and the third by probabilistic Naive Bayes (NB). We use tenfold cross-validation to estimate the accuracy. For more details and explanations on the classifiers' performance refer to Section 2.3.2.

In order to verify our claim that the strategy-related features are necessary and sufficient, we compare the results of two sets of experiments. In the first set we represent the data using personal pronouns, modal verbs, the verbs *do* and *be*, and their negative versions. That is, we form bags of words for each negotiation using the number of occurrences of the words in the negotiation. In the second set of experiments the data are represented by top 500 unigrams including function words; see subsection 7.4.1.

To justify adding the number of other words in negotiation when bags of words are built with word frequencies, we evaluate the attributes using the "Select attributes" option in Weka. For each of the reported data representations we evaluate the attributes with the Best First, Forward Selection, and Genetic Search methods. The additional attribute was selected in all cases.

We list the classification results in Tables 8.15 and 8.16. On the representations with strategy-related features and personal pronouns + modals + negations, SVM performed much more slowly than NB and C5.0.[1] For the top 500 unigrams representation N/A means that SVM was running at least 4 times more slowly than it did with our representation of data. We report the results of NB with kernel density estimation because it classified negotiations more accurately than NB with normal distribution.

We report the average tenfold cross-validation results for all the experiments. These were the results for the set of parameters for each classifier that gave the highest classification accuracy. The baseline *Acc* equals 55.8%, when all negotiations are classified as positives. Corresponding P, R, and F are equal to 55.8, 100 and 71.6 per cent respectively. # denotes the number of features in the representations, or the number of attributes in bags of words.

Features	#	NB	SVM	C5.0
strategy-related features	100	65.25	71.26	74.5
personal pronouns + modals + negations	26	63.57	67.4	71.4
top 500 unigrams	501	63.4	N/A	75.5

Table 8.15: Classification accuracy of the negotiation outcomes, strategy-related features.

Features	#	NB			SVM			C5.0		
		P	R	F	P	R	F	P	R	F
strategy-related features	100	58.3	74	65.2	73.2	74.8	74	72.5	87.6	79.25
pers pronouns + modals + negations	26	54.9	73.2	62.7	66.5	72.7	69.5	69.2	87.91	77.45
top 500 unigrams	501	46.4	71.2	55.83	N/A	N/A	N/A	73.39	88	79.8

Table 8.16: Classification of the negotiation outcomes, strategy-related features.

[1] This slowness is due to the specifics of Weka's implementation. Generally, SVM is a fast learner.

High values of P, R and expectedly high value of F mean that the overall good performance of C5.0 is due to the high accuracy of classification of the positive examples which dominate in the data. C5.0 performs poorly on the negative examples which makes its application to the "real world" negotiation data questionable. With P and R values closer than those in the C5.0 experiments, SVM has shown more balanced performance. It classifies the positive examples slightly worse than C5.0, and the negative examples considerably better.

For NB the low precision and moderate recall lead to expectedly low F-measure. With the number of negative examples less than the number of positive examples, this shows that NB classifies negative examples better than positive examples. This characteristics of NB might be very important when we gain access to the data of real world negotiations. The fact that the kernel density estimators perform better than normal distribution shows that the normality assumption does not hold for e-negotiation data. The latter is consistent with the classical conclusions Mandelbrot (1954) on natural language texts.

Now we consider three sets of features discussed earlier. Performance is the weakest when negotiations are represented only through personal pronouns, modals and negations. Although there is a statistically significant difference in their use in successful and unsuccessful negotiations, they do not provide enough information to separate the two classes.

The addition of mental and volition verbs to personal pronouns, modals and negations reduces the difference between the precision and recall values and improves the classification accuracy and F-measure for all three classifiers. However, the automatically selected features – 500 most frequent unigrams – worsen recall and F-measure for NB as compared to the strategy-related features. This means that positive examples are classified more accurately and negative examples less accurately, when represented by the most frequently used words. We attribute this to the fact that words not related to negotiation strategies, such as greetings, closure and casual words are similarly correlated in both successful and unsuccessful negotiations.

We conclude that a reliable classification of negotiations is possible if the features include elements of logical reasoning and appeal (personal pronouns, modals, negations), the attitude toward the issues (volition verbs in our case) and the intention on continuity of negotiation (mental verbs in our case). Representation only via logical reasoning is insufficient to produce reliable classification results.

8.7.1 The strategy-related features and the role classification

We have shown that the negotiation outcomes are reliably classified when the whole negotiation constitutes one data entry. Chapter 3 states that participants affect each other's negotiations in the negotiating process. The influence and negotiations strategies used depend on those of the negotiating counter-parts and are insufficient for understanding if considered isolated. This conclusion implies that the strategy-related features extracted from the negotiator's transcripts are related to ones extracted from the counter-part's transcripts.

We have used the strategy-related features to represent data in the role classification. The results are reported in Tables 8.17 and 8.18. As a baseline (BL) we consider a classifier that classifies everything as buyer data.

Data	BL	NB	SVM	C5.0
B and S	49.98	53.7	57.0	52.3
BSN and SSN	49.96	55.3	57.2	51.6
BUN and SUN	50.03	50.2	52.8	52.3

Table 8.17: Classification accuracy of the buyer and seller roles; the strategy-related features.

Data	NB			SVM			C5.0		
	P	*R*	*F*	*P*	*R*	*F*	*P*	*R*	*F*
B and S	53.5	57.1	55.2	56.4	61.4	58.8	52.7	51.2	51.9
BSN and SSN	54.9	59.3	57	57.4	56.6	57	51.6	52.2	51.9
BUN and SUN	50.3	26.3	34.6	52.8	64.5	58.1	52.3	52.3	52.3

Table 8.18: Classification of the buyer and seller roles; the strategy-related features.

The empirical results support the claim that taken separately for a buyer and a seller within a negotiation the strategy-related features do not provide solid grounds for a reliable classification. Thus we omit the strategy-related features from the multi-class problems, where classes are defined with respect to the roles.

8.8 Comparison of negotiation-related and strategy-related features

So far we have introduced different sets of features representing e-negotiation data. An obvious continuation of the work is to compare the informativeness of these features with respect to the negotiation outcome. We also want to find out which feature representation intensifies the learning ability of machine learning methods. In other words, we seek answers to the following questions:

1. which features are better for classification of successful and unsuccessful negotiations?

2. from which feature representation will classification benefit more?

8.8.1 Empirical setting and results

To answer Question 1 we first separate feature selection methods according to the automatic involvement in the selection process. Such methods belong to one of two generic categories:
a) partially-automatic or non-automatic, and
b) fully automatic selection.

The former category usually involves employing, to different extent, knowledge about the domain of the data origin. The latter category involves scoring of features according to various criteria. Within each of these two categories we concentrate on selection successfully used on e-negotiation data. Within the first, the so-called knowledge-based category, we work with the selection methods that use the knowledge about the process of negotiation. Hence the selected features are called *process-based.* Within the second, the so-called automatic category, we work with the selection methods that score features according to the corpus statistics. The features so are called *corpus-based.* For the two above-mentioned subcategories we explore the methods approaching the selection task from different directions. In the process-based subcategory we explore features specific to the process via *negotiation-related* features and features generic to the process via *strategy related* features. In the corpus-based subcategory we explore features common to different sub corpora via *most frequent* features and features on which the sub-corpora differ via *indicative* features. Figure 8.1 shows a hierarchical division of features.

To answer Question 2, we build two representations for each set of features:

1. with the numerical attributes whose values are the numbers of occurrences of the word in negotiation; in this case we add one more attribute, whose value is the number of occurrences of

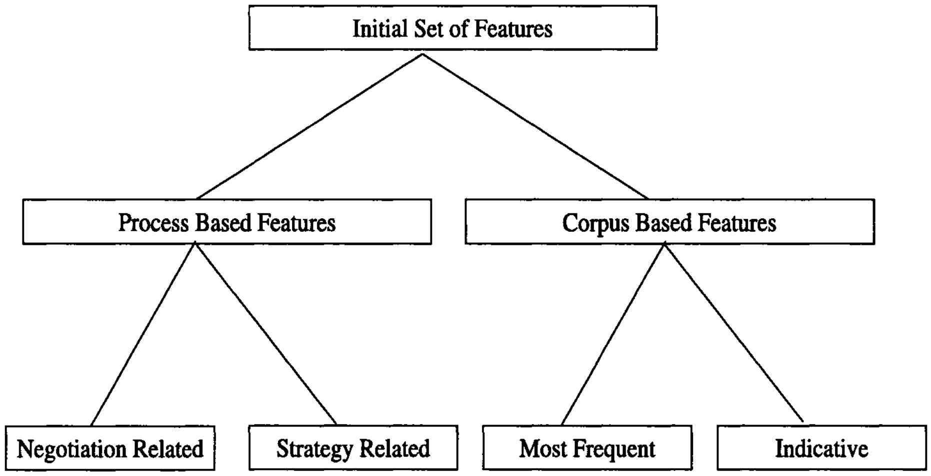

Figure 8.1: Sets of features

other unigrams in the negotiation;

2. with the binary attributes showing whether the feature appears in the negotiation; there is no additional attribute.

In Table 8.19 we summarize the proposed feature sets and their representations.

Features	Representations	Selection basis	Automatic
negotiation-related	binary	negotiation	semi
negotiation-related	numerical	negotiation	semi
strategy-related	binary	negotiation	semi
strategy-related	numerical	negotiation	semi
most frequent	binary	corpus	yes
most frequent	numerical	corpus	yes
indicative	binary	corpus	yes

Table 8.19: Features and representations.

In Tables 8.20, 8.22, 8.21 and 8.23 we have reported the highest accuracy and corresponding precision, recall and F-measure for each classifier on every feature set and feature representation.

The tables show that negotiation-based features give better accuracy results than corpus-based

Features	# of attr	NB	SVM	C5.0
negotiation-related	124	69.3	71.7	75.4
strategy-related	100	65.3	71.3	74.5
most frequent	201	64.3	73.4	71.5
indicative	201	64.2	72	74.4

Table 8.20: Classification accuracy of the negotiation outcomes; numerical representation.

Features	# of attr	NB	SVM	C5.0
negotiation-related	123	69.4	74.0	74.8
strategy-related	99	71.1	72.7	73.7
most frequent	200	64.2	71.5	73.3

Table 8.21: Classification accuracy of the negotiation outcomes; binary representation.

features. Specifically, the correct identification of successful negotiations as well as correct identification of unsuccessful negotiations decrease when automatically built sets of corpus-based features are used.

8.8.2 Difference between the feature sets

Two negotiation-based selected sets of features give higher classification accuracy than the corpus-based selected sets of features. However, there exists an important difference between the sets of negotiation-based features, the negotiation-related and the strategy-related.

We have shown that the negotiation-related features provide high accuracy in classification of positive examples, especially with the numerical representation. The strategy-related features improve the classification accuracy of negative examples. In the context of negotiations this means that knowing how often people discuss the topic of negotiations helps identify successful negotiations. Knowing how the participants implement their influence strategies helps identify unsuccessful negotiations. We explain such results by different trends in the use of the negotiation-related and strategy-related features in successful and unsuccessful negotiations. We present our conclusions in detail below.

For the negotiation-related features we have demonstrated that successful negotiations are classified much better than unsuccessful negotiations. This means that the classifiers extract sufficient information from positive examples but cannot extract sufficient information from negative examples.

Features	# of attr	NB			SVM			C5.0		
		P	R	F	P	R	F	P	R	F
negotiation-related	124	72.3	72.5	72.5	75.8	72.5	74	73.3	87.7	79.9
strategy-related	100	58.3	74	56.7	74.8	73.2	74.0	72.5	87.6	79.3
most frequent	201	54.4	74.6	62.6	72.9	75.3	74.1	84.2	72.4	80.0
indicative	201	54.6	74.6	62.9	73.2	75.8	74.5	85.9	73.0	79

Table 8.22: Precision and recall of the negotiation outcomes; numerical representations.

Features	# of attr	NB			SVM			C5.0		
		P	R	F	P	R	F	P	R	F
negotiation-related	123	66.4	72.3	69.5	73.1	84.6	78.4	72.6	88	77.3
strategy-related	99	71.5	80	75.6	71.4	85.3	77.7	71.3	87.4	78.9
most frequent	200	54.6	74.6	63.1	72.9	75.3	74.2	72.3	84.2	77.8

Table 8.23: Precision and recall of the negotiation outcomes; binary representations.

This corresponds to the situation when with respect to the selected features

1. the positive class either is homogenous or consists of a few well-represented subclasses;
2. the negative class is divided into several small subclasses, and some subclasses are under-represented.

Similarities among successful negotiations are easily revealed through the use of negotiation-related features and are strong enough to build a homogenous class whereas for unsuccessful negotiations this assumption does not hold.

The strategy-related features extract stronger similarities in their use in the negative class, thus improving the correct identification of the negative examples. On the other hand, with respect to the positive class the strategic features do not extract the same similarities as the negotiation-related features, thus giving less information to the classifiers.

It is natural to ask whether we can use the benefits of both features simultaneously. One possible way is to construct new features that benefit from the features from both sets. However, this might be a rather complicated task where automatic feature construction methods should be augmented by domain knowledge. We argue that the negotiation-related and strategy-related features differ with respect to the negotiation process:

- the negotiation-related set captures the negotiators' main goal towards the negotiating issues, the negotiation preferences and the scope (the width, depth and generality and specificality) of negotiations; with the numerical representation features revealing the intensity of the discussion of negotiation issues;

- the strategy-related set does not relate specifically to negotiation issues; it captures the intentions to continue a negotiation, the influence on a counterpart, self-obligations and motivations, openness to feedback or the opposite, the boundaries within personal communication, etc; the numerical representation of features reinforces the captured trends.

Thus negotiation-related and strategy-related features represent different sides of the same process.

Another opportunity to benefit from both sets of features comes from building an ensemble of classifiers where the classifiers built by the same learner use different sets of features to classify the data and then combine their results. SVMs with the high accuracy and the most balanced performance on the data are reasonable candidates. This approach seems to be the most promising one. We suggest it as one of the avenues for future work.

8.9 Analysis of classification performance on different sets of features

Here we discuss the resulting performance of the classifiers on different feature sets and their representations. From the practical point of view we want to show the successful negotiation trends and warn of unsuccessful negotiation in the best possible way. Thus, we need to know which combination of a feature set, the feature representation and a classifier gives a higher rate of the correct identification of positive examples and which combination is better in the correct identification of negative examples.

Precision and recall values show that successful negotiations are better identified by C5.0 than by other classifiers, with NB being the poorest classifier for successful negotiations. The best feature set for correct identification of successful negotiations is *negotiation-related features*. The opposite is true for unsuccessful negotiations: they are more correctly identified by NB than by other classifiers, with C5.0 being the poorest classifier. The best feature set for correct identification of unsuccessful classifiers is *strategic features*; see Table 8.22.

For C5.0, using numerical representation ensures higher accuracy than simple presence of the same words. This is explained by the more trustworthy calculation of the information gain value. Again as expected, SVM performed much better on the binary representations than on numerical representations. This can be explained by the original design of SVM.

NB has better accuracy on numerical representation than on binary representation for all features except the strategic features. We attribute this to the stronger conditional independence between the presence of strategic features than between their occurrences in negotiation. [2]

We have noted different trends in classifier performances for the correct identification of successful and unsuccessful negotiations. With C5.0, both classes were better identified via occurrences than the presence of words. With SVM, however, the identification results for each class varied. The increase of the overall accuracy on the presence-based representations was due to a much better identification of unsuccessful negotiations with the slight decrease in the correct identification of successful negotiations. The opposite is true for NB. Its much higher accuracy on the presence of strategy-related features is explained by a 40% increase in correct classification of successful negotiations, with the accuracy decline on unsuccessful negotiations.

From all classification options, the *most balanced* results, with the closest correct identification of successful and unsuccessful negotiations, were produced by SVM using the numerical representation of the strategic features. The classification results on the numerical (occurrence) and binary (presence) representations of automatically selected features are close to each other for each classifier. It is worth noting that for each classifier the tendency of the results was opposite to that shown on negotiation-based representations: NB and SVM perform better on the numerical representation and C5.0 performs better on the binary representation.

The summary of the performance is the following: C5.0 achieves its high accuracy due to the high correct identification of true positives. Its performance on negative examples is rather poor. SVM is more balanced in its identification of true positive and true negative examples, with better performance on positive than on negative examples. NB performs better among three classifiers in the correct identification of negative examples.

The results of this section are used in proposals for future work; see Section 9.5.4.

[2]This means that the word types in which negotiators express their strategies are more correlated in numbers than in presence/absence.

8.10 Text segments and the negotiation outcomes

The research problem that we present in this section is the influence of a deadline on the course of negotiations and the way it affects the classification of e-negotiation texts. We hypothesize that the presence of a deadline puts pressure on the participants, who become more motivated to conclude the negotiation process, and that this can be detected from the language used. From point of view of the language use, our hypothesis states that certain language features make the second part of negotiations (and its textual data) more indicative of the outcome of the negotiation than the first part.

This contradicts previous analysis of face-to-face negotiation data by Simons (1993) who proposed that the language patterns used in the first half of a negotiation are better predictors of the negotiation outcome than those in the second half. The explanation was that in the first phase people establish contact, exchange personal information and engage in general polite conversation, creating a foundation of trust between partners. No numerical data, however, supported this claim. From the text classification point of view, our hypothesis says that the classification of the second part of e-negotiation texts is more accurate with respect to the outcome than the classification of the first part. This makes e-negotiation texts different from those classified by Blaták et al. (2004), where texts showed better classification accuracy on their initial parts. We report the results of several sets of ML experiments. Performed on varying-size text data segments and different feature sets, they support our hypothesis.

