Attribute Value Acquisition through Clustering of Adjectives

Agnieszka Mykowiecka and Małgorzata Marciniak

Institute of Computer Science, Polish Academy of Sciences,
Jana Kazimierza 5, 01-248 Warsaw, Poland
{agn,mm}@ipipan.waw.pl

Abstract. In the paper we analyse Polish descriptive adjectives which occur in domain related texts. The experiments were done on data obtained from hospital discharge records. Prenominal adjectives selected from these texts were filtered out of presumably relative adjectives and clustered on the basis of a set of context related features and interword relations derived from Wordnet. We tested if this procedure can be used to automatically identify concept features, i.e. whether adjectives representing different values of one feature will form one cluster. The obtained results proved to be useful as a preprocessing step in a specialized subdomain ontology creation procedure.

Keywords: Adjectives, clustering, property identification, domain texts.

1 Introduction

An ultimate goal of research done within an area of natural language processing is automatic text and speech understanding. While this task is still much too hard to solve, a lot of simpler but practically useful applications are being developed like information extraction or semantic role labeling. To provide them, one needs a description of searched information, e.g. list of semantic concepts together with their features, and data on how they are expressed in texts. In the case of highly specialized topics, the task of domain model creation has to be performed by domain specialists who posses the required knowledge. But quite often, they are not prepared to organize the extracts from their knowledge into a taxonomy and they fail to enumerate "off-line" all concepts and features which are relevant to the chosen area. As an adequate domain model is crucial to the efficiency and reliability of any computer application built upon it, even partial automatisation of this process could be of practical use.

Domain model creation consists of several steps which can be addressed separately. In a very rough approximation, one can distinguish the steps of concept recognition, definition of concept attributes, and concept taxonomy creation. The first step is usually done by terminology extraction methods which rely on identifying (mainly nominal) phrases in the domain related texts and ranking them according to chosen criteria approximating domain relevance. Standard approaches to automatic terminology extraction are discussed in [15], while some

A. Przepiórkowski and M. Ogrodniczuk (Eds.): PolTAL 2014, LNAI 8686, pp. 92–104, 2014.

improvements to the ranking procedure are proposed for example in [5]. At this step we can identify *left kidney* or *acute appendicitis* as domain significant concepts. Afterwards it would be useful to recognize that *left* and *acute* are the values of some concepts' attributes, which can be named for example as *location* and *type*. The next step is the recognition of relations between concepts. These relations can be of very different kinds, but for ontology creation the most important one is hypernymy-hyponymy relation. Work on hypernymy detection was inspired by research connected with Wordnet population, [12]. In [14] the idea of automating pregrouping of concepts which should be located close in the hierarchy tree was proposed.

The goal of the research presented in this paper was to elaborate a method of attribute indication on the basis of an analysis of descriptive adjectives. Most research conducted in the area of automatic ontology learning was focused on nouns and nominal phrases. Adjectives as such were less studied, but for example [2] used relational adjectives for extraction of hyponyms from medical texts; while [3] and [8] proposed using adjectives for attribute learning. In [10] a semi-supervised machine-learning approach for the classification of adjectives into property denoting (like in *deep wound*) vs. relation denoting adjectives (like in *environmental science*) is presented. In [16] corpus-based methods were used to group Polish adjectives into semantic clusters. Contrary to this latter work, in our research we focus on domain specific texts and we analyse only property denoting adjectives to address the task of property identification. We postulate that adjectives describing values of one feature are used in similar contexts and thus can be automatically identified. In domain specific texts the degree of word ambiguity is lower than in general texts, so we expect to get not very much noise while grouping adjectives themselves not their senses.

Similar work was done by [11], which was aimed at automatic identification of adjectival scales, and performed clustering adjectives as the first stage of this process. Hatzivassiloglou and McKeown utilized only the Kendall τ coefficient counted for adjective-noun pairs and the information of adjective co-occurrence within one phrase.

