Query Expansion on Compounds

Bolette Sandford Pedersen

Center for Sprogteknologi, University of Copenhagen
Njalsgade 80, DK-2300 S
bolette@cst.dk

Abstract

Compounds constitute a specific issue in search, in particular in languages where they are written in one word, as is the case for Danish and the other Scandinavian languages. For such languages, expansion of the query compound into separate lemmas is a way of finding the often frequent alternative synonymous phrases in which the content of a compound can also be expressed. However, it is crucial to note that the number of irrelevant hits is generally very high when using this expansion strategy. The aim of this paper is to examine how we can obtain better search results on split compounds, partly by looking at the internal structure of the original compound, partly by analyzing the context in which the split compound occurs. We perform an NP analysis and introduce a new, linguistically based threshold for retrieved hits. The results obtained by using this strategy demonstrate that compound splitting combined with a shallow linguistic analysis focusing on the recognition of NPs can improve search by bringing down the number of irrelevant hits.

1. Search on Compounds in Danish – a Query Expansion Topic

During the last decades, several researchers and developers have experimented with language technology techniques in order to improve information retrieval, cf. Gonzales et al., 1998 for an experiment with EuroWordNet and Voorhees, 1994, Voorhees, 1994, Smeaton & Quigley, 1996 for experiments concerning the use of Princeton WordNet. In these experiments, synonymy and other semantic relations such as hyponymy have been applied for enhancing queries. Also the TREC reports describe experiments with language technology in information retrieval, and they are generally neutral in their conclusions, stating that natural language processing seem to provide only quite modest results at its current stages (see for instance Strzalkowski et al., 1996, Strzalkowski et al., 1997 and Strzalkowski et al., 1998).

In this paper we apply language technology in order to improve search on Danish *compounds*. Compounds constitute a specific issue in search, in particular in languages where they are written in one word, as is the case for Danish and the other Scandinavian languages. For such languages, expansion of the query compound into separate lemmas is a way of finding the often frequent alternative synonymous phrases in which the content of a compound can also be expressed. In other words, if we search for *apoteksovertagelse* (`farmacyFUGE`[1]`takeover`, 'farmacy takeover'), search results containing *overtagelse af apotek* (`takeover of a farmacy`) are presumably relevant, since the two expressions are synonymous.

Recent experiments for Swedish – a language which has the same characteristics as Danish regarding compounding – show that compound splitting improves search results substantially, mainly due to the fact that many queries containing compounds simply do not produce *any* search results without query expansion, cf. Chen & Gey (2003), and Dalianis (2005). In spite of these improvements in search, it is crucial to note that Dalianis gets 23% non-relevant hits when expanding to split compounds. The same scenario is reported on for Danish by Ankiro - a Danish company who develops language-based search engines for Danish, and who also performs compound splitting in several of their systems in order to improve search results (www.ankiro.dk).

On the basis of the findings concerning Danish and Swedish, we do not in this paper question the fact that compound splitting is a fruitful strategy in search. Query expansion seems to be a *precondition* for satisfactory search results – also in the light of the fact that both languages are becoming heavily influenced by English, meaning that misspelling of compounds (in two words) is becoming more and more frequent in texts. Rather we have asked ourselves the following questions: Does the internal structure of the compound tell us anything relevant about the likelihood of finding relevant search results based on an expanded compound query? And can we by performing a *shallow linguistic analysis* discard a substantial part of the irrelevant hits found by splitting compounds? Although Scandinavian languages (as well as other Germanic languages) are specific in this context, similar phrase normalization problems have been investigated for English and have been treated also by using linguistic analysis, for instance as described in Strzalkowski et al. 1996 and Strzalkowski et al. 1997. They use parsing techniques to locate so-called head+modifier pairs streams as a means to normalize English phrases like *weapon proliferation, proliferation of weapons,* and *proliferate weapons*. Although we limit ourselves to focus on the Danish compound problem, we pursue a similar approach in this work using a shallow noun phrase identification technique combining it, however, with detailed lexical information on predicate-argument structure.

---

[1] FUGE stands for linking element which can be either *-s-*, *-e-* or zero.