8.10.1 Time slots in negotiations

We propose that throughout a negotiation not all messages exchanged are created equal, and that the presence of a deadline will cause a more informative exchange closer to the conclusion of the process. To test this hypothesis, we take an ever smaller tail end of the negotiation, and see how well we can predict the outcome of the process, based only on the messages in this fragment. To support our claim that the language captured in the e-negotiation data changes as a negotiation progresses, and that such changes can be analyzed using ML, we conduct a series of classification experiments. We divide the text for each negotiation in half and for each half we build a bag of 123 negotiation-related features. The attributes are binary: the presence or absence of the word in its half of the text.

We concatenate the two bags, and label the resulting bag by the outcome of the whole negotiation: positive if the negotiation was successful, negative otherwise. We repeat this procedure for the split of the negotiation text into 3/4 and 1/4. The classification results, accuracy, precision P, and recall R (Manning and Schutze, 1999), are reported in Tables 8.24 and 8.25.

Features	Split	# of attr	NB	SVM	C5.0
negotiation-related	1/2 and 1/2	246	68.1	73.6	73.9
negotiation-related	3/4 and 1/4	246	69.1	73.7	75.4

Table 8.24: Classification accuracy of the negotiation outcomes; representation with positions.

Features	Split	# of attr	NB		SVM		C5.0	
			P	R	P	R	P	R
negotiation-related	1/2 and 1/2	246	73	68	75.4	78.2	72.1	86.8
negotiation-related	3/4 and 1/4	246	74.1	68.7	75.5	78.5	73.83	86

Table 8.25: Precision and recall of the negotiation outcomes; representation with positions.

The results show that C5.0 is sensitive to the positions of words in different parts of the negotiations. SVM's and NB's accuracy change only slightly. The precision and recall results give a better picture of the performance. The presence/absence of words recorded for different splits of negotiations influences the identification of true positive (successful negotiations) and true negative (unsuccessful negotiations) examples. Successful negotiations have the highest rate of true identification, achieved by C5.0, when the negotiations are split in half. Unsuccessful negotiations have the highest rate of true identification, achieved by NB, when the split is 3/4 and 1/4; this split lets us improve the worst rates of true classifications – unsuccessful negotiations for C5.0 and successful negotiations for NB. Generally, the unequal split allows us to reduce the difference between true positive and true negative results, and thus makes the classification of negotiations more balanced than the equal split.

8.10.2 Experimental setting

We use ML tools to investigate the influence of deadlines on negotiation outcomes. We test whether the final segment of a negotiation is better at predicting the success of the process than the initial segment. If we discover that this is true, we will also experiment with different segment sizes, to find out what segment size suffices to predict the outcome.

We need to determine the placement of the segment of a negotiation most salient to deciding whether the outcome is positive: at the beginning or at the end of the process. To do that, we split each negotiation in half, and build two parallel data sets, corresponding to the two halves. We classify each part using various ML tools. Next, we repeat the same classification tasks using smaller and smaller final segments, in order to monitor the variation in performance. Thus each negotiation text N consists of the head segment(h) and the tail segment(t):

$$N = h \cup t, h \cap t = \emptyset \tag{8.1}$$

where their sizes in tokens are $|t| = \frac{N}{i}$ and the corresponding $|h| = \frac{(i-1)N}{i}$. We stop when for two consecutive splits two classifiers have better accuracy on the head than on the tail. Each segment gets the same class label as the whole negotiation. The best results, in particular with high overall accuracy, are reported in Tables 8.26 and 8.27.

Classifier	full length	head 1/2	tail 1/2	head 2/3	tail 1/3	head 3/4	tail 1/4	head 4/5	tail 1/5	head 5/6	tail 1/6
C5.0	74.8	71.1	74.4	72.4	75.2	72.2	74.9	72.9	75	72.7	73.9
SVM	74	73.5	75.3	71.8	74.1	71.4	73.5	72.1	73.6	72.6	74.6
NB	67.5	67.6	68.8	67.6	70	67.7	70.1	68	70.8	68.1	70.9

Table 8.26: Classification accuracy on splits; the negotiation-related words, binary attributes.

Classifier	full length	head 1/2	tail 1/2	head 2/3	tail 1/3	head 3/4	tail 1/4	head 4/5	tail 1/5	head 5/6	tail 1/6
C5.0	73.2	72.3	73.8	72.9	73.7	72.6	73.4	73	72.6	73.1	72.9
SVM	72.7	72.55	73.8	71.3	73.6	72.4	72.8	72.8	72.4	72.7	73.4
NB	71.1	70.3	70.6	71	70	71.2	69.5	71.1	69.1	70.8	68.7

Table 8.27: Classification accuracy on splits; the strategic words, binary attributes.

The tail segments give more accurate outcome classification than the head segments when the data is represented using negotiation-related words. This holds for all splits and all classifiers. The increase in accuracy when the head segments are growing was to be expected, although it does not happen with C5.0 and SVM – only with NB. At the same time, there is no monotonic decline in accuracy when the length of the tail segments decreases. On the contrary, NB constantly improves the accuracy of the classification.

The results on strategy-related features are slightly different for the three classifiers. SVM classifies all tail segments better than head segments, C5.0 classifies tail segments better than head segments up to the 4/5 and 1/5 split, and NB classifies the tail segment better than the head segment only for the half-and-half split. The accuracy results are unstable for all three classifiers, with a sudden accuracy decrease when the head segments grow and increases when the tail segments shrink. The classifier performance means that the features reflect the process of negotiation less adequately than negotiation-related words.

To investigate which part of the tail segments is more important for classifying the outcomes, we introduced additional splits in the tail segments. We divided the second half of texts into 2 and 3 parts and repeated the classification procedures for every new split. The results appear in Table 8.28, where "tail" shows the classification results when the second half of the text was classified, and the next columns report the results on the tail splits; both splits satisfy the conditions $tail = \cup s_i, s_i \cap s_j = \emptyset$.

Classifier	tail	s_1	s_2	s_1	s_2	s_3
C5.0	74.4	71.9	74.9	72.5	71.9	73.9
SVM	75.3	70.52	73.5	70.79	69.9	74.6
NB	68.8	68.53	70.1	68.72	68.88	70.87

Table 8.28: Classification accuracy on tail splits; the negotiation-related words, binary attributes.

The results show that adding splits in the tail segments emphasizes the importance of the last part of a negotiation: in both experimental settings, the classification of the outcome on the last part of the tail is more accurate than on the other parts of the tail. This holds for all three classifiers. For the strategy-related features (not included for lack of space) the same is true for C5.0 and SVM, but not for NB. NB classifies the negotiation outcomes more accurately on splits s_1 than on splits s_2 and s_3.

To better analyze the classification of true successful and unsuccessful negotiations, we present the precision and recall values in Table 8.29. The results on the negotiation-related representation show that true negative negotiations are best classified using the end of a negotiation – on segment

s_2 for the first split and on segment s_3 for the second split. NB is the most accurate in finding true negative examples (unsuccessful negotiations). The same holds when negotiations are represented through strategic words. There is no obvious tendency in identifying true positive examples (successful negotiations) when we compare the accuracy on s_1 and s_2. All classifiers differ when classifying true positives: NB improves its accuracy, C5.0 achieves exactly the same accuracy, SVM decreases its accuracy.

Classifier	s_1		s_2		s_1		s_2		s_3	
	P	R	P	R	P	R	P	R	P	R
C5.0	70.92	84.35	74.2	84.3	72.07	82.95	70.38	83.51	68.61	83.33
SVM	71.2	79.32	76.3	76.3	73.4	74.8	72.3	74.8	75.3	78.3
NB	72.5	70.4	73.9	71.8	73.2	69.4	73.2	69.9	74.9	72

Table 8.29: Precision and recall on tail splits; the negotiation-related words, binary attributes.

8.10.3 Segmentation Results

Because of the results reported in section 8.10.2, we choose negotiation-related words as the feature set. We select for further analysis the half that performed better for a majority of the tools used. We focus on the last part of the negotiation, and we extract a gradually smaller fragment (1/2 - 1/9, where 9 is the average number of text messages in one negotiation). Figure 8.2 plots the results of the experiments performed with decreasing segment sizes. As we see, the tail segment of the length 1/7 gives a decline of the accuracy for SVM and NB, with a slight improvement on smaller tail segments.

A more detailed analysis comes from looking at precision and recall results on the segments; see Table 8.30. We use a generalization and specialization approach in our analysis (Mitchell, 1997; Japkowicz, 2001). On 1/7 and 1/9 tail segments all classifiers have improved the identification of true negatives. This means that the trends in the class of unsuccessful negotiations become more noticeable for the classifiers when the deadline is approaching. The 1/8 split is an exception, with the abrupt drop of true negative classification by C5.0. The correct classification of positive examples, however, diminishes when splits get smaller; this applies to the performance of all three classifiers. This means that at the end of the negotiations the class of successful negotiations becomes more diverse and the trends are more difficult to capture by the classifiers.

As in the previous experiments, NB's accuracy on the tail segments is higher than on the complete

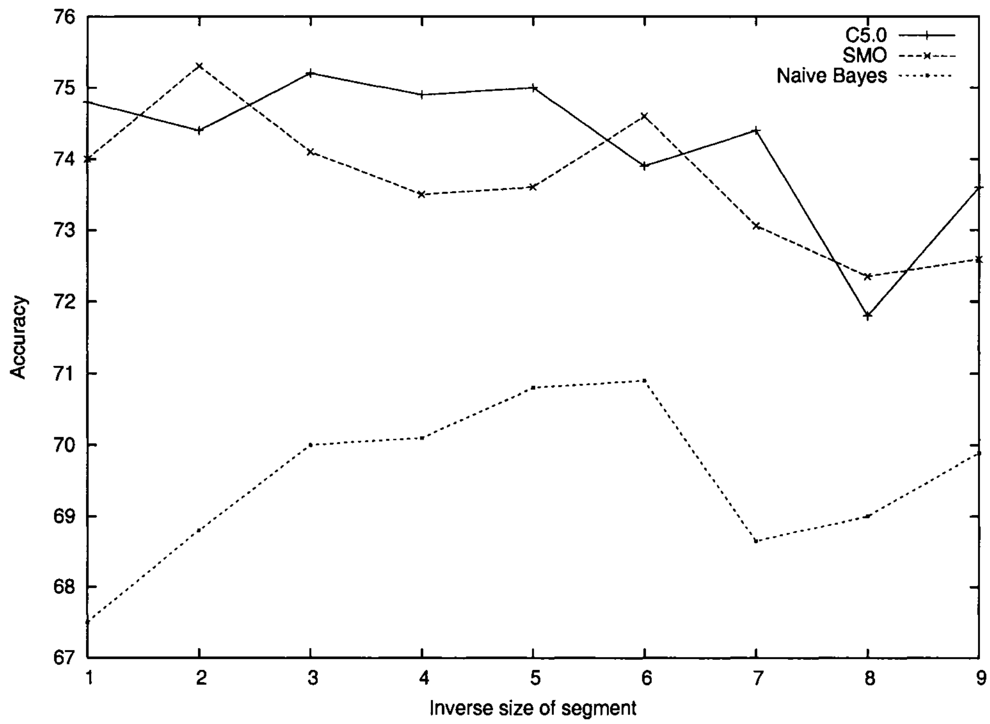

Figure 8.2: The evolution of classification accuracy with decreasing segment sizes.

data. The opposite is true for SVM and C5.0 – their accuracy on the tail segments is lower than on the complete data. We explain this by the fact that the sizes of tail segments in the last splits do not give these two classifiers sufficient information.

The results of this section are useful for information extraction because they pinpoint the text segments most densely conveying information about the negotiation outcomes. Another application of the section results lies in the sequential learning of e-negotiation data. Section 8.11 continues the study of e-negotiation segments but from a different perspective.

8.11 Preventing unsuccessful negotiations

In this section we investigate the ability of our approach to assist the negotiation process. We consider the task of possibly preventing unsuccessful outcomes early. The task is directly related to the segmentation of texts. Our goal is to find if an early reliable classification of unsuccessful negotiations is available.

For this we conducted several sets of experiments on the text segments. The segments are

Classifier	1/3		1/4		1/5		1/6		1/7		1/8		1/9	
	P	R	P	R	P	R	P	R	P	R	P	R	P	R
C5.0	74.2	85.3	74.2	84.3	75.2	82.3	73.61	83.02	74.53	82.39	72.1	81.62	74	81.3
SVM	76.1	78.1	76.3	76.3	77	75.3	78.3	75.3	77.2	73.4	76.9	72.3	77.6	71.6
NB	73.8	71.8	71.8	73.9	74.8	71.9	74.9	72	71.3	72.2	70.8	72.5	70.5	74.3

Table 8.30: Precision and recall on the tail segments; negotiation-related words.

1/6,1/5,1/4,1/3,1/2,2/3,3/4,4/5 and *5/6*

where *0.XX* is the negotiation length calculated from the start. Each set of experiments involves NB, SVM, DT and KNN. We apply the classifiers to negotiation- and strategy-related features, with two different representations each - by binary and numerical attributes. **True negative** and **true positive** rates are the results of the experiments.

Because of the large volume of the numerical results we have performed their generalization. We calculated polynomial regression from the result curves (Duda et al., 2000):

$$\hat{\mathrm{Y}} = \hat{\beta}_0 + \sum_{j=1}^{m} \mathrm{X}^j \hat{\beta}_j \qquad (8.2)$$

where X is the vector of curve data and $\hat{\mathrm{Y}}$ is the approximation of the output. For more details on formulae refer to Section 2.5 and Equations 2.22 – 2.24.

For each classifier, a regression was calculated across the segments on the obtained true accuracy rates. Such generalization preserves all necessary information about the tendency of the outcomes and helps to understand the trends of the negotiation process.

We divide the section into two parts: the first part presents batch learning, when a classifier has access to the complete data; the second part presents incremental learning when a classifier re-evaluates the results as soon as the next data entry comes in (Mitchell, 1997).

8.11.1 Batch learning results

Figure 8.3 presents the regressions of a true negative rate curves; see Equation 8.2. The resulting polynomials have the following orders: first for NB, forth for DT and sixth for SVM.

Naive Bayes provides much better prediction of true negatives than two other classifiers. This is true even on very early stages of negotiations. However, these high true negative rates might be nullified by a large number of misclassified positive examples. To test if this is true we calculate the

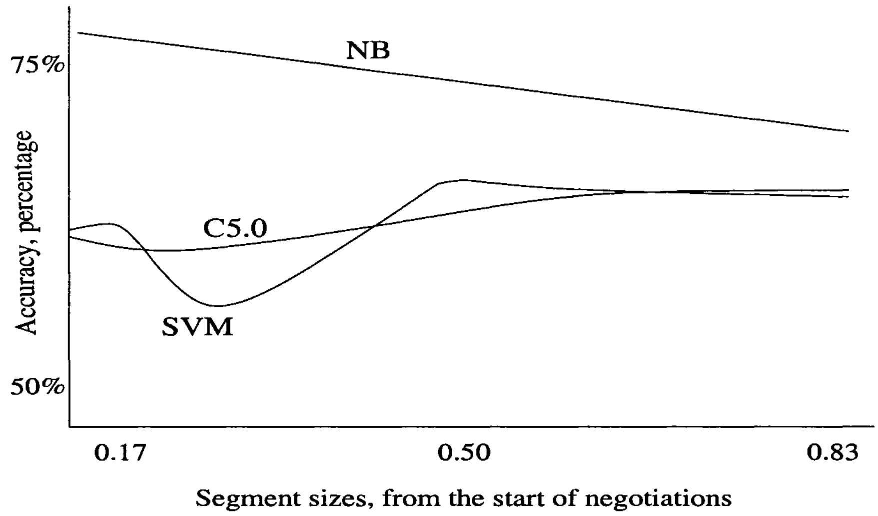

Figure 8.3: Correct classification of unsuccessful negotiations, polynomial regressions.

true positive rates. Figure 8.4 presents the regressions of the true positive rate curves. The regression polynomials have the following orders: third for NB, fifth for DT and SVM.

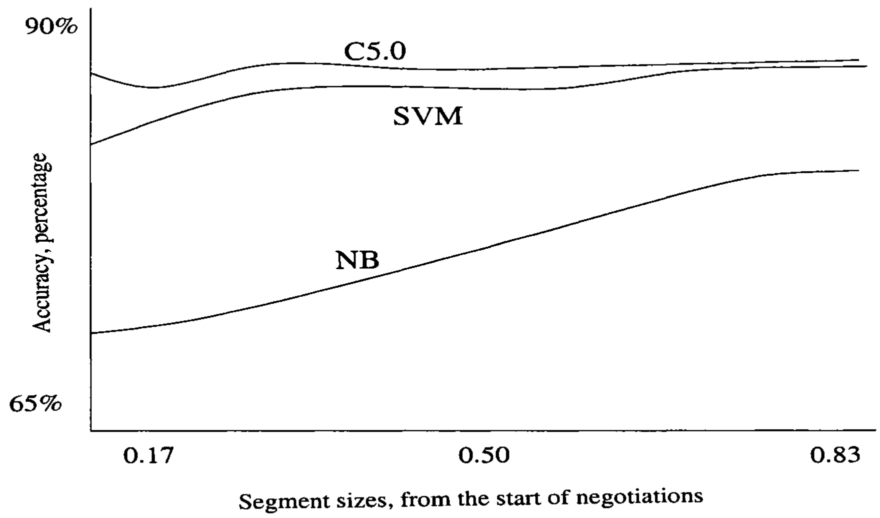

Figure 8.4: Correct classification of successful negotiations, polynomial regressions.

Regression of the true positive curves shows that NB misclassifies many positive examples, especially at the beginning of negotiations. Nevertheless, NB's performance is more steady than the performance of the other two classifiers. Thus, if a false warning can be acceptable, NB is a better predictor of the unsuccessful outcomes than C5.0 and SVM. Although the classification rates for C5.0

and SVM are almost the same, the latter has shown high fluctuation of the accuracy on the negative examples. In general, the results show that possibility of the reliable classification increases when we reach the second half of negotiations.