The paper is organized as follows. The characteristics of the collected data and the process of its linguistic analysis are briefly described in Sect. 2. Section 4 presents a set of features used in the classification process. The next section presents the results of the proposed approach, and finally an evaluation of the results is described in Sect. 6.

2 Data Characteristics

The experiments presented in the paper were performed on hospital discharge documents gathered at one of the Polish children's hospitals. Texts came from six departments (general, surgery, neurology, neonatology, infectious diseases and rehabilitation) and were written by several physicians of different specialties. Original MS Word files were converted into plain text and anonymised before further processing. Next, all texts were analysed using standard general purpose

NLP (Natural Language Processing) tools. The morphological tagger Pantera [1] based on the linguistic data from the Morfeusz analyser [21] was used to divide text into tokens and annotate them with morphosyntactic tags. The morphosyntactic description included a part of speech name (POS), a base form, as well as case, gender and number information, where they are appropriate. The end of sentence tags were also introduced. This information is then used by shallow grammars which were defined to recognize boundaries of nominal phrases.

The data set consists of 3,116 documents comprising 1,940,000 tokens. The tagset we use is quite detailed, thus words which could constitute a nominal phrase can be of one of four categories: *noun (subst), gerund (ger), brev* and *acron*. We are interested in descriptive features expressed both by *adjectives (adj)*, e.g. *duży* 'big', and *past participles (ppas)*, e.g. *powiększony* 'enlarged'. The Pantera tagger does not analyse out of dictionary forms. These forms (11,777 in total) were tagged as *ign* and were not taken into further considerations. Table 1 presents statistics for the relevant syntactical classes recognized within our data.

Table 1. Selected part of speech distribution

	POS	types	occurrences
acronim	acron	1,476	25,814
abreviation	brev	168	76,058
gerund	ger	444	13,405
substantive	subst	3,642	362,352
adjective	adj	1,743	130,611
past participle	ppas	291	23,278

3 Noun Phrases with Adjectival Modifiers

Adjective modification is one of the most typical constructions occurring inside nominal phrases. For English, the thorough studies of the role of adjectives in grammar and their types is presented in [19]. In Polish linguistic tradition, two main groups of adjectives are distinguished, basing on a difference in the type of properties they describe: qualitative (i.e. absolute, descriptive), and relative (i.e. classifying, distinguishing). A similar but a little more precise classification was introduced in [6]. This co-called BEO classification consisted of **b**asic (i.e. qualitative, property-denoting) adjectives (like in *old box*), **o**bject-oriented (relational) adjectives (like in *environmental science*) and a new class of **e**vent-related adjectives (like *eloquent person*). This division was used for example in [7] and [10] in which a semi-supervised machine-learning approach for the classification of adjectives was presented. In [13] a very detailed classification of Polish adjectives into 56 classes made from the point of view of a machine translation application was presented. In our work, we divide adjectives into two groups without introducing a differentiation between basic and event-related adjectives and we perform a clustering experiment on all but relational adjectives.

In Polish noun phrases, adjectives can occur at the both sides of a noun and their position can help in the classification task. This work concentrates on descriptive adjectives, which in Polish occur typically before nouns, e.g. *niski poziom* 'low level', in opposition to classifying adjectives, which are located generally after a noun, e.g. $jama_{subst}$ $brzuszna_{adj}$ 'abdominal cavity'. However, not all adjectives occurring before noun are actually descriptive. Apart from a limited number of examples of lexicalised connections of an adjective and a noun in this very order, e.g. *biały szum* 'white noise' or *ostry dyżur* 'ER', there is a systematic rule of placing at least one classifying adjective before a noun which is modified by more than one relative adjective, e.g. *wojewódzka poradnia alergologiczna* 'provincial allergology clinic' vs. *poradnia wojewódzka* 'provincial clinic' and *poradnia alergologiczna* 'allergology clinic'. In such cases an adjective which describes a more specific aspect of a concept is placed before a noun and a more general adjective is placed after a noun. As the specificity itself is subjective, the order of adjectives changes in different contexts. To be able to filter out classifying adjectives which occurred before nouns, we identified nominal phrases with adjectival modification both before and after nouns. For further processing we selected adjectives that occurred more often before nouns than after them. All phrases obeyed standard Polish gender, case and number agreement constraints. In Table 2 the numbers of types of phrases consisting of one noun and one or more adjectival modifiers which were recognized in the analysed texts, are given (subphrases included in wider phrases are also counted).