Summing up, the aim of this paper is to examine how we can obtain better search results on split compounds, partly by looking at the internal structure of the original compound, partly by analyzing the context in which the split compounds occur. The experiments reported on have been carried out in the Danish research project, VID (VIden og Dokumenthåndtering med Sprogteknologi – Knowledge and Document Handling with Language Technology) in co-operation with five Danish companies, among others the technology provider Ankiro. The project was funded by the Danish Centre for IT research. A full account of the query expansion experiment will be given in Pedersen (forthcoming).

2. Compounds and Synonymous NPs

We have had two assumptions as a basis for our experiments: 1) that some compounds are more likely to have synonymous 'split' counterparts than others; and 2) that search results where both the search words (obtained by splitting a compound noun) occur in the same noun phrase, are more likely to contain a synonymous phrase to the original compound query.

An investigation of the first assumption calls for an analysis of the internal structure of the compound to see for instance if compounds where the head is derived from a verb act in a particular way and are eventually safer candidates for expansion than more lexicalized, non-predicative compounds. From several studies on compound nominalizations such as Levi (1978), Grimshaw (1990), Lapata (2002), and Ørsnes (1995), we know that verb-derived compounds follow certain phrase-like patterns inherited from the verb from which they are derived. It is therefore reasonable to assume that the content of verb-derived compounds is expressed also by several other phrase-like constructions - which can only be found by query expansion. Since syntactic valency features on nominalizations are encoded in the Danish computational lexicon, STO (Braasch & Olsen, 2004)[2], this assumption can be pursued by associating these syntactic features to all compound heads.

To test the second assumption it is necessary to look at the context of the split compound. This strategy can be seen as a variant to the classical closeness criterion: hits where the search words are close to each other are generally better than hits where they are far apart. However, the syntactically-based criterion opens for a more fine-grained analysis where hits with the two search words far apart – but still in the same noun phrase – can be given a high priority. This is desirable for instance in example 1 where the following hit can be found for the query *apoteksovertagelse* (`pharmacyFUGEtakeover`):

1) *Udgifter i forbindelse med <u>overtagelse</u>, nyanlæg eller flytning af et <u>apotek</u>*
   ```
   Expenses   in   relation   to   takeover,
   foundation or movement of a pharmacy
   ```

In example 1, the two search words are five words apart and yet belong to the same, complex NP. The hit is highly relevant to the query since *overtagelse ... af et apotek* (takeover ... of a pharmacy) is synonymous to *apoteksovertagelse*. In contrast, in example 2 the search words are found within the distance of four words, but they do not belong to the same NP. The hit is in fact less relevant since we are no longer dealing with a synonym phrase to *apoteksovertagelse*:

2) *Han lod ved <u>overtagelsen</u> foretage en optælling af <u>apotekets</u> varelager*
   ```
   He let at takeover-THE make a counting of
   pharmacy-THE-GEN³ stock
   ```

In 2) the two query words are in the same sentence but not in the same NP. In conclusion, the second assumption opens for a more nuanced analysis of the context in which the split compound occur. By focusing on the NP we believe it is possible to locate with greater precision the split compound occurrences which are actually synonyms to the original compound.

3. The Test Bed

For test purposes the firm Ankiro generated a compound list of 1000 Danish compounds selecting randomly from a list of 60.000 compounds harvested from a text database, KommuneInformation, provided by the general Danish municipalities. This database (www.kommuneinformation.dk) is a large public database containing text material relevant to all Danish citizens such as government orders, laws, departmental circulars, as well as more general municipality information. When the experiment was run, the database contained 64 million running words. The text material is generally characterized by complex sentence structures influenced by the fact that it lies mostly within the juridical domain.

All compounds were automatically split using the firm's compound analyzer which applies a firm-specific computational lexicon and a set of exception rules. In cases of more than two lemmas in a compound, the coherent boundary was calculated in order to identify which two lemmas (of which one was itself a compound) to carry on with in search; as in *aktie / beskatningsregel* (`share / taxationFUGErule`) in contrast to *arbejdsgiver / medlem* (`workFUGEgiver / member`, 'employer member').

The query test set was modeled as a list containing (split) compounds in isolation - i.e. not with compounds as part of longer queries. The split compounds were used to search in the same database. Hits for split compounds have been defined as text excerpts where the two query words – in any inflectional form – were found within a

---

[3] GEN stands for genitive-*s*.

distance of 10 words. For each query a limit has been set to a maximum of 200 hits.