8.11.2 Results of incremental learning

We employ the same segmentation, feature selection and representation as described before. We apply IBK, an instance-based k-Nearest Neighbour; see Equation 2.12. $k = 10$ was found to provide higher overall accuracy than other ks. Thus, we present the results for $k = 10$. Again, we use polynomial regression to generalize accuracy obtained on different data representations; see Equation 8.2. The polynomials are third order for both curves. Results are reported in Figures 8.5 and 8.6.

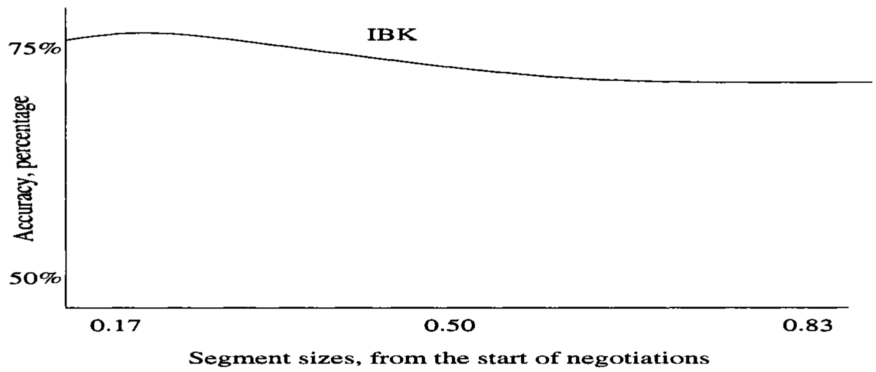

Figure 8.5: Correct classification of unsuccessful negotiations, polynomial regressions.

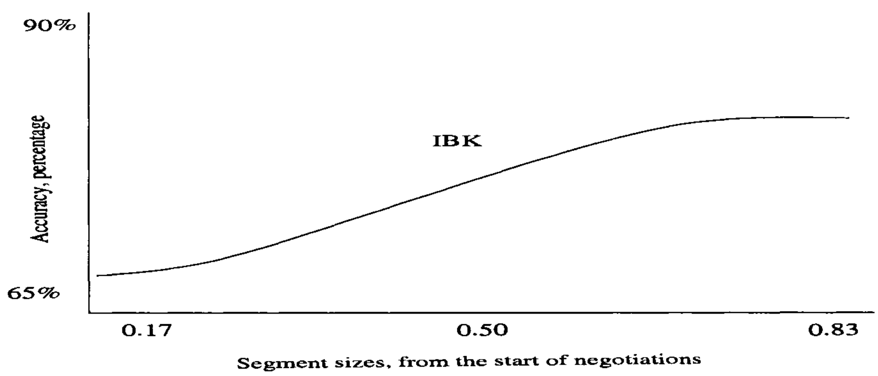

Figure 8.6: Correct classification of successful negotiations, polynomial regressions.

Incremental learning also improves the outcome predictability when a negotiation approaches its second part. The results of this section are important for future implementation in NSSs.

8.12 Strategy-related features of higher linguistic order

In this section we investigate how strategy-related features participate in dependencies between words in the data sentences .

8.12.1 Dependencies of strategy-related words on other words

Again, we have looked into relations between strategy-related words and other word types. We use the XIP parser; refer to Section 2.2.3 for a brief description. As with the negotiation-related features in Section 7.10, we apply the parser to the *Inspire* data and extract the dependencies which may contain strategy-related words:

– NUCL, a dependency between two verbal forms or a verbal form and a partical,

– NMOD, the noun modifier dependency,

– VMOD, a dependency that links a verb and the head of a nominal chunk.

These dependencies have the following modes:[3]

NUCL: MODAL, PERFECT, PASSIVE, PROGRESS, INSIST, SUBJCOMPL, PADJUNCT and VLINK modes, where MODAL, PERFECT, PASSIVE and PROGRESS modes tag when a verb form has modal, perfect, passive and progressive verbs respectively, INSIST mode tags relations between DO and the following verb, SUBJCOMPL mode tags subject compliments, PADJUNCT mode tags predicative adjuncts and VLINK mode tags BE and the related verb in a passive form;

NMOD: POST, PRE, RELATIV, DEDUCED, INFINIT and APPOS modes;

VMOD: POST, PRE, NEG, TEMP, COMPLTHAT, SENTENCE, UNSAFE, LOC, MANNER, DURATION and DEDUCED modes.

We look for dependency modes in which most parameters belong to the strategy-related category. Repeating the procedure described in Section 7.10 we find that the following modes contain mostly strategy-related words as their parameters: NUCL-PERFECT, NUCL-MODAL, NUCL-INSIST, VMOD-PRE, VMOD-NEG, VMOD-COMPLTHAT and VMOD-UNSAFE. Table 8.31 re-

[3]We provide details on NMOD and VMOD modes in Section 7.10.

ports the distribution results for the modes.

Data	NUCL-PERFECT	NUCL-MODAL	NUCL-INSIST	VMOD-PRE	VMOD-NEG	VMOD-COMPLTHAT	VMOD-UNSAFE
BSN	3090	15536	1437	7159	1979	4729	*728*
SSN	2832	15292	1196	6614	1820	4434	*754*
BUN	1179	6222	590	2737	875	1733	*289*
SUN	1189	6059	499	2683	799	1681	*290*

Table 8.31: Distribution of modes among four classes, strategy-related words.

The results show that NUCL-MODAL, VMOD-PRE, VMOD-NEG and VMOD-COMPLTHAT are uniformly distributed throughout the classes. NUCL-PERFECT is more often found in SUN than in other classes. This mode connects the auxiliary verb *have* with main verbs. Usually this grammar construction is used to emphasize actions and to make a stronger statement. This result corresponds to conclusions about language patterns made in Chapter 6. VMOD-UNSAFE is more often found in SSN and SUN than in BSN and BUN data. Table 8.32 reports the distribution of pair types across four classes. The type-the number of dependency pairs ratio is negatively correlated with the number of pairs, when the number of types is positively correlated with the number of pairs. The same holds for dependencies of negotiation-related words; refer to Section 7.10.

Data	NUCL-PERFECT	NUCL-MODAL	NUCL-INSIST	VMOD-PRE	VMOD-NEG	VMOD-COMPLTHAT	VMOD-UNSAFE
BSN	597	2235	353	3371	343	1749	452
SSN	559	2204	303	3164	304	1653	451
BUN	322	1333	206	1548	188	879	221
SUN	342	1328	186	1513	186	819	205

Table 8.32: Distribution of the dependency pair types among four classes, strategy-related words.

Tables D.5 – D.8 report 10 most frequent dependency pairs for each data class. We omit the VMOD-PRE most frequent pairs because they do not contain strategy-related words. The pairs show that the most frequent dependencies of strategy-related words coincide in four classes.

8.12.2 Dependencies between negotiation- and strategy-related words.

The NMOD-INFINIT mode finds dependencies between negotiation-related and strategy-related words in the *Inspire* data. We look at the similarities and differences between the dependencies used in four data classes. Table 8.33 reports the distribution of pair types and ten most frequent word pairs found for the dependency.

Data	# of types	**Word pairs**
BSN	2365	(pleasure,negotiate), (hope,hear), (us,pay), (us,accept), (me,accept), (you,consider),(you,accept),(me,pay), (you,make), (order,make)
BUN	1313	(you,consider), (hope,hear), (you,deliver),(you,make), (you,accept), (me,accept), (me,pay), (offer,be), (me,get),(order,make)
SSN	2549	(you,pay), (me,accept), (you,consider), (us,accept), (you,accept), (hope,hear), (pleasure,negotiate), (me,make), (us,make), (order,make)
SUN	1152	(hope,hear), (us,accept), (products,go), (control,make), (process,make), (order,get),(you,accept),(pleasure,negotiate),(you,consider), (you,find)

Table 8.33: Ten most frequent word pairs, NMOD-INFINIT.

An interesting observation comes from dependencies "*PersPron* **pay**". Successful buyers and sellers as well as unsuccessful buyers use this type of dependencies often whereas unsuccessful buyers do not. The dependencies "*PersPron* **pay**" are directly connected with the price of purchase. Recall from Section 4.6 that the price is a better predictor of the negotiation outcome than other negotiation issues. This results also correlates with the conclusion on the specifics of dependencies used by successful sellers; see Section 7.10. Note that successful sellers build more diversified NMOD-INFINIT pairs than other negotiators.

8.12.3 Time dependencies

Time in negotiations is one of the most problematic issues that we have accounted in this work. We have found, through the study of language patterns, that the intensity of electronic negotiations is directly correlated with the negotiation outcome. This confirms previous conclusion of Kersten and Zhang (2003). However, the results of the classification of the negotiation segments did not support the importance of the first part of negotiations obtained for face-to-face negotiations.

We look at the time dependencies found by XIP. Table 8.34 reports the results of the VMOD-TEMP mode for four classes.

Data	# of types	**Word pairs**
BSN	732	(deliver,days), (is,today), (accept,days),(have,time),(have,day), (pay,days), (are,today), (am,year), (is,days), (make,days)
BUN	349	(deliver,days), (pay,days), (is,today), (are,today), (have,day), (is,tomorrow), (accept,days),(have,time),(improves,time), (give,days)
SSN	733	(deliver,days), (is,today), (have,day), (is,time), (pay,days) (have,time),(deliver,time),(send,days), (is,days),(am,year)
SUN	372	(deliver,days), (supply,days), (have,days), (make,days), (have,day), (get,days), (want,days),(follow,mail), (accept,days), (out,time)

Table 8.34: Ten most frequent word pairs, VMOD-TEMP.

The reported pairs confirm that unsuccessful buyers do not pay the same attention to the price as members of other classes: the dependency *(pay,days)* is one of the most common for BSN, BUN and SSN but not for SUN. However, SUN have more different types of time dependency pairs than other classes. This partially supports the conclusion about many different small trends in the language of unsuccessful negotiations. We make this conclusion based on a poor classification of unsuccessful negotiations represented by negotiation-related words.

8.13 Conclusion

In this chapter we have explored how negotiation strategies are used in e-negotiations. Whenever possible, we compared the strategy implementation in e-negotiations and ftf-negotiations.

We base our conclusions on our machine learning results. Let us first remember how learning in binary classification problem works (Cherkassky and Muller, 1998). A learner builds a classifier. The classifier generalizes from characteristics provided by the positive examples and specializes from characteristics provided by the negative examples. The first process includes examples while the second process excludes them.

The reported results support the conclusion that the negotiation-related and strategy-related features express the difference between successful and unsuccessful negotiations better than automatically selected sets of features that do not take into account knowledge about the process. We have given an extensive experimental background to support our claims, exploring the influence on the overall performance introduced by numerical and binary feature representations.

We have shown that negotiation-related and strategy-related features catch different trends in the negotiation process: the trends of successful negotiations are better caught by negotiation-related features while the trends of unsuccessful negotiations are better caught by strategy-related features.

For e-negotiation text data we explain such results by different trends in the use of negotiation-related and strategy-related features in successful and unsuccessful negotiations. In successful negotiations the use of negotiation-related features constructs tendencies easily detected by ML methods. This condition is relaxed for unsuccessful negotiations, where common trends are not so strong. In general, ML methods have found more similarities among successful negotiations than among unsuccessful negotiations.

The reported results support a conclusion that the features selected with the knowledge about the negotiation process express the difference between successful and unsuccessful negotiations better than automatic corpus-based features that do not take into account special knowledge. We have given an extensive experimental background to support our claims.

We analyzed how different classifiers perform on e-negotiation data when the task is outcome classification. We showed that Naive Bayes is a likely candidate for predicting the true negative outcome and C5.0 is a likely candidate for predicting the true positive outcome. SVM is the most balanced classifier which gives the smallest difference between the true positive and the true negative rates. The obtained results are useful for implementation in a knowledge-based negotiation support system.

Application of the learning paradigm allowed us to analyze the structures of the two classes of negotiations: successful and unsuccessful. We have analyzed the classes of texts with respect to the preset negotiation deadline. We have concluded that the deadline influences language that the negotiators use, and that this influence can be detected by ML methods. The results have shown that the last part of a negotiation gives a more accurate classification of the negotiation outcome than the first part. We support the claim by an extensive set of experiments. The obtained results are important for applying Information Extraction and sequential learning to future studies of e-negotiation data.

We draw the following conclusion with respect to the negotiation strategies:

1. participants in successful negotiations show different attitude towards continuing the negotiations than participants of unsuccessful negotiations: the former signal to continue negotiation, the latter signal to stop negotiation;

2. participants in unsuccessful negotiations are more demanding than participants in successful negotiations;

3. denial varies in four data classes, being more implicit in successful negotiations than in unsuccessful negotiations, although "face-saving" technique is rarely used in both successful and unsuccessful negotiations;

4. implicit questions and answers are different among four classes;

5. strategies employed by sellers in successful negotiations show less authority and demand and more interest in the needs of the counterparts than strategies employed by the three other classes; this corresponds to the results of the analysis of seller and buyer behaviour in face-to-face negotiations (see Section 3.3.3).

Chapter 9

Conclusions and Future Work

Every solution breeds new problems.

Murphy's Law

In this chapter we draw overall conclusions and list the contributions of the current work. We generalize and report hypotheses supported or rejected in the dissertation. We finish this chapter with the list of directions that appeal to us as the most promising for future work. Our suggestions are rather flexible and can be applied to a variety of data as long as the data originate from communication.

9.1 Overview of completed work

Human communication provides a material unique in its richness for the study of various tasks related to humans. Humans who employ electronic means in their communication use text messages as the primary medium to fulfill their goals and obtain desirable outcomes. At the same time, the use of electronic means influences the exchange of messages, thus interfering with the activities which this communication accompanies. In our studies we have learned from the data supplied by electronic negotiations.

Electronic negotiations, conducted by email or other electronic means, are a phenomenon that recently attracted the attention of Computer Science and Artificial Intelligence. The volume of electronic negotiation increases due to the increasing volume of on-line communication in routine business. The negotiation-support systems work in such domains as legal and economic and in research and training. We worked with data gathered in bilateral electronic business negotiations, the most common type of e-negotiations.

In Chapter 2, we presented NLP and ML methods and their applicability to solving problems stated for electronic business communications. In Chapter 3, we discussed electronic negotiations. In Chapter 4, we analyzed the data that come from bilateral e-negotiations lasting long enough to provide interesting material for our research. When applicable, we analyzed face-to-face negotiation data. In Chapter 5, we built statistical and semantic models for e-negotiation data. We built statistical models for face-to-face negotiation data as well. In Chapter 6, we learned language patterns characteristic of successful and unsuccessful negotiations and of the buyer and seller roles. When applicable, we performed multi-class comparison and studied language patterns characteristic of successful and unsuccessful buyers and successful and unsuccessful sellers. In Chapter 7, we applied semantic categories for feature selection and learned the negotiation outcomes and roles within negotiations. In Chapter 8, we used negotiation strategies to find and select features representing electronic negotiations. Learning of the negotiation outcomes and roles came as a result.

In chapters 4 – 8, we have reported new results on language in electronic business negotiations and on text data of electronic business messages and on the process of electronic negotiation.

9.2 Overview of the contributions

The first phase of our research involved *corpus analysis and modelling.* We compare our data with various corpora used in NLP research. We use statistical methods to find what makes our data unique, and what common characteristics are shared by the negotiators involved in the recorded negotiations. This work presents two different data models. To find quantitative characteristics of the data we build a statistical language model. To find qualitative characteristics of the data we build a lexicon that contains semantic information for each word used.

Next, we selected (small) *subsets of features* that express essential characteristics of electronic negotiations, e.g. conditions in which these negotiations take place, rules by which they abide and goals that the participants want to achieve. The large number of actions taken by people make e-negotiations quite complex to be described by formal methods and the application of automatic analysis difficult. Moreover, when activities involve electronic means, automatic analysis is a necessity because of the vast amount of data. Additional challenges appear when activities are voluntary and do not have strict rules. In those cases people might (and do) make random choices and their behavior is unpredictable. Thus, e-negotiations are affected by uncertainty present inherently in human behavior.

From our point of view, the ***most interesting and promising contribution*** of this dissertation comes in the form of identifying two sets of features that characterize successful and unsuccessful communication respectively. Research in human communication shows that it is very difficult to find the characteristics of unsuccessful activities and communication corresponding to them. This applies to truth and falsehood recognition, where it is much easier to find features characterizing truth telling than features characterizing where somebody is not telling truth, or simply lies.

We empirically applied our research to the largest available collection of electronic negotiations, the *Inspire* data and, when appropriate, to the *Cartoon* data and the *SimpleNS* data. The size of the data requires the *application of Machine Learning* methods to explore and analyze it. We have found ML methods whose learning heuristics are favorable for the data. The last step includes identification of the data representation suitable for the methods of choice. Applying different learners we compare semi-automatically built, process-specific and process-dependent, representations of text data with the automatic ones. At the same time the data themselves pose a major challenge for learning. There are different reasons for this. For instance, the use of informal and short texts prevents characterization of this data in any standard manner which is possible for well-formatted and structured texts.

We analyze the classification performance of ML methods and their learning ability and apply our research on e-negotiation data. We apply (supervised) ML methods to different data classes, mostly in the binary classification but also in multi-class classification settings. The results provide insight into electronic negotiations and can be used in information extraction implemented within knowledge-based negotiation systems to predict the negotiation outcome and warn the negotiators when their language use may lead to the failure of negotiations.

9.3 The summary of hypotheses

We report here the hypotheses considered in the course of experiments. The hypotheses are separated into two groups according to the support received in the dissertation. Within each group we indicate whether these hypotheses relate to negotiations or learning, although some of them combine negotiation specifics with learning. It should be noted that our conclusions on negotiators' behaviour in e-negotiations support previously stated hypotheses on the behaviour of sellers and buyers in face-to-face negotiations (Drake, 2001), and relations between the intensity of e-negotiations and their outcomes (Kersten and Zhang, 2003). However, for electronic negotiations we cannot confirm the importance of language in the first phase of negotiations (Simons, 1993).