Table 2. Types of phrases in data

	nouns		length=2		length=3		length>3	
	types	occ.	types	occ.	types	occ.	types	occ.
A N	1,209	22,425	3,323	21,037	654	1,212	94	176
N A	1,070	63,857	3,067	54,510	754	6,349	375	2,998
A N A	310	4,355	—	—	906	3,426	408	929

Although in general not all adjectives which precede nouns are descriptive, an examination of the data confirmed that adjectives that occur mainly to the left of a noun being a concept are descriptive, and should be represented in the ontology as values of a feature of a concept. For further processing, from the list of adjective-noun phrases we selected adjectives which occurred at least 10 times as left modifiers and occurred relatively much more frequently (two times) as left than as right modifiers.

4 Similarity Features

Our goal was to elaborate a classification scheme which can be easily used for any new types of texts for which no specific ontological resources are available, so we had to rely on such parameters which can be derived directly from text or obtained from general linguistic resources or tools. Thus, we use standard window based features, and context derived coefficients defined specifically for this particular task. Additionally, we use information from the lexicon of closely related words and from Polish Wordnet (plWordnet [16]) – a general, very rich lexical database which contains information on different meanings of words (nouns, adjectives and verbs) and connections between them.

4.1 Lexical Context

We are looking for groups of adjectives which define similar attributes or features of concepts. Values of the same features can be identified by common lexical contexts they occur in. For example, after the expression *zwraca uwagę* 'draws attention' usually untypical aspects of a following concept are given, like *uwypuklone ciemię* 'arched crown', *spłaszczona potylica* 'flattened occiput' or*chrapliwy oddech* 'rasping breath'. Similarity of lexical contexts is a rather obvious and often used feature to define different kind of similarities of particular lexical items, e.g. [12] used patterns like "such NPx as NPy" to identify that NPx is a kind of NPy (NPx, and NPy are noun phrases). This characteristic behaviour can be described by a context consisting of a few words used either to the left or to the right of the expression. It may also be the case that in domain texts some syntactic structures are characteristic for introducing particular types of attributes, so we also decided to test a syntactic characterisation of the context by specifying grammatical classes that words surrounding an adjective belong to.

The chosen maximal length of the context is small, as the analysed documents are rather concise — their authors tend to use short informative phrases and they rapidly change topics.

The context features for an adjective are defined symmetrically for right and left text surrounding an expression consisting of the adjective and a noun (not taking into account a noun which is inside this expression but the subsequent one as the modified noun is taken into account by the $sim_{com-nouns}$ coefficient described in 4.2) and consist of:

- Sequences of base forms of 1, 2 and 3 tokens (afterwards abbreviated to: lb1...lb3 for left, and to rb1...rb3, for right contexts).
- Sequences of POS tags of 2, 3, 4 tokens (lpos2...lpos4, rpos2...rpos4).
- The base form of the nearest:
  - verb (if there are no verbs encountered within the sentence boundaries, this feature value is set to null) (lvp, rvp);
  - noun type token (e.g. nouns, gerunds, acronyms) (lnp, rnp);
  - adjective type token (e.g. adjectives, participles) (ladj, radj);
  - preposition (lpp, rpp).

While establishing contexts we do not go beyond the sentence/paragraph boundaries. In the case of the nearest base form, we ignore similarity which would arise from contexts being ends of sentences or paragraphs, punctuation marks and conjunctions (like *i* 'and'), i.e. in the formula below we do not count adjective occurrences in these contexts.