Seen from a morphological perspective, the corpus material is rather homogenous in the sense that it contains only nominal compounds. However, nominal compounds can be further subdivided with regards to the word class that the head is derived from. Since this aspect plays a role in our assumptions, we have classified the test material into three classes: 1) **Verb-derived compounds**, i.e. compounds where the head is derived from a verb as in *depotanbringelse* (depositinvestment), 2) **Other predicative compounds**, i.e. where the head is not derived from a verb but takes an argument anyhow: *bistandspligt – pligt til* (aidFUGEduty - 'duty to'); and 3) **Non-predicative compounds**, i.e. compounds where the head does not take arguments: *reaktortank* (reactortank).

These three classes are all automatically identifiable when using a computational lexicon with syntactic information; we have therefore made use of the Danish STO lexicon. In this lexicon, verb-derived nouns are encoded differently than other predicative nouns; likewise, the encoding of non-predicative nouns differs from that of the other two classes.

4. Introducing a New Threshold for Retrieved Hits

In this section we turn to the automatic shallow analysis of the search results where the aim is to introduce a new, more critical threshold for retrieved hits on split compounds. The material contains hits on enhanced queries only, search on the original compound is not performed as a part of this experiment. However, since the compounds are harvested from the same text corpus, we know that in these cases all examined compounds would give hits also to the original query.

The search results have been tokenised and pos tagged, and NPs have been tagged using a recognizer written in the CASS formalism (cf. Abney, 1996 and for the Danish implementation, Haltrup, 2000). Noun phrases including up to two prepositional phrases are recognised. Then, all search hits have been assigned a relevance score according to the following algorithm:

$$Score = 10(I - N) \text{ where}$$

I = number of found search words (in any morphological form) in the search hit and
N = number of NPs in the search hit which contains a search word (in any morphological form).

As discussed and illustrated in section 2, the motivation behind the algorithm is based on the assumption that search results where both the search words (obtained by splitting a compound noun) occur in the same noun phrase, are more likely to contain a synonymous phrase to the original compound query. In other words, the algorithm ensures that a higher score is assigned to hits that fulfill this criterion. For illustration, let us consider the first hit in figure 1 where the two search words

omsætning (turnover) and *opgørelse* (account) each occur twice. Following the algorithm, I is therefore set to 4. The number of NPs containing one or more search words in the hit is 2, so N is set to 2. In this case $I – N = 2$, amounting to a score of 20. In other words, it is implicitly calculated that the hit contains two NPs in which both search words occur (Jongejan et al., 2004). Likewise in the first three hits in Figure 2, two occurrences of the two search words *godkendelse* (approval) and *værdi* (approval) are found, so I is set to 2. However the number of NPs containing one or more search words in the hit is in these cases two, so N is set to 1. $I – N = 1$ and a score of 10 is therefore assigned to the hit. Finally, if we consider the last hit in Figure 2, we find two occurrences of the search words, so I is set to 2. The number of NPs containing one or more search words in the hit is now also two, so N is also set to 2. Therefore, the hit achieves a score of 0.

According to this algorithm the threshold has been set to 10; search results with a score below 10 are thus considered 'not retrieved', whereas results with a score of 10 or over are conceived as retrieved documents. In conclusion, hits where both search words are in the same NP are preferred to hits where they are not in the same NP.

Score for search words being in the same noun phrase	Enhanced query: opgørelse omsætning ('account turnover')
20.0	andelsforeninger som nævnt i § 16A , stk . 1 , og som ikke er medlemmer , skal hverken medregnes ved **[NP1 [NP [N \<T>opgørelsen\</T>]] [PRÆP af] [NP [N_GEN andelsforeningens] [N \<T>omsætning\</T>]]]** med henholdsvis medlemmer og ikkemedlemmer eller ved **[NP1 [NP [N \<T>opgørelsen\</T>]] [PRÆP af] [NP [PRON_DEMO den] [ADJ samlede] [N \<T>omsætning\</T>]]]**
10.0	vedrørende sine subsidiære påstand henvist til , at det ikke af Østre Landsretsdommen fremgår , at de af klageren udarbejdede **[NP1 [NP [N \<T>opgørelser\</T>]] [PRÆP over] [NP [V_PARTC_PAST mistet] [N \<T>omsætning\</T>]]]** i andet halvår af 1985 og første halvår af 1986 har dannet grundlag for rettens udmåling af erstatningen
0.0	**[NP [PRON_DEMO Denne] [N \<T>opgørelse\</T>]]** foretages på **[NP2 [NP1 [NP [N grundlag]] [PRÆP af] [NP [N \<T>omsætningen\</T>]]] [PRÆP i] [NP [PRON_DEMO det] [ADJ foregående] [N regnskabsår]]]** . Efter regnskabsårets udløb foretages endelig regulering af afgiftstilsvaret . For nyetablerede virksomheder sker den

Figure 1: Search results on *oversigt* and *omsætning* ('account' and 'turnover') enhanced from the original query *omsætningsoversigt*.