- Supported hypotheses
  - on learning:
    - machine learning methods more easily detect commonalities in successful negotiations than in unsuccessful negotiations;
    - decision-based classifiers outperform kernel and probabilistic classifiers for correct identification of the negotiation outcomes; sometimes this difference is statistically significant;
    - for electronic negotiations statistically selected features provide better multi-class classification results than knowledge-based features;
  - on language and learning:
    - language in negotiations allows the same and sometimes marginally better classification of the negotiation outcomes then the negotiation history records;
    - the negotiation-related words allow reliable classification of the negotiation outcomes, especially successful; they provide significantly better classification than statistically found features and "casual talk" words;
    - the strategic words allow reliable classification of the negotiation outcomes, especially unsuccessful; intentions are important for the outcome classification; necessity and appeal are not enough to classify the negotiation outcome;
    - the indicative words allow reliable classification of the negotiation roles;
  - on language and negotiations:
    - in electronic negotiations the negotiators use unrestricted language;
    - implicit and explicit question answering and rejection vary in successful and unsuccessful negotiations; sometimes difference is statistically significant;
    - with respect to the used language patterns:

      successful sellers differ more from unsuccessful sellers than successful buyers differ from unsuccessful buyers,

difference between successful counter-parts is bigger than between unsuccessful counter-parts;

- on negotiations:
  * the price is a better predictor of the negotiation outcomes than other negotiation issues;
  * the intensity of the discussion of the negotiation issues varies in successful and unsuccessful negotiations.

- Rejected or not confirmed hypotheses
  - on language and learning:
    * "casual talk" during negotiations helps to predict the negotiation outcome;
    * language used in the first part of electronic negotiations better predicts the negotiation outcome than language used in the second part of electronic negotiations;
    * the indicative words are features that provide reliable classification of the negotiation outcomes;
    * strategic words are features that provide reliable classification of the negotiation roles;
  - on language and negotiations:
    * the negotiation closure significantly contributes to the data predictability;
    * politeness expressions vary in successful and unsuccessful negotiations;
    * computer-mediated characteristics differ in successful and unsuccessful negotiations;
    * corpora of face-to-face and electronic negotiations exhibit different statistical characteristics.

9.4 Further development of e-negotiation means

In this dissertation we have studied the acquisition of three types of knowledge in the field of electronic negotiations: general with respect to the whole data, specific with respect to the negotiation outcome and specific with respect to the negotiators' roles.

These forms of knowledge acquisition could be combined to help in further development of e-negotiation means. We propose three areas:

- enhance the functional support in active and pro-active systems by introducing communication support;

- diversify decision support functions by making possible the analysis of texts exchanged by the negotiators;

- make an ML classifier a built-in system component in order to construct learning models of successful and unsuccessful negotiations.

Our recommendation would be text analysis which focusses on language patterns. A reliable forecast of the negotiation outcome may result from the use of patterns that indicate requirement, obligation, intensity of negotiations, etc. Our results for language patterns suggest that role-dependent text analysis shold be more effective than role-independent analysis. Patterns of appeal, logical necessity, intentions towards the subject and continuation of negotiations show how well negotiators adapt to their roles.

We propose ML on negotiation-related features if the prediction of success in negotiations is a principal goal. If the goal is the prevention of unsuccessful negotiations, then ML method benefits from receiving strategy-related features. Finally, if learning is concerned with negotiation roles, then indicative words are an appropriate representation of e-negotiation data.

Future work proposes other forms of learning for negotiation support systems. For example, Section 9.5.1 talks about incorporating bargaining data into dialogue analysis. Negotiation analysis on both types of data will certainly enrich insights into negotiation and improve the system's decision support. Another interesting research direction is the application of sentiment analysis methods to e-negotiation studies (see Section 9.5.2).

9.5 Future work

In this section we propose the directions of future work. The Internet text data can be divided into two broad categories: the mass media type category, e.g. articles and advertising, and the communication category, e.g. email and instant messaging. ML and NLP methods perform different tasks with respect to the data from the former category. Mostly, these tasks involve mining the large volumes of documents, e.g. well-structured and edited texts, written according to established, albeit relaxed, norms. Application of NLP and ML methods to the latter category of data concentrates on email classification and filtering and often works with relatively small collections of email texts. The current forecast (Allen, 2005; Baym et al., 2005) predicts that the volume of web-based communications will increase; for overview and details refer to (Mitra et al., 2005). Texts exchanged in electronic business communications present new realms and require new approaches to study them. We consider these new tasks and approaches in the proposals for future work and want to apply them to the publicly available data, e.g. the *Enron* data (Klimt and Yang, 2004).

9.5.1 Incorporating bargaining data into dialogue analysis

For advanced analysis of e-negotiation data we propose to analyze negotiators' strategic reaction to their partners. In this work we have focused on the textual data, only briefly working with the actual numbers exchanged during negotiations; see Section 4.6. The latter correspond to bargaining during negotiations (Dupont and Faure, 2002). The incorporation of bargaining data will contribute to the study of negotiation dialogues.

For example, the offer values, changing in a way favourable to the counterpart, can support claims about the intentions of negotiators of reaching an agreement (Hargie and Dickson, 2004). In this case language patterns corresponding to co-operation and achievement of mutual agreement may show the true intention of a negotiator. When the offer values or the history of the values do not support the negotiator's claim, then the language patterns may show intentional or unintentional deception of the counterpart. Incorporation of the negotiator's starting preferences and the final acceptable package will help understand the moves made during negotiation.

In the case of face-to-face negotiations Drake (2001) shows in a series of experiments that the buyer's or seller's role affects the relationship between a negotiator's opening bid and final profits and that participants may be sensitive to the competitive or cooperative climate established by the counterpart. For buyers, moderate opening offers yielded highest profits. On the contrary, extremely low or high opening offers decreased final profits. For sellers, opening offers were positively and linearly related to final profits.

Comparative studies of the explanatory ability of an approach without consideration of buyer and seller roles and an approach with consideration of the roles were carried out on a group of international and US graduate students of a Midwestern university. The emphasis of the experiments was on studying the information exchange within dyads and connecting it with the fixed-sum error. A popular variable-sum simulation negotiating the price of three appliances was employed. Pre-negotiation questionnaires were completed (role, desired profit, competitive preconception of negotiations, opening bid, profit available after the opening bid). Countries of origin, age, and gender were known for every participant. Individualism-collectivism (I-C) was measured by a Likert scale assessing I-C toward spouses, parents, kin, neighbours, friends and coworkers.

Three categories of information exchange were measured:

1. Ask for information: numerical, i.e. specific numbers; priority or preference; reaction, what the counterpart feels.

2. Give information: numerical; reaction; direction.

3. Heuristic trial and error: straightforward listing of potential prices; proposal of mutual sacrifice; proposal of mutual gain; unilateral concessions.

The buyer-seller roles were a more powerful predictor of fixed-sum error than collectivism (negative correlation). Buyers entered negotiations with nearly twice the fixed-sum errors of sellers. In the course of negotiations the fixed-sum errors of buyers and sellers did not differ significantly. The hypothesis was that collectivism is positively correlated with information exchange. The hypothesis was rejected. The finding indicates that no personal variables of negotiators could explain information exchange. This may mean that participants are sensitive to the competitive or cooperative climate established by the counterpart. The results support the importance of buyer-seller roles and the importance of seeing how negotiators react to the counterpart strategies (Rubin, 2002).

On the other hand, relations between texts of messages and accompanying offers can help analyze threats, emotions and their influence on negotiations.

9.5.2 Emotions, threats, opinions and "casual talk" in electronic negotiations

Personal power includes attractiveness, emotion, integrity, persistence and tenacity (Ströbel, 2000). It also includes complaining, bullying, emotional blackmailing, etc (Kowalski, 2003). The negotiators' main means of exercising power in e-negotiations is language. Expressing emotions and stating personal opinions are important elements of personal power as discussed in Sections 3.5 and 3.4. They are are studied by ML and NLP methods with respect to the different topics, specifically in monologues and dialogues (Devillers et al., 2003; Tomokiyo and Chollet, 2005). We propose to learn emotional aspects of e-negotiations. This topic is closely related with dialogue understanding (Zechner, 2002) and the recognition of dialogue acts (Popescu-Belis, 2005). One of the problems we foresee is that in recent dialogue analysis the regular exchange of messages and turns of interlocutors are generally presumed; for example, see Chu-Carroll and Carberry (2000). However, this is not the case in negotiations, both electronic and face-to-face; for examples of dialogues in face-to-face negotiations see Marriot (1995).

We propose to start the analysis with the identification of agreement - disagreement pairs in e-negotiation interaction; for the analysis of agreement-disagreement in the conversational interactions refer to Galley et al. (2004). After analyzing the semantic characteristics of the vocabulary of e-negotiation data we conclude that the data provide a solid ground for future research on the use of personal power in e-negotiations. We suggest the following directions:

- find *language patterns* in which participants express their threats; this is possible due to the fact that in e-negotiations negotiators express threats that can be related to the real process of negotiations, e.g. delay of replies, inability to connect with the system, or to the imaginary events, e.g. receiving approval from CEO, finding another supplier, or to personal life;
- learn how *expressions of personal power* depend on and affect negotiation moves and exchange of formal information; this study can be done by combining textual and non-textual data and concentrate on the moves of the sender of emotions, threats and opinions;
- explore *the effect of emotions, threats and opinions* on the negotiation process and outcome; this study can connect emotional and strategic choices (see Section 3.3.1) and concentrate on sender-receiver dialogue and on the reaction of a receiver on the emotions, threats, opinions of the sender.

The topic of personal power is closely connected with "casual talk" and its impact on electronic negotiations. Due to the limited means of communication, in text-based CMC people satisfy the need for "socializing" through the only available tool, namely the text of messages (Thompson and Nadler, 2002). The "social" part of text exchange is usually accompanied by personal information. This leads to the common presence of personal information in CMC text data, even when the exchange was in regard to business affairs. In Chapter 4 we show that e-negotiation data bear the features of CMC data. As we said before, during the negotiation process negotiators socialize and exchange personal information.

The interesting question arises about the relations of the casual talk and the dynamics of negotiations: does it appear at the beginning, at the middle or at the end of negotiations? A brief observation is that in the part of successful negotiations the casual talk happens after negotiators have reached an agreement.

For studying casual talk we propose the same avenues that are used in this dissertation, i.e.

- define data classes which will be meaningful for the analysis; for example, with respect to profession: students and non-students, with respect to English skills: ESL and native English, with respect to cultural background: high- and low-context cultures (Brett, 2001), where low-context cultures are characterized by explicit communication, e.g., US, Western European countries, and high-context cultures are characterized by implicit communication, e.g. Japan, other Asian countries,

- work with semantic categories, concentrating on Email addresses, Place addresses, Studies, and Hobbies; these categories indicate the exchange of personal information,

- find language patterns corresponding to defined classes and relate them to dynamics, outcome and the process of negotiations.

9.5.3 Studies of the dynamics of negotiations

So far we have approached the dynamics of negotiations by studies of the predictive power of different negotiation segments (see Section 8.10).

We propose to study the language used at different stages of negotiations. We carried out a preliminary study the initial stage of negotiations. Both company names, *Cypress* for buyers and *Itex* for sellers, have higher ranks in unsuccessful than in successful negotiations. We note that the negotiators mostly use the company name when they introduce themselves to the partner. The difference in the company name ranks points to the fact that the language of successful negotiations differs from the language of unsuccessful negotiations from the very beginning of the process. We state our hypothesis: the starting phase of successful negotiations is different in terms of language from the starting phase of unsuccessful negotiations.

To support this we first calculate that in successful negotiations *Cypress* and *Itex* together account for 3/4 of the amount they account for in unsuccessful negotiations. We also compare ranks of high frequent bigrams and trigrams containing company names that appear among 500 most frequent bigrams and 700 most frequent trigrams.

Cypress appears in successful negotiations in one bigram (*Cypress Cycles*, $rank_s = 96$), two trigrams (*Cypress Cycles Dear*, $rank_s = 467$; *of Cypress Cycles*, $rank_s = 532$). In unsuccessful negotiations *Cypress* appears in one bigram (*Cypress Cycles*, $rank_u = 55$), five trigrams (*represent Cypress Cycles*, $rank_u = 359$; *Cypress Cycles is*, $rank_u = 391$; *I represent Cypress*, $rank_u = 420$; *of Cypress Cycles*, $rank_u = 474$; *Cypress Cycles Dear*, $rank_u = 625$).

Itex appears in successful negotiations in one bigram (*Itex Manufacturing*, $rank_s = 285$), and in no trigrams. In unsuccessful negotiations *Itex* appears in one bigram (*Itex Manufacturing*, $rank_s = 295$), and in no trigrams.

The comparison of N-gram ranks leads us to conclude that in both successful and unsuccessful negotiations *Cypress* (buyers) is used more frequently than *Itex* (sellers). *Cypress* is noticeably more frequently used in unsuccessful than in successful negotiations. The use of *Itex* in successful negotiations is more frequent than in unsuccessful ones. Our working hypothesis is that buyers through the use of of the company name *Cypress* want to emphasize importance of purchase if it happens and importance of their demands and requirements to the products. We suggest that sellers in successful negotiations are more business-oriented than sellers in unsuccessful negotiations.

In the final part of a negotiation, or closure, we found that only 15% of unsuccessful negotiations, in which dialogue has occurred, had a negative closure message. 3% of such negotiations had a positive closure message. Among successful negotiations approximately 50% had a positive closure message. Remarkably, when the text data of the closure messages were deleted from the data, the cross-entropy of the statistical language model has changed by only 0.05. Taking into account the fact that the cross-entropy change less than 0.14 does not make substantial difference in the model's applicability to the data (Chen and Goodman, 1998; Rosenfeld, 2000), we can say that closure messages do not contribute significant information to the data or the negotiation process. We call the part of the data without closure messages a *prefix*. The statistical analysis of the prefix data has supported the hypothesis about different initial stages in successful and unsuccessful negotiations.

For future work we propose to start by building a dynamic model of negotiations and identifying a set of relevant dynamic attributes, e.g. attributes that change with time. The study of dynamic attributes is challenging. First, the dynamic attributes are hard to quantify in advance. This would be true in most practical scenarios. For instance, *offers and ratings*, *preference structures* are two such attributes used by Kersten and Zhang (2003) that are generally dynamic in nature. There is a high probability that these attributes change their values over time and over the course of negotiations. Other factors might affect such attributes. Here is a possible scenario. At the beginning of the negotiation, an offer is unacceptable to the buyer; during the course of the negotiations the buyer may accept this offer when it comes as part of some package deal that tends to be a compromise acceptable to both parties. Such situations might strongly affect the dynamic attributes (Drake, 2001).

Second, the dynamic attributes require specific learning approach. In current studies we have applied the classical supervised learning approach. For future work we suggest incorporating sequential supervised learning and sequence classification. The former allows to work with the dynamics of negotiations and the latter allows to predict an outcome from the input sequence. For review of ML methods on sequential learning see Dietterich (2002).

9.5.4 The study of heterogenous data

We have shown that i.i.d. assumption does not always hold for e-negotiation data if the data contain bargaining, textual and protocol data; see Section 4.2.1. In current Data Mining and Machine Learning fields such data are referred as *heterogenous* data. Currently the studies of heterogenous data are closely related to image classification (Jin and Liu, 2004) and to gene classification (Li et al., 2003), where i.i.d. assumption of data distribution does not hold. For the heterogenous data the learning methods of choice are ensemble methods; refer to Dietterich (1998b) where the case of the decision tree based ensembles is discussed and to Pavlidis et al. (2001) where the ensembles of kernels built by support vector machine are used. Another popular classifier for use in ensemble methods is Naive Bayes.

However, these studies do not usually address the case where data come from multiple sources. The latter is studied under the multiple views umbrella in the unsupervised and semi-supervised learning; the work on two different views on web pages written by Blum and Mitchell (1998) is considered to be classical. The emotion detection through multiple views, using image data, is studied by Cohen et al. (2003). The studies of multiple views involve the use of clustering methods and in its theoretical part are focused on finding the appropriate distance measures; for example, see

I.S.Dhillon et al. (2004). The studies of multiple views are interconnected with the hierarchical classification where the hierarchy of e-negotiation data can be considered as the following: message fragment – message – negotiation. This also will require the employment of evaluation measures specific to the hierarchical classification. For various F-measures used in hierarchical classification refer to Kazawa et al. (2005).

We suggest to extend the studies of e-negotiation data by studying the heterogeneous data obtained from different sources, e.g. interaction, bargaining, and protocol data. The *proposed task* is to classify negotiation outcomes based on different data sources. The *proposed method* is to construct the ensemble of classifiers or the similar approaches. Based on the classification results presented in Chapters 7 and 8 we suggest to use either boosting of C5.0 or the combination of different kernels built by SVM. The basic idea is to apply the same learner to build classifiers on different types of data, say, bargaining/numerical and interaction/textual, and build the final classifier from their outputs.

The boosting of C5.0 is suggested because of performance of C5.0 on classifying of positive e-negotiation examples (successful negotiations) in both numerical and textual data. The high accuracy was detected for a wide range of data representations, in fact for all 19 data representations used for classification of the e-negotiation outcomes. At the same time, C5.0 classifies the negative e-negotiation examples (unsuccessful negotiations) more accurately than Decision List Machine which also showed a high accuracy on the positive examples.

The use of SVM is suggested because it gives the most balanced performance in terms of the true positive and true negative classification rates. This is the case for all 19 data representations when the negotiation outcomes were classified.

We do not suggest to use Naive Bayes because of its poor performance in identification of positive results from the textual data, extreme instability of performance and high dependency on feature representation while classifying the numerical data.