All lexical similarities were calculated according to the Jaccard coefficient scheme used on form frequencies (we tested also Jaccard coefficient on types and Dice coefficient in both versions, the results were a little lower, but did not differ much; we did not explore more existing similarity measure variants given for example in [20]). In the equation below, t is a context type, C_t is a set of all contexts of the type t, and $ctx(c, a_i)$ is the number of occurrences of an adjective a_i in a context c.

$$\begin{aligned} sim_t(a_i, a_j) & \\ &= \frac{\textit{number of all occurrences of common } C_t \textit{ type contexts}}{\textit{number of occurrences of all } C_t \textit{ contexts of both adjectives}} \\ &= \frac{\Sigma_{c \in C_t} min(ctx(c, a_i), ctx(c, a_j))}{\Sigma_{c \in C_t} ctx(c, a_i) + \Sigma_{c \in C_t} ctx(c, a_j) - \Sigma_{c \in C_t} min(ctx(c, a_i), ctx(c, a_j))} \end{aligned} \quad (1)$$

4.2 Common Modified Nouns

The classification task concerns modifying adjectives, so the most important context in this case is a context consisting of a modified noun. If more than one concept have a given attribute, e.g. size, it is likely that nouns representing these concepts will occur with the same adjectives expressing it, e. g.*mały, powiększony, niewielki* 'small, enlarged, slight'. To account for this observation we established a similarity measure whose value is equal to the division of a number of commonly modified nouns by a maximum number of nouns modified by one adjective in the analysed text:

$$\begin{aligned} sim_{com-nouns}(a_i, a_j) &= \frac{\textit{number of commonly modified nouns}}{\textit{maximum number of nouns modified by one adjective}} \\ &= \frac{\Sigma_{n \in N:\ frq(a_i n) > 0 \textit{ and } frq(a_j n) > 0} \ 1}{max_{n \in N} \left(\Sigma_{a_i \in A: frq(a_i n) > 0} \ 1\right)} \end{aligned} \quad (2)$$

where A is a set of analysed adjectives, N is a set of nouns, $frq(a_i n)$ is a frequency of a bigram consisting of a_i and n.

4.3 Wordnet Similarity

Polish Wordnet is a large net of lexical meanings connected via different relations. In particular, it contains information on a lot of adjectives. Unfortunately, the adjective hypernymy hierarchy is very flat and most adjectives have one or at most two nodes above them so adjectives' similarity is hard to derive from there. None of the wordnet-based similarity measures provided at the plWordnet site, works for adjectives. Thus, we used information from plWordnet relations (antonymy, synsets relatedness and hypernymy) directly to implement similarity measures for adjectives. As adjectives in our data are not semantically disambiguated, we aggregate information given for all meanings.

The relation which is the most numerous in the adjectival part of plWordnet is antonymy. Although this relation, of course, does not link words with the same meaning, it may link adjectives expressing different values of one attribute, e.g. *duży, mały* 'big small', so we use it as a source of one of the similarity measures ($\mathrm{sim}_{w_{a-ant}}$).

Two other plWordnet relations which we use are hypernymy and relatedness. They are directly connected with meaning similarity, and information coming from these two sources is represented jointly by one similarity feature ($\mathrm{sim}_{w_{a-hyp}}$). This coefficient takes into account adjectives from the same synset.

In some cases additional information can be obtained indirectly from the antonymy relation. Words which share the same antonyms (for some of their senses) are similar so we added the coefficient $\mathrm{sim}_{w_{a-sim}}$ basing on this information.

The three plWordnet similarity features are defined as:

$$\begin{aligned} sim_{w_{a-X}}(a_i, a_j) &= \frac{number\ of\ senses\ in\ relation\ X}{smaller\ number\ of\ senses} \\ &= \frac{\Sigma_{si \in S(a_i), sj \in S(a_j):\, ((si\ X\ sj)\ or\ (sj\ X\ si))} 1}{min\,(S(a_i), S(a_j))} \end{aligned} \tag{3}$$

where a_i and a_j are adjectives, $S(a_i)$ is a set of senses of a_i and X is one of three possible relations: an antonymy, a sum of hypernymy and synsets relatedness, or the similarity relation defined above.