Score for search words being in the same noun phrase	**Enhanced query: godkendelse værdi ('approval value')**
10.0	**[NP1 [NP [N <T9>Godkendelse</T9>]] [PRÆP af] [NP [V_PARTC_PAST forhøjede] [N <T9>værdier</T9>]]]** meddeles for et bestemt tidsrum , som højst kan være 5 år .
10.0	til kl . 22 , en dag ugentlig dog til kl . 24 . Klageren gjorde over for landsskatteretten til støtte for **[NP1 [NP [N <T9>godkendelse</T9>]] [PRÆP af] [NP [PRON_DEMO den] [ADJ selvangivne] [N <T9>værdi</T9>]]]** af den fra arbejdsgiveren modtagne frie kost gældende , at han kun havde modtaget ét måltid om dagen i
10.0	medføre uacceptable skattemæssige konsekvenser , hvis kurs 80 ikke blev godkendt , idet denne kurs havde været **[NP2 [NP1 [NP [PRON_UBST en] [N forudsætning]] [PRÆP for] [NP [N_GEN skattemyndighedernes] [N <T9>godkendelse</T9>]]] [PRÆP af] [NP [N <T9>værdierne</T9>]]]** . I pådømmelsen deltog dommerne Sigrid Ballund , Jochimsen og Jakob Jakobsen (kst.). Det måtte lægges til grund , at sagsøgerens omdannelse
0.0	blev erlagt i form af en græsningsret for svigerfaderens kreaturer på klagerens ejendom . **[NP1 [NP [N <T9>Værdien</T9>]] [PRÆP af] [NP [PRON_DEMO de] [ADJ købte] [N kreaturer]]]** efter de af **[NP1 [NP [N_GEN statens] [N_GEN lignings] [N direktorat] [PRÆP med] [NP [N_GEN ligningsrådets] [N <T9>godkendelse</T9>]]]** fastsatte normalpriser var på købstidspunktet 1. 100 kr . eller 700 kr . lavere end

Figure 2: Search results on *godkendelse* and *værdi* ('approval' and 'value') enhanced from the original query *værdigodkendelse* (value approval).

5. Evaluation

The manually evaluated material amounts to 410 randomly picked compound queries, resulting in 3.367 hits. However, all in all, only one forth of the evaluated enhanced queries produced hits.

In order to calculate the proportion of retrieved material that is actually relevant (precision), as well as the proportion of relevant material actually retrieved in answer to a query (recall), we have as mentioned applied the strategy of defining retrieved hits as search results with a score of 10 or more. According to standards, precision is calculated on the basis of the intersection of the relevant hits and the retrieved hits (as mentioned corresponding in our context to the hits that have passed the threshold) divided with the number of retrieved hits. Recall on the contrary is calculated on the basis of the intersection of the relevant hits and the retrieved hits, divided with the number of relevant hits. Relevant hits are defined on a manual judgment basis: if the search result was considered by the evaluator to contain a synonym to the search query, the retrieved document was tagged as relevant; in other words we apply a rather strict relevance

criterion (even if synonymy can be difficult enough to judge) which proved the best strategy since the compound queries were run without context.