9.6 Directions of research

Finally, we propose the research directions for electronic negotiations that might benefit both Negotiation Theory and Artificial Intelligence, especially Natural Language Processing.

Many aspects of e-negotiations remain beyond the scope of this dissertation. For example, we omit the relations between gender and negotiation behaviour (Sutter et al., 2003). From our point of view, this subject has been well-studied. The same applies to the studies of gender and CMC (Mitra et al., 2005) and gender and text classification (Koppel et al., 2002). Analysis of relations between age and CMC is another well-developed area (Pujolar, 2000). For the list of additional references see Section 3.8.

Our suggestions focus on behavioural and cultural issues of negotiations.

9.6.1 Studying e-negotiation behavioural biases

We propose to investigate how behavioural biases affect e-negotiation. This direction is interconnected with the psychological study of negotiations; for an overview see Bazerman et al. (2000). Till early 1980ties, prescriptive research on negotiations focused primarily on game theory, the mathematical analysis of fully rational negotiators. In early 1980ties the focus switched to providing the best advice to a negotiator (Raiffa, 1982). Behavioural Decision research argues that people rely on simplifying strategies, or cognitive heuristics, that have systematic and predictable biases. In bilateral face-to-face negotiations negotiators tend to:

- be more concessionary to a positively framed specification of the negotiation than to a negatively framed specification,
- be inappropriately affected by readily available information,
- be overconfident and overly optimistic about the likelihood of attaining outcomes that favour them,
- falsely assume that the negotiation pie is fixed and miss opportunities for mutually beneficial trade-offs between the parties,
- falsely assume that their preferences on issues are incompatible with those of their opponent,
- escalate conflict even when a rational analysis would dictate a change in strategy,
- ignore the perspective of other parties,
- reactively devalue any concession made by the opponent,
- overweigh the views that favour them,

- remember better facts that favoured them.

To look for behavioural biases we suggest working with the exchanges within a negotiating pair, or dyad.

We also suggest studying how behavioural biases and social relations within negotiations are interrelated. In social relationships in negotiations two levels of research are well distinguished. The first level studies the influence of social context (what the counterpart gets, and relations with him) on the judgement and preference of a negotiator; the second level studies how social relationships within dyads can influence the negotiation process and outcome (seemingly irrational behaviour of a negotiator may result in negotiation outcomes that outperform game-theoretic models).

How negotiators understand the game is of critical importance to how they play the game; see Section 3.8. Integrative negotiation depends fundamentally on the negotiators' ability to trade issues against each other. An unwillingness to consider trades because of sacred values may constrain the set of permissible agreements. Negotiators can hurt themselves by claiming certain issues to be sacred, thus constraining the game and the ability to find integrative trade-offs. Within a dyad negotiators share the understanding of the situation, roles, and the rules of acceptable behaviour. Their behaviour is developed in interaction. Communication research takes as a given that negotiation is dynamic and built by both participants (Hargie and Dickson, 2004; Brett, 2001). All this is expressed in the exchange within dyads.

The important characteristic of the exchanges within a dyad in the e-negotiations is that they differ from "regular" spoken or written dialogues. E-negotiation[1] messages include exchange of formal business information, sometimes written in a structured format, and informal business information. They can also include casual talk (see Section 3.5). The different type of dialogues requires different, and possibly novel, models and methods of studies. As mentioned in Section 3.5, the CMC texts, to which e-negotiation dialogues belong, include many speech-related elements. That is why we include references to speech recognition models and methods in the models as well as the ones proposed for the future work.

We propose to implement role-dependent information in models of e-negotiation dialogues, as it was done in the speech recognition model for court procedures in (Kenne and O'Kane, 1996), or implement domain-specific information through the topic representation (Popescu-Belis et al., 2004).

[1]Recall that we consider business e-negotiation

We propose to work on existing learning methods and approaches to the dialogue analysis and incorporate information which comes from decision-support functions of NSS or peculiarities of CMC texts; for the list of references see Sections 2.7 and 3.8.

9.6.2 Language, culture and negotiations

The Inspire collection with its population of more than 5000 data contributors, coming from more than 50 countries of all continents, gives a unique opportunity to study relations between the language of e-negotiations, negotiation behaviour and culture. Below we summarize a few recent results obtained on the culture and international business negotiations. These results can be a starting point for future studies of culture in e-negotiations.

Culture uniquely represents a social group by the characteristic behaviour, norms, and values of its members, and by legal, economic, social, and religious institutions. Culture influences negotiation interests and priorities (what is important), behavioral strategies, and context (laws, economic conditions). Cross-cultural differences in the negotiation game can be conceptualized along four basic dimensions: collectivism-individualism, power distance, communication context, and concept of time. Most research has focused on the first dimension (Brett, 2001; Faure, 2002; Vetschera et al., 2004).

In face-to-face conversations the use of language patterns by representatives of different social groups was intensively and extensively studied (Preiser, 1986), although without applying ML and NLP tools. We suggest that the relations between the professional background of negotiators and their behaviour in negotiations, especially in international negotiations, is a promising direction of future research. The topic of the influence of a profession on a negotiator is understudied in current research (Sjostedt, 2003). The e-negotiation data gives opportunities to study

- similarities between the language of negotiators belonging to the same professional group, including differences in syntax, semantics and pragmatics,
- differences between the language of negotiators belonging to various professional groups, including differences in syntax, semantics and pragmatics,
- implementation of the negotiation strategies and the influence strategies by different professional groups.

Salacuse (1998) has proposed that culture influences ten negotiation factors **Goal, Attitudes,**

Personal Styles, Communications, Time Sensitivity, Emotionalism, Agreement Form, Agreement Building, Team Organization, Risk Taking. Culture causes different responses

Negotiation Factors	**Range of Cultural Responses**
Goal	Contract $\longleftrightarrow$ Relationship
Attitudes	WinLose $\longleftrightarrow$ WinWin
Personal Styles	Informal $\longleftrightarrow$ Formal
Communications	Direct $\longleftrightarrow$ Indirect
Time Sensitivity	High $\longleftrightarrow$ Low
Emotionalism	High $\longleftrightarrow$ Low
Agreement Form	Specific $\longleftrightarrow$ General
Agreement Building	Bottom Up $\longleftrightarrow$ Top Down
Team Organization	One Leader $\longleftrightarrow$ Consensus
Risk Taking	High $\longleftrightarrow$ Low

Table 9.1: The impact of culture on negotiations.

between two extremes. **Goal: Contract or Relationship** defines to the purpose of the negotiators. American negotiators mostly aim to sign the contract, while for many cultural Asian groups relationships are important, with the emphasis of utility over friendship (Smith, 2000). **Attitudes: WinLose or WinWin** distinguishes between distributive and integrative negotiation. **Personal Styles: Informal or Formal** corresponds to interactions between negotiators. **Communications: Direct or Indirect** distinguishes between the direct, or simple, and indirect, or complex, manner of communications. **Time Sensitivity: High or Low** relates to different attitude towards time, e.g. polychronic for East Asians and monochronic for Americans, and the length of time devoted to negotiation itself. **Emotionalism: High or Low** distinguishes between appropriateness and inappropriateness of displaying emotions during negotiations. Positive moods tend to increase the chances of selecting a cooperative strategy. Angry negotiators are less accurate in judging the interests of their counterparts and achieve lower joint gains (Bazerman et al., 2000). **Agreement Form: Specific or General** relates to preferable form of written contract agreement. Americans prefer detailed contracts, while Asians prefer general principles, which correlates with their understanding that negotiation is the beginning of relationship (see the Goal factor). **Agreement Building: Bottom Up or Top Down** distinguishes between negotiators starting with specific points and summarizing totals at the end, and negotiators starting with general principles and going into details later. Americans prefer the former approach, while the French and East Asians prefer the latter. **Team Organization: One Leader or Consensus** relates to intra-group organization. **Risk Taking: High or Low** emphasizes whether negotiators divulge information, try new approaches and tolerate uncertainties in a course of actions. Due to the bilateral nature of the Inspire negotiations we do not consider **Team Organization**. The pre-determined structure of Inspire agreements and fixed negotiated issues make **Agreement Form** and **Agreement Building** redundant for our research.

There are different cultural approaches to negotiation. Culture as shared values (CASV) concentrates on contrasting cultural values represented by individualism and collectivism, particularly the contrast between competitive and cooperative approaches. The strongest argument against CASV that it reinforces a largely untested assumption that intra- and intercultural negotiating behaviour are similar. Evidence suggests that this assumption is incorrect (Drake, 2001). Intercultural dyads are more involved and more likely to solicit partner participation than intracultural dyads. Culture in context (CIC) treats negotiators as regulators of a complex negotiation system. For communication researchers, CIC may appear the more heuristically valuable perspective because it is consistent with assumptions of interdependence and mutual accommodation. However, these approaches were not tested on large data. With access to the *Inspire* and other data future work will apply ML and NLP methods to the data of social studies.

9.7 Conclusion

We finish the dissertation with our favourite *Murphy's Law*:

What we learn after we know it all is what counts.

References

M. Aery and C. Sharma. 2004. emailsift: Adapting graph mining techniques for email classification. Technical Report Technical Report CSE-2004-7, Department of Computer Science and Engineering, University of Texas at Arlington. http://www.cse.uta.edu/research/publications/Downloads/CSE-2004-7.pdf.

D. E. Allen and R. F. Guy. 1974. *Conversation Analysis: The Sociology of Talk.* Mouton, The Hague.

M. Allen. 2005. Content or connectivity: understanding the meaning of the internet. In *Proc. of Internet Generations, the 6th International and Interdisciplinary Conference of the Association of Internet Researchers*, pages –.

D. Angell and B. Heslop. 1994. *The Elements of E-mail Style.* Addison-Wesley.

M. Argyle, A. Furnham, and J. Graham. 1981. *Social situations.* Cambridge: Cambridge University Press.

Y. Bakos. 1998. The emerging role of electronic marketplaces on the internet. *Communications of the ACM*, 41(8):35–42.

N. Baym, F. Turner, D. Weisberg, and S. Cornish. 2005. Generations of internet reception, 1968-2005. In *Proc. of Internet Generations, the 6th International and Interdisciplinary Conference of the Association of Internet Researchers*, pages –.

M. H. Bazerman, J. R. Curhan, D. A. Moore, and K. L. Valley. 2000. Negotiation. http://www.findarticles.com/cf0.

S. Bickel, P. Haider, and T. Scheffer. 2005. Learning to complete sentences. In *Proc of ECML'2005.* http://www.informatik.hu-berlin.de/ scheffer/publications/ecml2005-seco.pdf.

J. Blaták, E. Mráková, and L. Popelínsky. 2004. Supporting user task based conversations via email. In *Proc ACL 2004*, pages 226–229.

A. L. Blum and P. Langley. 1997. Selection of relevant features and examples in machine learning. *Artificial Intelligence*, 97:245–271.

A. Blum and T. Mitchell. 1998. Combining labeled and unlabeled data with co-training. In *Proc COLT'1998*, pages 92–100.

J. Boparai and J. Kay. 2002. Supporting user task based conversations via email. In *Proc 7th Australasian Document Computing Symposium.*

M. Bowling. 2005. Convergence and no-regret in multiagent learning. In *Advances in Neural Information Processing Systems(Proc of NIPS'2004)*, volume 17, pages 209–216. The MIT Press.

J. M. Brett. 2001. *Negotiating Globally.* Jossey-Bass, San Francisco.

E. Brill. 1995. Transformation-based error-driven learning and natural language processing: A case study in part of speech tagging. *Computational Linguistics*, 21(4):543–565. http://www.cs.jhu.edu/brill/.

P. Brown and S. Levinson. 1978. Universals in language usage: Politeness phenomena. In E. N. Goody, editor, *Questions and Politeness*, pages 256–289. Cambridge University Press.

L. Buddress, M. E. Smith, and A. Raedels. 2003. The path to effective online negotiations. In *Proc 88th Annual International Supply Management Conference.* http://www.ism.ws/ResourceArticles/Proceedings/2003/BuddressCE.pdf.

B. D. Burns and A. P. Danyluk. 2000. Feature selection vs theory reformulation: A study of genetic refinement of knowledge-based neural networks. *Machine Learning*, 38(1-2):89–107.

N. Burrell and R. Koper. 1998. The efficacy of powerful/powerless language on attitudes and source credibility. In M. Allen and R. Preiss, editors, *Persuasion: advances through meta-analysis*, pages 203–215. Hampton Press.

J. Carey. 1980. Paralanguage in computer mediated communication. In N. K. Sondheimer, editor, *The 18th annual meeting of the association for computational linguistics and parasession on topics in interactive discourse: Proc*, pages 67–69.

P. Carnevale and D. Pruitt. 1992. Negotiation and mediation. *Annual Review of Psychology*, 43:531–582.

M. Caropreso, S. Matwin, and F. Sebastiani. 2001. A learner-independent evaluation of the usefulness of statistical phrases for automated text categorization. In A. G. Chin, editor, *Text Databases and Document Management: Theory and Practice*, page 78102. Idea Group Publishing.

J. Carroll, G. Minnen, and T. Briskoe. 2003. Parser evaluation. In A. Abeillé, editor, *Treebanks: Building and Using Parsed Corpora*, pages 299–316. Kluwer Academic Publishers.

C. Cellich and S. C. Jain. 2004. *Global Business Negotiations : A Practical Guide.* Thomson, South-Western.

J.-P. Chanod et al., 2001. *Xerox Incremental Parser.* Xerox Research Centre Grenoble.

E. Charniak. 1993. *Statistical Language Learning.* The MIT Press.

S. F. Chen and J. Goodman. 1998. An empirical study of smoothing techniques for language modeling. Technical Report Center for Research in Computing Technology Tech. Report TR -10-98, Harvard University.

V. Cherkassky and F. Muller. 1998. *Learning from Data.* Wiley.

J. Chu-Carroll and S. Carberry. 2000. Conflict resolution in collaborative planning dialogues. *International Journal of HumanComputer Studies*, 53(6):969–1015.

K. Church and P. Hanks. 1989. Word association norms, mutual information and lexicography. In *Proc 27th Annual Meeting of the ACL*, pages 76–83.

S. Climent, J. Mor, A. Oliver, M. Salvatierra, I. Snchez, M. Taul, and L. Vallmanya. 2003. Bilingual newsgroups in catalonia: A challenge for machine translation. *Journal of Computer-Mediated Communication*, 9(1). http://www.ascusc.org/jcmc/vol9/issue1/climent.html.

I. Cohen, N. Sebe, F. G. Cozman, and T. S. Huang. 2003. Semi-supervised learning for facial expression recognition. In *Proc of the 5th ACM SIGMM international workshop*, pages 17–22.

W. Cohen, V. Carvalho, and T. Mitchell. 2004. Learning to classify email into "speech acts". In *Proc EMNLP*, pages 309–316.

W. Cohen. 1996. Learning rules to classify emails. In *Papers from AAAI Spring Symposium on Machine Learning in Information Access*, pages 18–25.

R. Cooley, P.-N. Tan, and J. Srivastava. 2000. *Discovery of Interesting Usage Patterns from Web Data.* Springer-Verlag.

T. Cormen, C. Leiserson, and R. Rivest. 2000. *Introduction to Algorithms.* Cambridge: MIT Press.

R. Cowie, E. Douglas-Cowie, B. Apolloni, J. Taylor, A. Romano, and W. Fellenz. 1999. What a neural net need to know about emotion words. In N.Mastorakis, editor, *Computational Intelligence and Applications (CSCC'99)*, pages 109–114.

R. Cowie, E. Douglas-Cowie, S. Savvidou, E. McMahon, M. Sawey, and M. Schroder. 2000. "feeltrace": an instrument for recording perceived emotion in real time. In *Proc ISCA Workshop on Speech and Emotion*, pages 19–24.

N. Cristianini and J. Shawe-Taylor. 2000. *An Introduction to Support Vector Machines and other kernel-based learning methods.* Cambridge University Press.

N. Cristianini, J. Shawe-Taylor, and H. Lodhi. 2001. Latent semantic kernels. In *Proc Eighteenth International Conference on Machine Learning (ICML'2001)*, pages 66–73. Morgan Kaufmann.

K. Crowston and E. Kammerer. 1998. Communicative style and gender differences in computer-mediated communications. In B. Ebo, editor, *Cyberghetto or Cybertopia: Race, Class and Gender on the Internet*, pages 185–204. Praeger.

G. Demetriou and E. Atwell. 2001. A domain-independent semantic tagger for the study of meaningassociations in english text. In *Proceedings of the Fourth International Workshop on Computational Semantics*, pages 67–80, Tilburg, Netherlands.

L. Devillers, L. Lamel, and I. Vasilescu. 2003. Emotion detection in task-oriented spoken dialogs. In *Proc ICME*, Baltimore.

T. G. Dietterich. 1998a. Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Computation*, 10:18951923.

T. G. Dietterich. 1998b. An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. *Machine Learning*, 40(2):139–157.

T. G. Dietterich. 2002. Machine learning for sequential data: A review. In T.Caelli, editor, *Lecture Notes in Computer Science.* Springer-Verlag.

L. E. Drake. 2001. The culture-negotiation link. *Human Communication Research*, 27(3):317–349.

D. Druckman and P. T. Hopman. 2002. Content analysis. In V. Kremenyuk, editor, *International Negotiation: Analysis, Approaches, Issues*, pages 288–314. Jossey-Bass, second edition.

R. Duda, P. Hart, and D. Stork. 2000. *Pattern Classification.* Wiley, second edition.

C. Dupont and G. O. Faure. 2002. The negotiation process. In V. Kremenyuk, editor, *International Negotiation: Analysis, Approaches, Issues*, pages 39–63. Jossey-Bass, second edition.

L. Fais and K. Ogura. 2001. Discourse issues in the translation of japanese e-mail. In *Proc of Conference of the Pacific Association for Computational Linguistics, PACLING 2001.* http://afnlp.org/pacling2001/pdf/fais.pdf.