Although plWordnet contains numerous adjectives, quite a number of the words from our list are not represented there. As many adjectives are derived from nouns, for these forms we used additional information from the noun hyperonymy hierarchy. Using a set of rules we transformed adjectives into possible noun forms and calculated $\mathrm{sim}_{w_{n-hyp}}$ coefficient in the same way as above for those nominal forms which were found in plWordnet (nearly no links between adjectives and nouns are present in this data at the moment). The fourth plWordnet similarity feature defined using this information is:

$$sim_{w_{n-hyp}}(a_i, a_j) = \frac{\Sigma_{si \in S(n(a_i)), sj \in S(n(a_j)):\, ((si\ hypernim\ sj)\ or\ (sj\ hypernim\ si))} 1}{min\,(S(n(a_i)), S(n(a_j)))} \tag{4}$$

where $n(a_i)$ is a noun derived from a_i.

4.4 Data from the Dictionary of Polish Synonyms

As hypernymy relation among adjectives in plWordnet is not very elaborated, we additionally used data included in an open source dictionary of synonyms (`http://synonimy.ux.pl`). It contains 13,180 groups with 44,550 words or phrases. From this data set we obtained 85 similarity pairs for the considered list of adjectives. As, in this case, no frequency data nor number of senses described are available, the similarity is defined as:

$$sim_{dict}(a_i, a_j) = \frac{number\ of\ groups\ with\ both\ elements}{bigger\ number\ of\ groups\ for\ both\ elements} \tag{5}$$

5 Clustering

Automatic clustering of adjectives was done using Multidendrograms tool [9] – a program which implements hierarchical clustering and solves the non-uniqueness problem found in the standard pair-group algorithm by grouping more than two clusters at the same time when ties occur. As input, it takes singular similarity values for all pairs of adjectives which were counted as a weighted sum of 26 features described in the previous section. In (6) S_{sim} is the set of all similarity coefficients described in Sect. 4.

$$sim(a_i, a_j) = \Sigma_{sim_t \in S_{sim}} weight(sim_t) \times sim_t(a_i, a_j) \tag{6}$$

We performed several experiments with different weights assigned to different groups of features. So, we tested if any features help in obtaining better results, and which of them are the most valuable in this task. The complete linkage strategy was chosen as the clustering method.

In the first model, the weights were tuned on the basis of manual grouping of the 28 most popular adjectives that appeared at least 50 times in adjective-noun phrases in the corpus. The procedure of selecting these adjectives was the following. From the whole set of 105 adjectives that appeared at least 50 times in adjective-noun phrases we selected those that were used three times more frequently to the left of an adjective than to the right. After this step we obtained 65 adjectives. From this set we selected those that create at least two-element groups according to the information obtained from the plWordnet and the thesaurus of synonyms. This set consists of adjectives that are frequent in general language so it was relatively easy to group them manually and create a development set for tuning weights of coefficients. The grouping reflects projection of annotator's knowledge about the domain into synsets represented in plWordnet. This step resulted in 12 groups given in Table 3.

The method of constructing the starting set of adjectives was motivated by our desire to obtain a data set which includes some multi-element groups whose automatic identification can be then later tested. Due to the limited capacity of our hospital data, many properties are represented there by only one possible value so randomly chosen adjectives might create mostly one element groups.