The evaluated test set contains approx. 15% *verb-derived* compounds, and out of these approx. two thirds have the role 'theme' as non-head. For all compounds with 'theme' as non-head, synonymous 'of' constructions were very frequent in the search results, as for *kirkegårdsudvidelse* (graveyardFUGEextension): *udvidelsen af kirkegårde* (extensionTHE of graveyards). Also several of them display the objective genitive construction: *kirkegårdens udvidelse* (the graveyards' extension).
The rest of the group was not quite as systematic regarding synonymous constructions, however, to most of them corresponded synonymous NPs with other prepositions than *af* (of), as for instance in *depotanbringelse* (depositplacement): *anbringelse i depot* ('placement in a deposit') corresponding in this case to a locative role of the non-head.
For the whole group of verb-derived compounds the evaluation results were the following: precision: 0.91 and recall: 0.61. Precision is quite satisfactory, only relatively few of the search results that had passed the threshold were irrelevant to the query. Recall, however, needs a closer examination, since a relatively large number of relevant hits had not passed the threshold. Two factors seem to play a role namely 1) errors in the NP analysis: (complex NPs are not always recognized by the NP analyzser) and 2) copula constructions where the subject predicate denotes a property: whole sentences in terms of copula constructions have not been analyzed as relevant by the system even if they are actually synonymous to the query, as in: *badevandskvalitet – badevandet er af en bestemt kvalitet* (bathwaterFUGEquality – the bath water is of a certain quality).

Also *deadjectivals* (i.e. compounds where the head is derived from an adjective) and *other predicative compounds* follow a predictable pattern. They constitute around 10% of the evaluated material. Actually, since the heads in these compounds subcategorize for a specific preposition, they show even better evaluation results than event nominals, namely a precision of 0.94 and a recall of 0.70. The NPs found in the search results are thus very predictable and follow the encodings in STO, as in example 4:

> 4) *tyveririsiko – risiko for tyveri*
> theftrisk – risk of theft
> (STO encoding on *risiko*: preposition=*for*)

An additional explanation to the higher recall is that complex coordinated NPs do not occur to the same extent with these compounds, probably because they are not derived from verbs.

Around three fourths of the evaluated queries are constituted by *non-predicative compounds*. For one third of these the split query strategy produced no results at all.
For the rest of the non-predicative compounds we get a precision of 0.78 and a recall of 0.52. These results are not as good as for the other two categories. Generally, this

group seems to be less predictable and much more heterogeneous. In other words, some queries give very good results, *husholdningsaffald* (`householdFUGErubbish`) (precision:1.0, recall 0.8), whereas others give very bad results, e.g. *ledelsesstilling* (`managingFUGEposition`) (precision: 0.33, recall: 0.12).

In figure 3 we sum up the test results for the three categories of compounds as well as the average precision and recall for all the evaluated categories. For comparison, the figure also contains the precision obtained when applying no threshold, i.e. when searching on split compounds without any further linguistic restrictions. It should be noted again that the relevance criterion applied is very strict; only when a hit is considered to contain a fully synonymous phrase to the original compound, the hit is categorized as relevant.

	Precision	Recall
All evaluated material without applying the threshold[4]	0.40	not calculated
Verb-derived compounds, applying the threshold	0.91	0.61
Other predicative compounds, applying the threshold	0.94	0.70
Non-predicative compounds, applying the threshold	0.78	0.52
All evaluated material, applying the threshold	0.81	0.54

Figure 3: Search results on enhanced compound queries

6. Conclusions

In conclusion, both assumptions put forward seem to hold. Some compounds are more likely to have synonymous 'split' counterparts than others, namely compounds where the head takes arguments, as is the case for verb-derived compounds or other predicative compounds. For almost all compounds in these two groups, the split compound strategy produces search results, and they generally follow a predictable pattern relating to the subcategorisation structure of the compound head. In contrast, non-predicative compounds in many cases do not produce search results whatsoever to a split query, and when they do, the number of irrelevant hits is higher. Furthermore, this group proved to be much more heterogeneous and less predictable regarding patterns of context.

---

[4] Precision without a threshold is calculated by taking the intersection of all the originally retrieved hits (column D in figure 3) and the relevant hits (subtracting all the non-relevant hits from the total number of hits) and dividing it by the number of originally retrieved hits.

These results relate directly to our second assumption, according to which search results where both the search words (obtained by splitting the compound) are in the same noun phrase, are more likely to contain a synonymous phrase to the original compound query. Again, the first two groups clearly demonstrate that this is the case: precision is high for these two groups meaning that very few of the documents passing the NP threshold, are irrelevant. Non-predicative compounds show a much more blurred picture, both precision and recall being lower and the results in general being very heterogeneous. One serious defect of the shallow linguistic analysis generally influenced the results negatively: the NP recognition in several cases proved to be unsuitable to the document types which include many juridical texts with very complex NPs. In future work we will experiment with an extended NP recognizer capable of recognizing complex coordinated NPs in order to see if recall can be improved without aggravating precision too much.