G. O. Faure. 2002. International negotiation: The cultural dimension. In V. Kremenyuk, editor, *International Negotiation: Analysis, Approaches, Issues*, pages 392–415. Jossey-Bass, second edition.

C. Fellbaum. 1998. *WordNet: An Electronic Lexical Database.* Cambridge: MIT Press.

G. Forman. 2003. An extensive empirical study of feature selection metrics for text classification. *Journal of Machine Learning Research*, 3:1289–1305.

G. Forman. 2004. A pitfall and solution in multi-class feature selection for text classification. In *Proceedings of 21st International Conference on Machine Learning (ICML'2004)*, pages 107–114. Morgan Kaufmann.

W. N. Francis and H. Kucera, 1979. *Brown Corpus Manual.* http://helmer.aksis.uib.no/icame/brown/bcm.html.

W. Frawley. 1992. *Linguistic Semantics.* Lawrence Erbaum Associates.

E. Gabrilovich and S. Markovitch. 2004. Text categorization with many redundant features: Using aggressive feature selection to make svms competitive with c4.5. In *Proc of the 21st International Conference on Machine Learning (ICML-2004)*, pages 321–328. Morgan Kaufmann.

M. Galley, K. McKeown, J. Hirschberg, and E. Shriberg. 2004. Identifying agreement and disagreement in conversational speech: Use of bayesian networks to model pragmatic dependencies. In *Proc ACL 2004*, pages 669–676.

J. Gebauer and A. Scharl. 1999. Between flexibility and automation: An evaluation of web technology from a business process perspective. *Journal of Computer-Mediated Communication*, 5(2). http://www.ascusc.org/jcmc/vol5/issue2/gebauer.html.

I. J. Good. 1953. The population frequencies of species and estimation of population parameters. *Biometrika*, 40(3,4):237–269.

S. T. Gries. 2003. *Multifactorial Analysis in Corpus Linguistics.* Continuum, New York.

I. Guyon and A. Elisseeff. 2003. An introduction to variable and feature selection. *Special issue on variable and feature selection, Journal of Machine Learning Research*, 3:1157 – 1182.

U. Hahn, M. Romacker, and S. Schulz. 1999. How knowledge drives understanding matching medical ontologies with the needs of medical language processing. *Artificial Intelligence in Medicine*, 15:25–51.

O. Hansberger. 1985. *Say it: 888 expressions pour negocier en anglais.* Mazarine, Paris.

O. Hargie and D. Dickson. 2004. *Skilled Interpersonal Communication: Research, Theory and Practice.* Routledge, fourth edition.

O. Hargie, D. Dickson, and D. Tourish. 1999. *Communication in management.* Gower, Aldershot.

O. Hargie. 1997. Interpersonal communication: a theoretical framework. In O. Hargie, editor, *The handbook of communication skills*, pages 29–63. Routledge, London.

T. Hastie, R. Tibshirani, and J. Friedman. 2001. *The Elements of Statistical Learning.* Springer-Verlag.

M. Hearst. 2003. What is text mining? http://www.sims.berkeley.edu/ hearst/text-mining.html.

D. G. Hendry. 2004. Communication functions and the adaptation of design representations in interdisciplinary teams. In *Proceedings of the 2004 conference on Designing interactive systems*, pages 123–132.

R. Herbrich. 2002. *Learning Kernel Classifiers.* The MIT Press.

S. C. Herring. 2001. Computer-mediated discourse. In D. Tannen, D. Schiffin, and H. Hamilton, editors, *Handbook of discourse analysis*, pages 612–634. Oxford, Blackwell.

C. W. Holsapple and A. B. Whinston. 2000. *Decision Support Systems: A Knowledge-based Approach.* West Publishing Company.

J. Hu. 2003. Message in the bottleneck. http://news.com.com/2009-1033-992348.html.

B. Hutchinson. 2004. Acquiring the meaning of discourse markers. In *Proceedings of ACL'04*, pages 685–692.

I.S.Dhillon, Y.Guan, and B.Kulis. 2004. Kernel k-means, spectral clustering and normalized cuts. In *Proc of The Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 551–556.

2001. ispell. http://www.gnu.org/directory/text/wordproc/ispell.html.

N. Japkowicz. 2001. Supervised versus unsupervised binary-learning by feedforward neural networks. *Machine Learning*, 42(1/2):97–122.

M. Jarmasz and S. Szpakowicz. 2004. Roget's thesaurus and semantic similarity. In N. Nicolov, K. Bontcheva, G. Angelova, and R. Mitkov, editors, *Current Issues in Linguistic Theory. Recent Advances in Natural Language Processing III: Selected Papers from RANLP 2003*, volume 206, pages 111–120. John Benjamins, Amsterdam/Philadelphia.

R. Jin and H. Liu. 2004. Switch: A novel approach to ensemble learning for heterogeneous data. In *Lecture Notes in Computer Science, Proc of ECML'2004*, pages 560–562. Springer-Verlag.

T. Joachims. 2002. *Learning to Classify Text Using Support Vector Machines : Methods, Theory and Algorithms.* Kluwer Academic Publishers.

K. Jokinen, T. Hurtig, K. Hynna, K. Kanto, M. Kaipainen, and A. Kerminen. 2001. Self-organizing dialogue management. In *Proc The Second Workshop on Natural Language Processing and Neural Networks (NLPNN 2001)*, pages 77–84.

D. Jurafsky and J. H. Martin. 2000. *Speech and Language Processing.* Prentice Hall.

S. M. Katz. 1987. Estimation of probabilities from sparse data for the language model component of a speech recognizer. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 35(3):400–401.

H. Kazawa, T. Izumitani, H. Taira, and E. Maeda. 2005. Maximal margin labeling for multi-topic text categorization. In *Advances in Neural Information Processing Systems(Proc of NIPS'2004)*, volume 17, pages –. The MIT Press.

P. E. Kenne and M. O'Kane. 1996. Topic change and the local perplexity in the spoken dialogs. In *Proc Fourth International Conference on the Spoken Language Processing.* http://www.asel.udel.edu/icslp/cdrom/vol2.htm.

G. E. Kersten and S. J. Noronha. 1999. Www-based negotiation support: Design, implementation, and use. *Decision Support Systems*, 25:135–154.

G. E. Kersten et al. 2002. Electronic negotiations, media and transactions for socio-economic interactions. http://interneg.org/enegotiation/index.html.

G. E. Kersten and G. Zhang. 2003. Mining inspire data for the determinants of successful internet negotiations. *Central European Journal of Operational Research*, 11(3):297–316.

G. E. Kersten. 2000. Modeling distribution and integrative negotiations. review and revised characterization. *Group Decision and Negotiation*, 10(6):493–514.

A. Kilgarriff. 2001. Comparing corpora. *International Journal of Corpus Linguistics*, 6(1):97–133.

S. Kiritchenko and S. Matwin. 2001. Email classification with co-training. In *Proc. of CASCON 2001*, pages 192–201.

S. Kiritchenko, S. Matwin, and S. Abu-Hakima. 2004. Email classification with temporal features. In M. Klopotek, S. Wierzchon, and K. Trojanowski, editors, *Intelligent Information Processing and Web Mining*, pages 523–533. Springer.

B. Kirkpatrick. 1998. *Rogets Thesaurus of English Wordsand Phrases*. Harmondsworth, Middlesex,England: Penguin.

B. Klimt and Y. Yang. 2004. The enron corpus: A new dataset for email classification research. In *Proc of ECML'2004*, pages 217 – 226.

R. Kneser and H. Ney. 1995. Improved backing-off for m-gram language modeling. In *Proc of the IEEE International Conference on ASLP*, volume 1, pages 181–184.

R. Kohavi. 1995. Wrappers for performance enhancement and oblivious decision graphs. PhD thesis, Department of Computer Science, Stanford University.

M. Koppel, S. Argamon, and A. Shimoni. 2002. Automatically categorizing written texts by author gender. *Literary and Linguistic Computing*, 17(4):401–412.

R. Kosala and H. Blockeel. 2000. Web mining research: A survey. *ACM SIGKDD Explorations*, 2(1):1–15.

R.M. Kowalski. 2003. *Complayining,Teasing, and Other Annoying Behaviors*. Yale University Press.

A. H. F. Laender, B. A. Ribeiro-Neto, A. S. da Silva, and J. S. Teixeira. 2002. A brief survey of web data extraction tools. *ACM SIGMOD Record*, 31(2):84–93.

Y. Lebrun. 1965. *Can and May in Present-day English*. Presses Universitaires de Bruxelles.

G. N. Leech. 1987. *Meaning and the English Verb*. Longman, second edition.

T. Li, S. Zhu, Q. Li, and M. Ogihara. 2003. Gene functional classification by semi-supervised learning from heterogeneous data. In *Proceedings of the 2003 ACM symposium on Applied computing*, pages 78 – 82.

C. Li, J.A. Giampapa, and K. Sycara. 2004. Bilateral negotiation decisions with uncertain dynamic outside options. In *Proc The 1st IEEE International Workshop on Electronic Contracting*, pages 54–61.

W. Lin, R. Yangarber, and R. Grishman. 2003. Bootstrapped learning of semantic classes from positive and negative examples. In *Proc ICML-2003 Workshop on The Continuum from Labeled to Unlabeled Data.* http://nlp.cs.nyu.edu/publication/papers/lin-icml03.pdf.

D. Lin. 2001. Latat: Language and text analysis tools. In *Proceedings of Human Language Technology Conference*, pages 222–227.

2004. Lancaster corpus of british english. http://helmer.aksis.uib.no/icame/lobman/lob1.html.

E. Ludovik, R. Zachnoki, and J. Cowie. 1999. Language recognition for mono-and multi-lingual documents. In *Proc VEXTAL Conference, Venice*, pages 209–214.

J. Lyons. 1995. *Linguistic semantics: an introduction.* Cambridge University Press.

J. Mahoney. 1996. Combining symbolic and connectionist learning methods to refine certainty-factor rule-bases. Ph.D. Dissertation, University of Texas, Austin.

B. Mandelbrot. 1954. Structure formelle des textes et communication. *Word*, 10:1–27.

C. D. Manning and H. Schutze. 1999. *Foundations of Statistical Natural Language Processing.* The MIT Press, fourth edition.

M. Marchand and M. Shah. 2005. Pac-bayes learning of conjunctions and classification of gene-expression data. In *Advances in Neural Information Processing Systems(Proc of NIPS'2004)*, volume 17, pages 881–888. The MIT Press.

M. Marchand and J. Shawe-Taylor. 2002. The set covering machine. *Journal of Machine Learning Reasearch*, 3:723–746.

M. Marchand and M. Sokolova. 2005. Learning with decision lists of data-dependent features. *Journal of Machine Learning Reasearch*, 6:427–451.

M. Marcus, B. Santorini, and M. A. Marcinkiewicz. 1993. Building a large annotated corpus of english: The penn treebank. *Computational Linguistics*, 19(2):313–330.

H. E. Marriot. 1995. Intercultural business negotiation. In A. Firth, editor, *The Discourse of Negotiation*, pages 247–268. Pergamon.

S. Martin, A. Sewani, B. Nelson, K. Chen, and A. Joseph. 2005. Analyzing behaviorial features for email classification. In *Proc of 2nd Conference on Email and Anti-Spam (CEAS 2005).* http://www.ceas.cc/papers-2005/123.pdf.

M. Mast, H. Niemann, E. Noth, and E. G. Schukat-Talamazzini. 1996. Automatic classification of dialog acts with semantic classification trees and polygrams. *Learning for Natural Language Processing*, 1040:217–229.

J. Matthews. 1995. Syntax, semantics, pragmatics. In F. R. Palmer, editor, *Grammar and meaning: Essays in honour of Sir John Lyons*, pages 116–135. Cambridge University Press.

A. McCallum, A. Corrada-Emmanuel, and X. Wang. 2005. Topic and role discovery in social networks. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI 2005)*, pages 786–791. Edinburgh, Scotland.

T. McEnery and A. Wilson. 2001. *Corpus Linguistics.* Edinburg University Press.

J. Miller. 1995. Does spoken language have sentences? In F. R. Palmer, editor, *Grammar and meaning: Essays in honour of Sir John Lyons*, pages 116–135. Cambridge University Press.

T. M. Mitchell. 1997. *Machine Learning.* McGraw-Hill.

A. Mitra, J. Willyard, C. A. Platt, and M. Parsons. 2005. Exploring web usage and selection criteria among male and female students. *Journal of Computer-Mediated Communication*, 10(3). http://www.ascusc.org/jcmc/vol10/issue3/mitra.html.

M. Morris, J. Nadler, T. Kurtzberg, and L. Thompson. 2002. Schmooze or lose: Social friction and lubrication in e-mail negotiations. *Group Dynamics*, 6:89–100.

J. Mulholland. 1991. *The Language of Negotiation: A Handbook of Practical Strategies for Improving Communication.* Routledge.

T. Mullen and N. Collier. 2004. Sentiment analysis using support vector machines with diverse information sources. In *Proc EMNLP*, pages 412–418.

H. Murakoshi, A. Shimaru, and K. Ochimizu. 2000. Construction of deliberation structure in email communication. *Computational Intelligence*, 16(4):570–577.

D. E. Murray. 2000. Protean communication: The language of computer-mediated communication. *TESOL Quarterly*, 34(3):397–421.

A. Namatame, T. Terano, and K. Kurumatani, editors. 2002. *Agent-Based Approaches in Economic and Social Complex Systems.* IOS Press.

V. Nastase and G. Kersten. 2005. Concession curve for inspire data. In *Proceedings of Workshop on Informal and Formal Information Exchange during Negotiations (FINEXIN 2005)*, pages 97–105.

V. Nastase and S. Szpakowicz. 2003. Augmenting wordnet's structure using ldoce. In *Proceedings of CICLing'03*, pages 281–294.

T. Nasukawa and J. Yi. 2003. Sentiment analysis: Capturing favorability using natural language processing. In *Proc ICKC*, pages 70–77.

K. Nigam and M. Hurst. 2004. Towards a robust metric of opinion. In *Proc AAAI-EAAA 2004.* http://www.kamalnigam.com/papers/metric-EAAT04.pdf.

S. Nirenburg and V. Raskin. 2004. *Ontological Semantics.* MIT Press.

M. P. Oakes. 1998. *Statistics for Corpus Linguistics.* Edinburg University Press.

D. W. Opitz. 1995. An anytime approach to connectionist theory refinement: Refining the topologies of knowledge-based neural networks. Ph.D. Dissertation, Department of Computer Science, University of Wisconsin-Madison.

B. Pang, L. Lee, and S. Vaithyanathan. 2002. Thumbs up? sentiment classification using machine learning techniques. In *Proc EMNLP'2002, Philadelphia*, pages 79–86.

S. Park, E. Durfee, and W. Birmingham. 2004. Use of markov chains to design an agent bidding strategy for continuous double agents. *Journal of Artificial Intelligence Research*, 22:175–214.

D. B. Paul and J. M. Baker. 1992. The design for the wall street journal-based csr corpus. In *Proc ICSLP'92, Kobe*, pages 357–361.

J. Pavlidis, J. Weston, and W. Grundy. 2001. Gene functional classification from heterogeneous data. In *Proc of the Fifth International Conference on Computational Molecular Biology*, pages 249–255.

M. R. Perkins. 1983. *Modal Expressions in English.* Ablex Publishing Corporation.

R. M. Perloff. 2003. *The Dynamics of Persuasion.* Lawrence Erlbaum Associates, second edition.

A. Popescu-Belis, M. Georgescul, A. Clark, and S. Armstrong. 2004. Building and using a corpus of shallow dialogue annotated meetings. In *Proc LREC'2004 (Fourth International Conference on Language Resources and Evaluation)*, volume IV, pages 1451–1454.

A. Popescu-Belis. 2005. Dialogue acts: One or more dimensions? Technical Report ISSCO Working Paper 62, School of Translation and Interpretation, University of Geneva. http://www.issco.unige.ch/publications/working-papers/papers/apb-issco-wp62.pdf.

B. Preiser. 1986. *Linguistic Sex Roles in Conversation.* Mouton de Gruyter.

J. Pujolar. 2000. *Gender, heteroglossia and power. A sociolinguistic study of youth culture.* Mouton de Gruyter.

L. Putnam and M. Roloff, editors. 1992. *Communication and Negotiation.* London: Sage.

J. R. Quinlan. 1986. Induction of decision trees. *Machine Learning*, 1(1):81106.

J. R. Quinlan. 1993. *C4.5: Programs for Machine Learning.* Morgan Kaufmann Publishers, San Mateo, California.

H. Raiffa. 1982. *The Art and Science of Negotiation.* Cambridge, MA: Belknap.

P. Rayson and R. Garside. 2000. Comparing corpora using frequency profiling. In *Proc Workshop on Comparing Corpora (ACL 2000)*, pages 1–6.

J. Rennie, L. Shih, J. Teevan, and D. Karger. 2003. Tackling the poor assumptions of naive bayes text classifiers. In *Proc of the 20th International Conference on Machine Learning (ICML-2003)*, pages 616–623. Morgan Kaufmann.

E. Riloff, J. Wiebe, and W. Phillips. 2005. Exploiting subjectivity classification to improve information extraction. In *Proc AAAI-2005*, pages 1106–1111.

E. Riloff. 1996. Using learned extraction patterns for text classification. In S. Wermten, E. Riloff, and G. Schieler, editors, *Connectionist, Statistical, and Symbolic Approaches to Learning for Natural language Processing*, pages 275–289. Springer-Verlag, Berlin.

J. Rojot. 1991. *Negotiation: from theory to practice.* Macmillan Academic and Professional.

R. Rosenfeld. 2000. Two decades of statistical language modeling: Where do we go from there? In *Proc IEEE*, volume 88, pages 1270–1278.

J. Z. Rubin. 2002. The actors in negotiation. In V. Kremenyuk, editor, *International Negotiation: Analysis, Approaches, Issues*, pages 97–109. Jossey-Bass, second edition.