In the process of manual tuning of the set of weights we performed several experiments described below. Their results were compared with a manually created classification using the B-cubed measure [4] which is sensitive to the presence and absence of the elements of the compared groups. The weights from the best manually tuned model are presented in Table 4. For this model we obtained F-measure of 0.799 for the division into 12 groups, while the highest F-measure

Table 3. Manual grouping of 28 adjectives

group 1: *drobny* 'small', *niewielki* 'small', *duży* 'large', *powiększony* 'enlarged';
group 2: *pozostały* 'remaining';
group 3: *stały* 'stable, constant';
group 4: *kolejny* 'next', *ponowny* 'repeated', *początkowy* 'initial', *ostatni* 'last', *pierwszy* 'first' *drugi* 'second';
group 5: *obfity* 'abundant', *liczny* 'numerous', *nieliczny* 'not numerous';
group 6: *nieznaczny* 'minor, insignificant', *znaczny* 'considerable, substantial', *istotny* 'important';
group 7: *obustronny* 'two-sided', 'mutual;
group 8: *rozluźniony* 'loose' 'relaxed';
group 9: *podwyższony* 'increased', *wzmożony* 'enhanced', *niski* 'low', *wysoki* 'high',
group 10: *różny* 'different';
group 11: *płynny* 'liquid', 'floating',
group 12: *silny* 'strong', *słaby* 'weak'

of 0.813 was obtained for 13 groups. If the information from the thesauri was neglected (the appropriate weights in the model given in Table 4 were set to a 0 value) the F-measure dropped to 0.758. The highest F-measure of 0.809 was obtained for 15 groups.

Table 4. The model

	coeff.	value	coeff.	value	coeff.	value	coeff.	value	coeff.	value
left/right POS										
	lpos2	0.12	lpos3	0.12	lpos4	0.12				
	rpos2	0.09	rpos3	0.09	rpos4	0.09				
left/right base form										
	lb1	0.12	lb2	0.12	lb3	0.12				
	rb1	0.09	rb2	0.09	br3	0.09				
left/right nearest verb, noun, adjective, preposition										
	lvp	0.12	lnp	0.30	ladj	0.25	lpp	0.06		
	rvp	0.90	rnp	0.25	radj	0.20	rpp	0.04		
thesauri										
	w_{a-sim}	0.20	w_{a-ant}	0.40	w_{a-hyp}	0.10	w_{n-hyp}	0.15	dict	0.20
common nouns		0.50								

To discover how important the particular type of coefficient is, we compared models with the only one non-zero value of the weight related to this coefficient. For all types of coefficients we checked the F-measure for the division consisting of 12 groups. We observed how many groups were created, how many adjectives were not linked and how quickly the set of adjectives was divided. The last information could be obtained by analysing thresholds for which the set of adjectives is divided into chosen number of groups. One of the questions we set out to explore was whether left and right contexts are equally important. In order to answer this, we set all left contexts based on POS and base forms to the same value. We obtained an F-measure of 0.601 in comparison with manual model, while the right contexts gave a slightly worse result (0.556). Similar differences were found comparing left and right sets of coefficients based on nearest (results for the left contexts): noun (0.609 – for threshold 0.043), verb (0.569 – 0.106), adjective (0.565 – 0.051), and preposition (0.541 – 0.0015). In this case combining left contexts give a 0.622 F-measure while the right one 0.577. So in the manual model the left contexts have slightly higher weights than the right one.

Table 5. Results for groups of coefficient types

type of similarity	F-measure
left noun/adjective/verb/preposition	0.622
right noun/adjective/verb/preposition	0.577
both nouns/adjectives/verbs/prepositions	0.699
left strings of base forms and pos	0.601
right strings of base forms and pos	0.556
both strings of base forms and pos	0.636
wordnet relations	0.655
common nouns	0.698–0.75

Table 5 presents the results for several groups of coefficient types. They show that the division closest to the manually created one is obtained for the coefficient based on common nouns modified by adjectives. Although there is no division into 12 groups, the division into 11 groups gives an F-measure of 0.698, while for 14 groups – 0.75. A somewhat surprising result was obtained for coefficients based on plWordnet and the synonym thesaurus, as it is slightly worse than the result for the combination of nearest noun/adjective/verb/preposition coefficients. The reasons for this may be twofold. First, we do not perform any word sense disambiguation, so for example, the word *stały* 'stable, constant' is connected in one group with *płynny* 'liquid', so the first meaning is preferred, while in our data it is used in another meaning *stała opieka* 'constant care', or *stały ból* 'constant pain'. The second problem is that the plWordnet hierarchy of adjectives is not very elaborated. It is being developed intensively at the moment, so we expect that the results will improve in the future.