In spite of the reservations on the technical level of the applied language tools, the experiments performed above clearly demonstrate that compound splitting combined with a shallow linguistic analysis focusing on subcategorisation information and on the recognition of NPs, can improve search by bringing down the number of irrelevant results. When we compare with the precision obtained where no threshold is applied, we see that all three categories of compounds present improved results when applying a linguistically based threshold.

References

Abney, S. (1996). Partial Parsing via Finite-State Cascades. In: *Journal of Natural Language Engineering Vol. 2(4):337-344.*

Braasch, A. & Olsen, S.(2004). STO: A Danish Lexicon Resource - Ready for Applications In: *Proceedings of LREC-2004,* 4:1079-1082, Lisboa.

Chen, A. and F. Gey (2003). Combining Query Translation and Document Translation in Cross Language Retrieval. *Proceedings of CLEF Conference p. 108-121*, 2003.

Dalianis, H, (2005). Improving search engine retrieval using a compound splitter for Swedish. *Proceedings of Nodalida,* Joensuu, Finland, May 20-21, 2005.

Gonzales J., F. Verdejo, C. Peters, & N. Calzolari (1998). Applying EuroWordNet to Cross-lingual Text Retrieval, in: *Computers and the Humanities Vol. 31,*

185-207, Kluwer Academic Publishers, The Netherlands.

Grimshaw, (1990) *Argument Structure*. MIT Press, Cambridge, London.

Haltrup, D. (2000). NP-genkendelse i OntoQuery*. Technical Report.* Center for Sprogteknologi, Københavns Universitet.

Jongejan, B., B.S. Pedersen, C. Navarretta (2004). Automatisk analyse af Zaccos og Ankiros tekstmateriale. *VID rapport no. 3*, Center for Sprogteknologi, University of Copenhagen.

Lapata, M. (2002) The Disambiguation of Nominalzations. In: *Computational Linguistics Vol. 28, Number 3 pp. 357-388*, MIT Press.

Levi, J. (1978) *The Syntax an Semantics of Complex Nominals*, Academic Press. N.Y.

Pedersen, B.S. (Forthcoming). Using shallow linguistic analysis to improve search on Danish compounds. In: J. Tate (ed.) *Natural Language Engineering*. Cambridge University Press.

Smeaton, A. & I. Quigley (1996). Experiments on Using Semantic Distances between Words in Image Caption Retrieval, in : *Proceedings of the 19th International Conference on Research and Development in IR.* 174–180. Zurich, Switzerland.

Strzalkowski, T., L. Gurthrie, J. Karlgren, J. Leistensnider, F. Lin, J. Perez-Carballo, T. Straszheim, J. Wang, J. Wilding (1996). Natural Language Information Retrieval: TREC-5 Report p.291-314. *The Fifth Text Retrieval Conference*, NIST Special Publication.

Strzalkowski, T., F. Lin, J. Perez-Carballo (1997). Natural Language Information Retrieval: TREC-6 Report p.347-366. *The Sixth Text Retrieval Conference*, NIST Special Publication.

Strzalkowski, T., G. Stein, G. Bowden Wise, J. Perez-Carballo, P. Tapanainen, T. Jarvinen, A. Voutilainen, J. Karlgren (1998). Natural Language Information Retrieval: TREC-7 Report p.217-226. *The Seventh Text Retrieval Conference*, NIST Special Publication.

Voorhees, E. (1993). Using WordNet to disambiguate word senses for text retrieval. In: Korfhage, Robert, Edie Rasmussen and Peter Willett, eds., *Proceedings of the 16th Annual ACM SIGIR Conference on Research and Development in Information Retrieval*, 171–180, Pittsburgh.

Voorhees, E. (1994). Query expansion using lexical-semantic relations. In: Croft, W. Bruce and C. J. van Rijsbergen, eds*., Proceedings of the 17th Annual ACM SIGIR Conference on Research and Development in Information Retrieval*, 1994, pp. 61 - 69.

Ørsnes, B. (1995). *The Derivation and Compounding of Event Nominals in Modern Danish – an HPSG Approach with an Implementation in Prolog.* PhD dissertation, University of Copenhagen.