J. Rudanko. 1989. *Complementation and Case Grammar.* State University of New York Press.

J. Rusiecki. 1985. *Adjectives and Comparison in English.* Longman.

M. Sahlgren. 2002. A study of two different machine learning approaches to dialogue management. http://www.ida.liu.se/ nlplab/gslt/papers/.

J. W. Salacuse. 1998. Ten ways that culture affects negotiation style: Some survey results. *Negotiation Journal*, 14(3):221–240.

T. Scheer. 2004. Email answering assistance by semi-supervised text classification. *Intelligent Data Analysis*, 8(5). http://www.informatik.hu-berlin.de/ scheffer/publications/idajournal2004.pdf.

M. Schoop. 2003. A language-action approach to electronic negotiations. In *Proc 8th International Working Conference on the Language-Action Perspective on Communication Modelling (LAP 2003)*, pages 143–160.

L. C. Schourup. 1985. *Common Discourse Particles in English Conversation.* Garland Publishing Inc., New York.

F. Sebastiani. 2002. Machine learning in automate text categoriazation. *ACM Computing Surveys*, 34(1):1–47.

M. Shah, M. Sokolova, and S. Szpakowicz. 2004a. The role of domain-specific knowledge in classifying the language of e-negotiations. In *Natural Language Processing (Proceedings of ICON'2004)*, pages 99–108.

M. Shah, M. Sokolova, and S. Szpakowicz. 2004b. Using domain-specific knowledge to classify e-negotiations. http://interneg.org/interneg/research/papers/.

M. Shaw and D. Garland. 1996. *Software Architecture: Perspectives on an Emerging Discipline.* Prentice Hall.

J. Shawe-Taylor and N. Cristianini. 2004. *Kernel Methods for Pattern Analysis.* Cambridge University Press.

H. Sichel. 1986. Word frequency disributions and type-token characteristics. *Mathematical Scientist*, 11:45–72.

2002. Simple negotiation system. http://interneg.org/enegotiation/projects/.

T. Simons. 1993. Speech patterns and the concept of utility in cognitive maps: the case of integrative bargaining. *Academy of Management Journal*, 38(1):139–156.

G. Sjostedt, editor. 2003. *Professional Cultures in International Negotiations.* Lexington Books.

D. Sleator and D. Temperley. 1991. Parsing english with a link grammar. Technical Report Computer Science Technical Report CMU-CS-91-196, Carnegie Mellon University. http://bobo.link.cs.cmu.edu/link/papers/index.html.

M. Smith. 2000. International business negotiations: a comparison of the theory with the perceived reality of australian practitioners. http://marketing.byu.edu/ams/maxsmith.htm.

M. Sokolova and S. Szpakowicz. 2004. Language in electronic business negotiations. In *Proceedings of InterNeg International Seminar "Markets, negotiations and dispute resolution in the new economy"*, pages –.

M. Sokolova and S. Szpakowicz. 2005. Analysis and classification of strategies in electronic negotiations. In *Advances in Artificial Intelligence (Proc 18th Conference of the Canadian Society for Computational Studies of Intelligence)*, volume 18, pages 145–157. Springer-Verlag.

M. Sokolova, M. Marchand, N. Japkowicz, and J. Shawe-Taylor. 2003. The decision list machine. In *Advances in Neural Information Processing Systems (Proc of NIPS'2002)*, volume 15, pages 921–928. The MIT Press.

M. Sokolova, V. Nastase, and S. Szpakowicz. 2004a. Language in electronic negotiations: Patterns in completed and uncompleted negotiations. In *Natural Language Processing (Proceedings of ICON'2004)*, pages 142–151.

M. Sokolova, S. Szpakowicz, and V. Nastase. 2004b. Automatically building a lexicon from raw noisy data in a closed domain. http://interneg.org/interneg/research/papers/.

M. Sokolova, S. Szpakowicz, and V. Nastase. 2004c. Using language to determine success in negotiations: A preliminary study. In *Advances in Artificial Intelligence (Proc 17th Conference of the Canadian Society for Computational Studies of Intelligence)*, pages 449–453. Springer-Verlag.

M. Sokolova, V. Nastase, M. Shah, and S. Szpakowicz. 2005a. Feature selection for electronic negotiation texts. In *Proceedings of RANLP'2005*, page to appear. -.

M. Sokolova, V. Nastase, S. Szpakowicz, and M. Shah. 2005b. Analysis and models of language in electronic negotiations. In M. Draminski, P. Grzegorzewski, K. Trojanowski, and S. Zadrozny, editors, *Issues in Intelligent Information Systems. Models and Techniques*, page in press. EXIT.

M. Sokolova, M. Shah, and S. Szpakowicz. 2005c. Comparative analysis of text data in successful face-to-face and electronic negotiations. In *Proceedings of Workshop on Informal and Formal Information Exchange during Negotiations (FINEXIN 2005)*, pages 33–42.

M. Sokolova, M. Shah, and S. Szpakowicz. 2006. Comparative analysis of text data in successful face-to-face and electronic negotiations. *Group Decision and Negotiations*, 15(2):22 pages.

R. A. Solsona, E. Fosler-Lussier, H.-K. J. Kuo, A. Potamianos, and I. Zitouni. 2002. Adaptive language models for spoken dialogue systems. In *Proc of International Conference on Acoustics, Speech, and Signal Processing.* http://www.cse.ohio-state.edu/ fosler/papers/icassp2002-adaptlm.pdf.

D. J. Stracuzzi and P. E. Utgoff. 2004. Randomized variable elimination. *Journal of Machine Learning Research*, 5:13311362.

M. Ströbel. 2000. Effects of electronic markets on negotiation process. http://www.ipsi.fraunhofer.de/m̃ahleko/seminars/ part%208-electronic%20markets/strobel/.

D. Summers, editor. 2003. *Longman Dictionary of Contemporary English.* Pearson Education: Longman, fourth edition.

M. Sutter, R. Bosman, M. Kocher, and F. van Winden. 2003. Experimental evidence of the importance of gender pairing in bargaining. http://www1.fee.uva.nl/creed/pdffiles/PTTgenderpaper20FINAL.pdf.

J. Suzuki, H. Isozaki, and E. Maeda. 2004. Convolution kernels with feature selection for natural language processing tasks. In *Proceedings of ACL-2004*, pages 119–126.

L. Thomos, H.P. Pfister, and P. Peterson. 2004. Issues related to the construction of a purpose-built domain-specific word corpus. *Australian Journal of Educational and Developmental Psychology*, 4:13–28.

C. A. Thompson and R. J. Mooney. 2003. Acquiring word-meaning mappings for natural language interfaces. *Journal of Artificial Intelligence Research*, 18:1–44.

L. Thompson and J. Nadler. 2002. Negotiating via information technology: Theory and application. *Journal of Social Issues*, 58(1):109–124.

M. Tomokiyo and G. Chollet. 2005. Voiceunl: A semantic representation of emotions within universal networking language formalism based on a dialogue corpus analysis. In *Proceedings of CICLing*, pages 441–451.

G. Tottie. 1991. *Negation in English Speech and Writing.* Academic Press Inc.

J. Tsujii et al. 2005. Genia. http://www-tsujii.is.s.u-tokyo.ac.jp/.

F. J. Tweedie and R. H. Baayen. 1998. How variable may a constant be? measures of lexical richness in perspective. *Computers and the Humanities*, 32:323–352.

V. Vapnik. 1998. *Statistical Learning Theory.* Wiley, New-York, NY.

F. Vega-Redondo. 2003. *Economics and the Theory of Games.* Cambridge University Press.

R. Vetschera, G. Kersten, and S. Köszegi. 2004. The determinants of nss success: An integrated model and its evaluation. http://interneg.org/interneg/research/papers/.

C. Walcott and P. T. Hormann. 1978. Interaction analysis and bargaining behavior. In Golembiewski, editor, *The Small Group in Political Science*, pages 251–261. The University of Georgia Press.

J. B. Walther. 1992. Interpersonal effects in computer-mediated interaction: A relational perspective. *Communication Research*, 19(1):52–90.

S. Wang, D. Schuurmans, and F. Peng. 2003. Latent maximum entropy approach for semantic n-gram language modeling. In *Proc AISTATS.*

B. Warren. 1984. *Classifying Adjectives.* Acta Universitatis Gothoburgensis.

C. Watson. 1994. Gender versus power as a predictor of negotiation behavior and outcomes. *Negotiation Journal*, 10:117–127.

I. Witten and E. Frank. 2000. *Data Mining.* Morgan Kaufmann. http://www.cs.waikato.ac.nz/ml/weka/.

M. Wolska and I. Kruijff-Korbayov. 2004. Analysis of mixed natural and symbolic input in mathematical dialogs. In *Proceedings of ACL'04*, pages 25–32.

A. Yakushiji, Y. Tateisi, Y. Miyao, and J. Tsujii. 2004. Finding anchor verbs for biomedical ie using predicate-argument structures. In *Companion Volume to the Proceedings of ACL'04*, pages 157–160.

R. Yangaber and R. Grishman. 2000. Extraction pattern discovery through corpus analysis. In *Proc LREC*, pages 31–35.

J. A. Yates and W. J. Orlikowski. 1993. Knee-jerk anti-loopism and other e-mail phenomena: Oral, written, and electronic patterns in computer-mediated communication. Technical Report Center for Coordination Science Technical Report 150, MIT. http://ccs.mit.edu/papers/CCSWP150.html.

S. J. Yates. 1997. Gender, identity and cmc. *Journal of Computer Assisted Learning*, 13(4):281–290.

K. Zechner. 2002. Automatic summarization of open-domain multiparty dialogues in diverse genres. *Computational Linguistics*, 28(4):447485.

J. Zeleznikow. 2002. Risk, negotiation and argumentation – a decision support system based approach. *Law, Probability and Risk*, 1:37–48.

R. Zhang and A.I. Rudnicky. 2002. Improve latent semantic analysis based language model by integrating multiple level knowledge. In *Proc. of ICSLP*, pages 893–896.

Y. Zhang, A. Nur Zincir-Heywood, and E. E. Milios. 2004. Term-based clustering and summarization of web page collections. In *Proc 17th Conference of the Canadian Society for Computational Studies of Intelligence*, pages 60–74. Springer-Verlag.

Index

Appendix A

Semantic and behavioural categories and language groups

Table 1: LDOCE3 NLP Database Subject Codes

ARTS (A)

Architecture (AA) Literary (AL) Media (AM): Film (AMF) Television & Radio (AMT) Performing (AP): Dance (APD) Music (APM) Theatre (APT) Visual (AV): Design (AVD) Painting (AVP) Sculpture (AVS)

BUSINESS (B)

Basics (BB): Advertising (BBA) Companies (BBC) Business Management (BBB) Marketing (REM) Offices (BBO) Trade (RET) Employment (BE): Conditions of Employment (BEC) Labour Relations/Unions BEL) Wages (BEW) Finance (BF): Banking (BFB) Loans (BFL) Stocks & Shares (BFS) Insurance (BFI) Occupations (BO)

COLOURS/SOUNDS/ETC (C)

Colours (CC) Forms/shapes (CF) Odours (CO) Surface/Texture (CS) Tastes (CT)

DAILY LIFE (D)

Clothes & Fashion (DC): Beauty (DCB) Clothes (DCC) Jewellery (DCJ) Food(DF): Cooking (DFC) Drink (DFD) Food Dish (DFF) Nutrition (DFN) Tobacco (DFT) Utensils (DFU) Games (DG): Board Games (DOB) Cards (DGC) Darts (DOD) Gambling (DOG) Other Games (DGO) Snooker (DOS) Household (DH): Babies (DHB) Cleaning (DHC) Furniture (DHF) House (DHH) Pets(DHP)

Toys(DHT) Leisure (DL): Antiques (DLA) Clubs (DLC) Gardening (DLG) Hobbies (DLH) Outdoor (DLO) Tourism (DLT) Nature (DN) Sport (DS): American Football (DSA) Baseball (DSB) Cricket (DSC) Football (DSF) Golf (DSG) Horses (DSH) Other Sports (DSO) Swimming (DSS) Tennis (DST)

HARD SCIENCE (H)

Astronomy (HA) Biology (HB): Animals (MBA) Birds (HBB) Fish(HBF) Human (HBH) Insects (HBI) Microbes/Genetics/Biochem(HBM) Plants (HBP) Chemistry (HC): Compounds (HCC) Elements (HCE) Metals (HCM) Plastics (HCP) Earth Sciences (HE): Geology (MEG) Meteorology (HEG) Oceanography (HEO) Maths (HM): Numbers (HMN) Statistics (HMS) Physics (HP): Electricity (HPE) Optics (HPO)

MEDICINE (M)

Alternative Medicine (USA) Birth (MB) Death (MX) Drugs/Medicines (MD): Drug Culture (MDD) Hospital (MH) Illness & Disability (Ml) Nurses, Doctors, etc (MN) Psychology & Psychiatry (MP)

POLITICS/WORLD AFFAIRS (P)

Economics (PE): Currencies (PEC) Grants (PEG) Tax(PET) Welfare (PEW) Government (PG): Citizenship (PGC) Parliaments (POP) Officials (PGO) Military (PM): Army(PMA) Defence (PMD) Navy(PMN) Planes (airforce) (PMP) Weapons (PMW) Politics (PP): Groupings (PPG) Policies (PPP) Voting (PPV) Third World (PT)

RELIGION & THOUGHT (R)

Folklore (RF) Mythology (RM) Occult (RO): Astrology (ROA) Magic (ROM) Philosophy (RP) Religion (RR): Buddhism (RRB) Christianity (RRC) Hinduism (RRH) Islam (RRI) Judaism (RRJ)

SOCIETY/SOFT-SCIENCE (S)

Anthropology (SA): Nationalities (SAN) Archaeology (SX) Crime & Law (SC): Bombs & Terrorism (SCB) Crime (SCC) Jail(SCJ) Law(SCL) Police (SCP) Trials (SCT) Education (SE): College (SEC) Pre School (SEP) School (SES) Training (SET) Geography (SG): Ecology (SGE) Pollution (SOP) History (SH) Linguistics (SL): Alphabets (SLA) Grammar (SLG) Languages (SLL) Sex(SY) Sociology (SS): Children (SSC) Family (SSF) Organisations (SSO) Race Relations (SSR) Women (SSW) Youth (SSY)

TECHNOLOGY (T)

Agriculture (TA): Biotechnology (TAB) Crops (TAC) Forestry (TAF) Soil(TAS) Building (TB): Construction (TBC) Buildings (TBB) Communications (TC): Broadcasting (TCB) Mail (TCM) Newspapers/Printing/ Publishing (TCN) Photography (TCP) Recording (TCR) Telephone Telegraph (TCT) Data Processing & Computing (TD) Engineering (TE): Civil (TEC) Electrical (TEE) Mechanical (TEM) Industry: (TI) Crafts (TIC) Factories (TIF) Glass (TIG) Material & Textiles (TIM) Pottery & Ceramics (TIP) Measurement (TM): Chronology (TMC) Temperature (TMT) Power (TP): Electricity (TPE) Gas Coal Oil (TPG) Wind Water Sun (TPW) Tools (TZ) Transport (TT): Air (TTA) Bicycles & Cans (TTB) Cars (TTC) Roads (TTR) Space (TTS) Trains & Railways (TTT) Water (TTW)

Categories	Actions
Substantive behaviour	initiations, accommodations, retractions
Strategic behaviour	commitments, threats, promises
Task behaviour	agreements, disagreements, questions, answers
Affective behaviour	positive affect, negative affect
Procedural behaviour	subject change

Table A.1: BPA behavioral categories and actions.

Group	Subgroups
Practical arrangements and ceremonial manners	Suggesting a negotiation, making appointments, greetings, manners declaring good intentions, announcements, speeches, meeting, requesting, volunteering help or services, apologizing, thanking, withdrawing from a negotiation
Bargaining	stating a position, holding to a position, offering and proposing, accepting an offer, refusing an offer, persuading, promising and assuring, being conciliatory, sympathizing with the other side delaying, showing patience, showing impatience, complaining, warning, arbitraining
Discussing	stimulating and directing a discussion, showing attention, interrupting, concentrating on a subject, calming a discussion, speaking briefly, speaking frankly, speaking in secret, clarifying and explaining, asking for something to be repeated, interviewing, introducing an opinion, agreeing, disagreeing, showing skepticism, calling for critisism, stating the view of group, stating proof
Deciding	saying yes, saying no, saying you don't know, avoiding a definite answer, expressing indecision, changing your mind
Frequent subjects in a negotiation	assumptions, priorities, conditions, interests, permission, precedent, advice, support, authority and responsibility, rights and accusations
Technical points	money, figures, estimates about figures, details, a document, charts
Commenting on negotiations	describing the other side, expressing suspicion of the other side, comparing advantages with the side, stating the progress or failure of negotiations, commenting on an agreement

Table A.2: Groups and subgroups of language examples in negotiations, by O. Hansberger.