To test to what extent the suggested method can be used in an automatic mode, without any manual parameter tuning step, we check the result for a model in which weights assigned to all similarity coefficients were given the same nonzero value. For this model we obtained the F-measure of 0.651 for 12 groups, while the best F-measure of 0.718 was for 20 groups. So, although the results are lower, they are not very much different from those obtained via manual tuning.

6 Evaluation

An evaluation of the method was done on the basis of 101 adjectives that appeared in the data at least 30 times in adjective-noun phrases and were used two times more often to the left of a noun than to the right.

A manual grouping of these 101 adjectives was done according to the subjective knowledge of an annotator familiar with the domain and data. The starting point of this task was the result of grouping the most common 28 adjectives. The grouping of 101 adjectives was checked by another annotator, who suggested only one change. This process resulted in a division consisting of 52 groups, 28 of them containing only one element.

The automatic clustering of the evaluation set using the manually tuned model described in the previous section, compared with the manual grouping for the same number of groups, gave following results: precision 0.676, recall 0.653 and F-measure: 0.664.

The evaluation data contained 28 adjectives that took part in tuning the model. If we removed these 28 adjectives from the final divisions we got 34 one element groups (6 more) and that affected the results of comparison making them better than they were. It turned out that these initially selected 28 adjectives were ‘seeds’ of larger groups.

The results obtained for the evaluation set using the model consisting of equal coefficients for the division into 52 groups were following: precision 0.626, recall 0.620 and F-measure 0.623. Thus, the results obtained using this non tuned model were only slighty lower for this bigger set than for the initial 28 adjectives. Again, the difference between results for the tuned model and the simple one was not very big.

Table 6 presents one group obtained by automatic clustering of 101 adjectives together with equivalents from the manual division. In the later data *dodatni* was included in the separate group together with *negatywny* ‘negative’. The last line presents the respective group obtained by automatic clustering of 28 adjectives (the other 2 adjectives are not present in this smaller set).

Table 6. An examplary result

method	set	group				
automatic	101	*dodatni*,	*niski*,	*wysoki*,	*obniżony*,	*podwyższony*
		positive,	low,	high,	reduced,	increased
manual	101		*niski*,	*wysoki*,	*obniżony*,	*podwyższony*
			low,	high,	reduced,	increased
manual	28		*niski*,	*wysoki*,		*podwyższony*
& automatic			low,	high,		increased

7 Conclusion

From the performed experiments it appears that automatic clustering basing on the suggested parameters set can be used to preliminary identify adjectives describing one property in big enough and coherent data. Adjectives that are not related with other adjectives in plWordnet or other thesauri can be clustered only on the basis of the syntactic information available in the corpus and the results are still acceptable.

The most informative feature in our model was the coefficient based on common nouns modified by both compared adjectives. In further work we plan to test different ways of defining this particular similarity measure to check if these results could be even better. Moreover, the left contexts turned out to be only slightly more important than the right ones, and contexts based on the nearest nouns/adjectives/verbs/prepositions are only slightly better than contexts based on POS and base forms.

To improve the results, we plan to extend the experiments presented here by combining them with adjectives' sense disambiguation task using clustering methods allowing for the placement of one word in more than one cluster like in [18].