Appendix B

Indicative words in Inspire and Cartoon

The indicative words for buyer and seller *Inspire* data (non-directional case)

itex manufacturing cycles cypress bikes suppliers pay supplier parts our sprocket $4 purchase 12 4 37 department below bike low reduce lower mountain products expensive equipment policy finance 71 return 75 10 deliver meet costs bicycle market 45 raw your as down president cover product why need line favorable reduced spoilage lowest the special delivery buy 47 am lowered assure can terms best gears payment i day firm ru new still industry from other 30 important full anyway clients refund too soon thank 12$ high office meeting didn't cycle further another pleased kong de just 98 small me work with degree i'm staff shipment sprockets well afford 37$ interest gear respond returned much you paid supplying hong chemnitz taiwan and drop guess there satisfaction im ten was us assemblies u win step fair upon present canada don't sales components concessions i?m companies working large reduction days price shipping begin ottawa sincerly discussions whole since additional back like we're 50 cooperation offering material within if messages kindly concerned issue guarantee promise paying com it conditions decrease side pls 5 should welcome possibility 98$ big profit mail negociation start expect willing then seem seems inspire get match mike cut sorry give decided worth wondering loss thing confident quality concede value satisfy ok them such project competitive offers excellent finally heard service discount more only who looking having provide introduce ourselves would met therefore condition wrong manager under increase hallo suggested negotiating financial favourable opinion time firms standard customer support hear extra going quite let germany contact producing prepared requested what flexible discussed you're ratings system for both company think nice contract excuse that's at put orders require hybrid convinced complete understand rate not may percent maybe sending hesitate help shorter worked table main years improve its set faster review actually every required 21 believe totally be probably hi tried profitable slightly concerning expectations any represent confidence bring administration country suggestion four benefit which submit mind it?s remember care allowed i'll regards bit $3 mutual middle earlier matter reconsider basis representative speak regarding end suggest higher items final suit please but into difference 20 one difficult satisfactory future in are hearing control little decide although minimum ensure good without always mr school second might yours above or opportunity great following grant due demand average don?t will using lets apologize technical once study how evaluate sell receiving response few must options change e otherwise free fact left deal see made solution materials were realize relation thoughts that continue agreeable kind better being doing plan internet allow become limit 0 order isn't closer concession else problems customers 71$ been possible supply accept m mine keeping details trade five write late dear discussion find around yourself personal proposed on sure payments discuss open night long partner hence yesterday unacceptable almost go hey answer rather 3 class agreed sign negotiation live number satisfied highest requirements getting wanted discussing message my cannot respect feel revised choice say according course interesting running major can?t management won't world hopefully

The indicative words for buyer and seller *Inspire* data

(directional case - sellers vs. buyers)

itex manufacturing our $4 department below low 12 reduce 37 deliver return 4 lower meet policy terms products favorable 10% can delivery lowest gears costs 45 payment 75% spoilage down industry refund day negociation further with possibility 98 30 shipping special best provide sincerly confident step and components customer i'm working meeting upon opportunity within taiwan pleased 75 work days time only interest chemnitz service end price guarantee sales let i?m such late inspire set manager suggested faster 12$ maybe sorry offering companies opinion excellent fair standard understand under shorter 5% 20 it?s contact present discuss require solution probably basis give producing without both but submit items satisfy favourable met people allow point apologize value free percent cut win concessions order study condition system quality yesterday representative following put go game next at counter decrease client agreement 98$ financial ensure moment kindly germany made improve concession any may attractive nice interesting contract trade benefit change review profitable can't difficult week regards expectations is alex decide discussed receiving world request acceptance busy orders profit deadline 10 keeping suggest wishes we've respect else certainly find conditions early to own

The indicative words for buyer and seller *Inspire* data (directional case - buyers vs. sellers)

cypress cycles parts pay bikes suppliers supplier purchase need bike bicycle other 71 product new high too i your full market why u gear defective line still am mountain how manufacturer another didn't ok sooner you're should buy hi increase feel client side acceptance earlier 47 since $3 willing they there were as cycle competitors yet rating concerned reputation company help view thank assemblies fast back me deadline canada each demand future cash afford thanks respond paying so much school which reply higher returned done sprockets flexible messages part does whether negotiating boss their opportunity yourself main here guess actually reach expect moment hand suitable carefully wait class reviewed second better us build otherwise satisfactory final two for touch mine consideration keep people like to start decided think parties least seem having counter several also we're them question all doing quickly especially heard not afraid today program closer hybrid

The indicative words for buyer and seller *Cartoon* data (non-directional case)

skip paying you got my we've certainly sell might 4 also we relative pay within profit think help type again folks contract talk our talking 8 in program payment let's look you've into take purchase interest we're understand buying success we'll can children offer me 5 buy 7 obviously tie different accept don point production else information definitely position negotiate terms 208 with ins ranger deal 103 feel okay very syndication cash revenue mean between had have really alright assuming side i difference right 202 percentage thirty able 15 eight part arrangement here ahead package ultra possibly called from another now how nothing beneficial is rather available saying there nine has because additional um that i would it's bit group people it finance out see age sorry percent cannot may TRUE leave 11 calculate add be already i've third 303 million yes correct up 111 55 properties competition 106 yours fairly attractive been wanted 80 kids 9 split financing wwin you're certain to aware kind months means there's total place thinking level opportunity product goes 6 zero interesting schedule meet fact are can't series base range sounds taking about situation could was sharing looks equal almost we 0 will many oh particular last what doing shows management 201 901 both payments i'd dollars so broad rest focus feeling 65 such either starting bet doesn't us possible less looking given 25% sure lot instead competitive increase seems during what's 50% depending asking big

Appendix C

Distribution of strategic words in buyer and seller data

Pattern	*Inspire*				*Cartoon*				*SimpleNS*	
	BS	SS	BU	SU	BJ	SJ	BUS	SUS	B	S
	544961	525049	209025	205524	8053	8757	56753	63225	51253	47493
must/will/have to										
I/we ...	1897	1749	760	700	25	8	22	42	233	179
you ...	821	852	390	372	5	15	16	22	71	80
can/may										
I/we ...	2281	2214	965	1028	49	24	203	278	188	203
you ...	861	473	401	361	-	1	4	5	87	61
could/would/should										
I/we ...	2126	1684	788	765	38	45	274	279	167	177
you ...	559	488	243	234	4	7	29	59	42	38

Table C.1: Occurrences of the subgroups of *PresPron Modal* in buyer and seller data.

Mental Verb	*Inspire*				*Cartoon*				*SimpleNS*	
	BS	SS	BU	SU	BJ	SJ	BUS	SUS	B	S
	544961	525049	209025	205524	8053	8757	56753	63225	51253	47493
think	1612	1475	562	492	26	38	238	405	149	180
know	1053	1041	503	454	46	29	408	516	112	122
understand	496	525	184	192	12	3	86	66	70	67
consider	487	469	202	195	7	-	8	14	37	32
guess	116	73	41	23	3	3	59	58	7	7

Table C.2: Occurrences of the mental verbs in buyer and seller data.

Volition Verb	*Inspire*				*Cartoon*				*SimpleNS*	
	BS	SS	BU	SU	BJ	SJ	BUS	SUS	B	S
	544961	525049	209025	205524	8053	8757	56753	63225	51253	47493
hope	2453	2362	895	931	-	2	12	10	178	159
want	751	715	281	277	75	73	148	216	73	98
wish	263	255	107	88	2	1	2	2	21	18
like	1288	1125	493	451	45	67	174	199	129	114
prefer	102	104	30	32	-	-	6	15	12	18
agree	1819	1832	670	626	16	16	-	-	140	159
promise	344	361	97	120	-	1	-	2	8	2
ask	344	340	134	111	12	17	51	34	48	30
afford	104	50	53	27	1	6	2	9	4	-
aim	53	34	23	13	-	-	-	-	1	-
choose	44	50	20	22	1	1	-	-	9	8
decide	253	211	81	75	2	6	8	17	16	-
intend	33	17	11	8	-	-	1	1	2	1
look	1397	1321	535	523	-	-	87	296	29	19
plan	225	122	84	52	4	10	10	13	10	3
propose	229	226	86	94	-	-	6	7	32	45
make	1660	1642	679	666	31	24	125	128	158	187
manage	220	250	91	87	1	-	2	1	3	5
move	126	115	54	52	-	-	24	19	13	13
proceed	40	39	16	21	1	-	-	1	1	-
try	539	550	209	237	7	3	99	104	30	30
decline	12	13	4	6	-	-	-	1	-	-
refuse	7	8	10	10	-	-	-	-	-	-
reject	18	34	11	10	-	-	-	-	-	1
disagree	15	14	9	10	-	-	2	2	-	-
delay	181	149	89	78	-	-	2	2	11	9
hesitate	31	47	10	14	-	-	-	-	5	4

Table C.3: Occurrences of volition verbs and their wordforms in buyer and seller data.

BE-negatives	*Inspire*				*Cartoon*				*SimpleNS*	
	BS	SS	BU	SU	BJ	SJ	BUS	SUS	B	S
am not/n't	5.2	4.7	5.5	4.2	1.2	1.4	5.2	2.7	2.9	1
are not/n't	7.7	7.2	8.6	7.4	1.2	-	0.8	0.1	6.4	4.7
is not/n't	13.7	7.9	14.4	30.8	1.2	1.4	1.1	1.3	16.5	13.5
will not/n't	7.5	7	7.6	8.8	4.8	-	1.6	2.1	8.5	9.6

Table C.4: Percentage covered by the subgroups of BE-negatives in buyer and seller data.

Not-negatives	*Inspire*				*Cartoon*				*SimpleNS*	
	BS	SS	BU	SU	BJ	SJ	BUS	SUS	B	S
I/we cannot/can't	19.3	22.2	19.3	23	20.5	11.1	4.3	1.9	11.7	12
I/we do not/don't	11.6	10.34	12	11	3.6	19.4	11.6	11.6	7.4	8.6
I am not/I'm not/ we are not/we aren't	7.7	7.4	8.5	7.3	2.4	4.1	5.5	2.8	4.5	2.7
I/we have not/haven't	4.7	3.6	3.7	3.5	-	2.7	1.8	1.1	1.4	0.7

Table C.5: Percentage covered by the not-negatives with first-person in buyer and seller data.

Appendix D

Features of higher linguistic order

Dependency	**Word pairs**
ADJMOD-POST	(acceptable,you),(important,me), (important,us), (come,agreement),(back,me), (back,you),(forward,response),(important, you),(forward,reply), (acceptable,us)
NMOD-POST	(you,offer),(payment,delivery), (you,very), (price,3), (thank,offer), (refund,spoilage),(days,delivery),(thanks,offer),(offer,you)
NMOD-PRE	(offer,last), (offer,new), (delivery,days), (time,delivery),(*hope*,i), (delivery,after),(policy,return),(assemblies,gear),(price,full),(assemblies,wheel)
NMOD-RELATIV	(me,think),(offer,is),(*hope*,accept), (*hope*,will),(*hope*,be), (price,is),(agreement,is),(hope,can), (feel,is),(offer,think)
NMOD-DEDUCED	(relationship,long),(terms,returns),(look,let),(supplier,assemblies), (you,you),(offer,returns),(offer,payment),(offer,hope), (offer,offer), (offer,reasonable)

Table D.1: Ten most frequent pairs with negotiation-related words, BSN.

Dependency	Word pairs
ADJMOD-POST	(acceptable,you), (come,agreement), (back,you),(important,us),(important,me) (forward,response),(forward,reply),(back,me), (acceptable,us), (acceptable,me)
NMOD-POST	(you,offer),(payment,delivery), (you,very), (price,3), (refund,spoilage), (days,delivery),(thanks,offer),(offer,you), (delivery,days)
NMOD-PRE	(offer,last), (offer,new), (delivery,days), (time,delivery),(delivery,after), (assemblies,gear),(offer,counter), (*hope*,i),(line,new),(price,full)
NMOD-RELATIV	(hope,accept), (offer,is), (hope,will),(me,think),(offer,propose), (offer,make),(*hope*,be), (suppliers,can),(offer,think),(*hope*,is)
NMOD-DEDUCED	(relationship,long), (offer,final),(terms,payment),(supplier,assemblies), (relationship,prosperous),(offer,best),(offer,last),(look,let), (you,offer), (product,qualitative)

Table D.2: Ten most frequent pairs with negotiation-related words, BUN.

Dependency	Word pairs
ADJMOD-POST	(acceptable,you), (come,agreement), (important,you),(forward,reply), (forward,response),(important,me),(back,me), (acceptable,us), (important,us), (close,agreement)
NMOD-POST	(you,offer),(payment,delivery), (refund,spoilage), (price,3), (time,days), (offer,you),(thanks,offer),(days,delivery), (you,very),(thank,offer)
NMOD-PRE	(offer,new), (offer,last),(time,delivery),(policy,return), (delivery,days), (*hope*,i),(delivery,after),(offer,counter),(terms,payment), (quality,high)
NMOD-RELATIV	(offer,is),(*hope*,be), (*hope*,will), (*hope*,accept), (price,offer), (*hope*,is), (me,think), (agreement,is),(price,is), (days,is), *(delivery,is)*
NMOD-DEDUCED	(relationship,long),(degree,finish),(business,long),(terms,payment), (payment,payment), (time,time), (you,offer),(terms,returns), (time,payment),(relationship,Itex)

Table D.3: Ten most frequent pairs with negotiation-related words, SSN.

Dependency	Word pairs
ADJMOD-POST	(acceptable,you), (come,agreement), (important,you),(forward,reply), (important,us),(back,me),(back,you),(forward,response), (important,me),(good,you)
NMOD-POST	(payment,delivery), (you,offer), (refund,spoilage),(offer,you), (price,3), (time,days),(price,4),(days,delivery),(you,very),(delivery,days)
NMOD-PRE	(offer,last), (offer,new), (time,delivery),(policy,return),(delivery,days), (*hope*,i),(delivery,after),(offer,counter),(assemblies,gear), (products,quality)
NMOD-RELATIV	(*hope*,will),(us,offering),(*hope*,accept),(*hope*,be), (offer,is), (*hope*,can), (price,offer),(*hope*,is),(offer,think),(offer,find)
NMOD-DEDUCED	(delivery,assemblies), (offer,final),(terms,terms),(terms,payment), (offer,assemblies),(offer,last),(offer,fair),(supply,assemblies), (terms,returns),(price,price)

Table D.4: Ten most frequent pairs with negotiation-related words, SUN.

Dependency	Word pairs
NUCL-PERFECT	(have,been), (have,received), (have,made),(has,been),(have,prepared), (have,agreed),(have,decided),(have come),(have,sent), (have, accepted)
NUCL - MODAL	(will,be),(would,like),(would,be),(can,accept),(will,accept), (should,be), (will,have),(will,find),(cannot,accept),(can,be)
NUCL-INSIST	(don't,know),(do,have),(don't,want),(don't,think),(do,want), (do,think),(don't,have),(did,get),(do,know),(do,understand)
VMOD-NEG	(is,not),(are,not),(accept,not),(be,not),(am,not), (have,not),(willing,not), (received,not), (was,not), (want,not)
VMOD-COMPLTHAT	(think,is),(hope,accept),(hope,find),(think,be),(think,are), (hope,be), (hope,have), (hope,understand), (hope,is), (think,have)
VMOD - UNSAFE	(know,is),(understand,is), (know,are), (understand,are), (understand,accept), (see,are),(know,be),(know,have), (know,looking), (understanding,are)

Table D.5: Ten most frequent pairs with strategy-related words, BSN.

Dependency	Word pairs
NUCL-PERFECT	(have,been), (have,received),(has,been),(have,made), (have,sent), (have,agreed), (have,prepared),(have,heard), (have,reviewed),(have,decided)
NUCL - MODAL	(will,be),(would,like),(would,be),(can,accept),(should,be), (will,have),(cannot,accept),(will,find),(will,accept),(can,be)
NUCL-INSIST	(do,have),(don't,know),(don't,want),(don't,have),(do,know), (don't,think),(do,think),(didn't,get), (do,want), (do,understand)
VMOD-NEG	(is,not),(are,not),(accept,not),(be,not),(am,not), (willing,not),(have,not),(was,not),(make,not),(agree,not)
VMOD-COMPLTHAT	(think,is),(hope,find),(hope,accept),(think,be),(hope,be), (hope,agree),(hope,reach), (hope,understand),(believe,is), (hope,is)
VMOD - UNSAFE	(know,is),(know,are), (understand,have), (know,have), (understand,is), (see,agree), (see,is), (understand,make), (hear,are), (seems,are)

Table D.6: Ten most frequent pairs with strategy-related words, BUN.

Dependency	Word pairs
NUCL-PERFECT	(have,been),(have,made), (has,been),(have,received),(have,decided), (have,agreed),(have,come) (have,sent),(have,reviewed), (have,accepted)
NUCL - MODAL	(will,be),(would,like),(would,be),(can,accept),(will,have), (will,accept),(will,find),(can,be),(cannot,accept),(can,offer)
NUCL-INSIST	(do,have),(don't,know),(don't,want),(do,think), (don't,think), (do,want),(do,know), (don't,have),(did,receive),(do,agree)
VMOD-NEG	(is,not),(are,not),(accept,not),(be,not),(am,not), (have,not),(was,not), (know,not), (go,not), (like,not)
VMOD-COMPLTHAT	(think,is),(hope,accept),(hope,find),(think,be),(hope,be), (think,are), (hope,is), (hope,have),(hope,understand), (hope,reach)
VMOD - UNSAFE	(understand,is), (know,is), (understand,are),(know,need), (understand,have), (know,are),(find,is),(see,are),(know,have), (seems,are)

Table D.7: Ten most frequent pairs with strategy-related words, SSN.

Dependency	Word pairs
NUCL-PERFECT	(have,been), (has,been),(have,received), (have,made),(have,agreed), (have,decided),(have,prepared),(have,sent), (have,given),(have,come),
NUCL - MODAL	(will,be),(would,like),(would,be),(can,accept), (can,do), (will,find),(can,be),(will,have), (can,offer), (cannot,accept)
NUCL-INSIST	(do,have),(don't,know), (don't,have), (don't,think), (don't,want), (do,think), (did,get),(do,want), (do,agree), (do,know),
VMOD-NEG	(is,not),(are,not),(accept,not),(be,not),(am,not), (have,not), (agree,not),(received,not), (was,not), (willing,not)
VMOD-COMPLTHAT	(think,is),(hope,find),(hope,accept),(hope,be),(hope,is), (think,be),(think,are),(hope,agree), (hope,understand),(hope,are)
VMOD - UNSAFE	(know,is),(know,are),(understand,is),(understand,are),(know,have), (considering,is),(understand,have),(regret,accept),(trust,is),(know,interested)

Table D.8: Ten most frequent pairs with strategy-related words, SUN.