References

1. Acedański, S.: Tager morfosyntaktyczny PANTERA. In: [17], pp. 197–207
2. Acosta, O., Aguilar, C.A., Sierra, G.: Using relational adjectives for extracting hyponyms from medical texts. In: Lieto, A., Cruciani, M. (eds.) AIC@AI*IA. CEUR Workshop Proceedings, vol. 1100, pp. 33–44. CEUR-WS.org (2013)
3. Almuhareb, A., Poesio, M.: Attribute-based and value-based clustering. An evaluation. In: Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing, pp. 158–165 (2004)
4. Bagga, A., Baldwin, B.: Algorithms for scoring coreference chains. In: The First International Conference on Language Resources and Evaluation Workshop on Linguistics Coreference, pp. 563–566 (1998)
5. Barrón-Cedeño, A., Sierra, G., Drouin, P., Ananiadou, S.: An improved automatic term recognition method for Spanish. In: Gelbukh, A. (ed.) CICLing 2009. LNCS, vol. 5449, pp. 125–136. Springer, Heidelberg (2009)
6. Boleda, G.: Automatic acquisition of semantic classes for adjectives. PhD thesis, Pompeu Fabra University (2006)
7. Boleda, G., Badia, T.: Morphology vs. syntax in adjective class acquisition. In: Proceedings of the ACL-SIGLEX Workshop on Deep Lexical Acquisition, DeepLA 2005, pp. 77–86. Association for Computational Linguistics, Stroudsburg (2005)
8. Cimiano, P.: Ontology Learning and Population from Text. Algorithms, Evaluation and Applications. Springer (2006)
9. Fernández, A., Gómez, S.: Solving non-uniqueness in agglomerative hierarchical clustering using multidendrograms. Journal of Classification 25, 43–65 (2008)
10. Hartung, M., Frank, A.: A semi-supervised type-based classification of adjectives: Distinguishing properties and relations. In: Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC 2010). ELRA, Valletta (2010)

11. Hatzivassiloglou, V., McKeown, K.: Towards the automatic identification of adjectival scales: Clustering adjectives according to meaning. In: Schubert, L.K. (ed.) ACL, pp. 172–182. ACL (1993)
12. Hearst, M.A.: Automated discovery of WordNet relations. In: Fellbaum, C. (ed.) WordNet: An Electronic Lexical Database, pp. 131–153. MIT Press (1998)
13. Jassem, K.: Semantic classification of adjectives on the basis of their syntactic features in Polish and English. Machine Translation 17(1), 19–41 (2002)
14. Mykowiecka, A., Marciniak, M.: Combining wordnet and morphosyntactic information in terminology clustering. In: Proceedings of the 24th International Conference on Computational Linguistics (COLING 2012), Mumbai, India (2012)
15. Pazienza, M., Pennacchiotti, M., Zanzotto, F.: Terminology Extraction: An Analysis of Linguistic and Statistical Approaches. In: Sirmakessis, S. (ed.) Knowledge Mining. STUDFUZZ, vol. 185, pp. 255–279. Springer, Heidelberg (2005)
16. Piasecki, M., Szpakowicz, S., Broda, B.: A Wordnet from the Ground Up. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław (2009)
17. Przepiórkowski, A., Bańko, M., Górski, R.L., Lewandowska-Tomaszczyk, B. (eds.): Narodowy Korpus Języka Polskiego. PWN, Warsaw (2012)
18. Tomuro, N., Lytinen, S.L., Kanzaki, K., Isahara, H.: Clustering using feature domain similarity to discover word senses for adjectives. In: Semantic Computing, ICSC 2007, pp. 370–377 (September 2007)
19. Warren, B.: Classifying Adjectives. Acta Universitatis Gothoburgensis, Göteborg (1984)
20. Weeds, J., Weir, D.J.: Co-occurrence retrieval: A flexible framework for lexical distributional similarity. Computational Linguistics 31(4), 439–475 (2005)
21. Woliński, M.: Morfeusz — a practical tool for the morphological analysis of Polish. In: Intelligent Information Processing and Web Mining. AISC, vol. 35, pp. 511–520. Springer, Heidelberg (2006)