

COMPARING METHODS OF AUTOMATIC VERB-NOUN COLLOCATION EXTRACTION

Svetlana Koshcheeva, Victor Zakharov

Abstract: *Automatic verb-noun collocation extraction is an important natural language processing task, the results of which can be applied in various spheres including machine translation, language teaching, summarization, information extraction, disambiguation, etc. The paper describes a set of experiments the aim of which is to compare several approaches to automatic verb-noun collocation extraction. The main subjects under observation are the impact of span size and POS-filtering on the quality of collocation extraction. The experiments have shown that collocations lists extracted by means of POS-filtering are significantly more precise than those obtained without POS-filtering, whereas the extension of a span size has an ambiguous effect. On the one hand, it enables the extraction of distant collocates, but on the other hand it results in erroneous collocates, which leads therefore to consider the use of syntax-based approach for verb-noun collocation extraction.*

Keywords: *Corpora, statistical methods, collocations, automatic collocation extraction.*

ACM Classification Keywords: *I.2.7 Natural Language Processing*

Introduction

The study of collocations is a state-of-the-art approach to the analysis of compatibility of lexical units. Collocation extraction is applied in many fields of linguistic technologies such as semantic and lexicographic research (including the creation of dictionaries and grammars of a new type), language teaching, machine translation, automatic analysis and disambiguation.

Statistical methods of collocation extraction have become widely used in modern corpus linguistics. The simplest way to detect collocations in texts is making up frequency lists of words, which appear to the left or to the right of a node word within the indicated span. The size of this span is usually 5 words to the left or to the right of a node. Statistical association measures (log-likelihood, MI, t-score etc.) have become extremely widespread in modern linguistic research. These measures are based on cooccurrence frequencies of word pairs and frequencies of each constituent (e.g., see [Peina, 2009]), which can as well be calculated within a certain span.

In our research as an association and ranking measure we use Mutual Information (MI) [Church, 1991] which can be understood as a coefficient of syntagmatic strength between collocation constituents [Evert, 2004]. MI for a bigram is calculated by the following formula:

$$MI = \log_2 \frac{f(n, c) * N}{f(n) * f(c)} \quad (1)$$

where MI is mutual information measure; $f(n, c)$, $f(n)$, $f(c)$ are absolute frequencies in a corpus of cooccurrence of n (node) and c (collocate) and words n , c respectively; N is a corpus size.

However, it is necessary to take into account the fact that words are syntactically related, they do not occur in texts haphazardly or by chance. Therefore, collocation extraction requires not only statistical methods but

also syntax-based approaches, which take into consideration morphological and syntactic properties of words in a corpus.

The aim of this paper is to examine and compare methods of automatic collocation extraction. As a tool for our investigation we have chosen IntelliText, a system developed by the Centre for Translation Studies (CTS) at the University of Leeds (<http://corpus.leeds.ac.uk/it/>). IntelliText offers ample opportunities for linguistic research and has representative corpora, including morphologically annotated Russian corpora.

Methodology

The subject of our investigation are verb-noun collocations of the type “*verb + noun* (in the accusative without a preposition)”. As a tool for verb-noun collocation extraction we used the system IntelliText. The corpus RNC2010-MOCKY, a 2010 version of the Russian National Corpus of 116 mln words, served as a material for our research.

For our experiments we have chosen the following verbs: *выполнять* ('to carry out'), *нарушать* ('to violate'), *принимать* ('to accept'). On the basis of the Dictionary of Collocations in the Russian Language [Slovar' sochetaemosti slov russkogo jazyka, 1983], the Dictionary of Russian Verb-Noun Collocations [Deribas, 1983] and Small Academic Dictionary [Slovar' russkogo jazyka, 1981 – 1984] we made up a list of verb-noun collocations for the chosen verbs with which we compared word-combinations extracted by IntelliText. Below one can see the list of verb-noun collocations for the verb *выполнять*:

- | | |
|-----------------------------|----------------------------|
| 1. выполнять директивы | 18. выполнять поручение |
| 2. выполнять долг | 19. выполнять правила |
| 3. выполнять желание | 20. выполнять приказ |
| 4. выполнять задание | 21. выполнять приказание |
| 5. выполнять задачу | 22. выполнять программу |
| 6. выполнять заказ | 23. выполнять просьбу |
| 7. выполнять заявку | 24. выполнять работу |
| 8. выполнять инструкцию | 25. выполнять распоряжение |
| 9. выполнять каприз | 26. выполнять решение |
| 10. выполнять команду | 27. выполнять роль |
| 11. выполнять нагрузку | 28. выполнять совет |
| 12. выполнять наказ | 29. выполнять требование |
| 13. выполнять норму | 30. выполнять указание |
| 14. выполнять обещание | 31. выполнять упражнение |
| 15. выполнять обязанности | 32. выполнять условие |
| 16. выполнять обязательства | 33. выполнять установку |
| 17. выполнять план | 34. выполнять функцию |

The aim of the first experiment was to study the impact of part of speech (POS) filtering on the quality of collocation extraction (specifying the POS of a candidate collocate). The spans [-1, 1], [-2, 2], [-3, 3] were under observation.

The second experiment consisted in investigating the impact of a span size on the quality of verb-noun collocation extraction with POS-filtering. Spans up to 5 words to the right of the verb were studied.

In the course of these experiments it was discovered that among phrases extracted by IntelliText there are those which are not fixed in the dictionaries mentioned above but which nevertheless can be considered as set expressions, i.e. collocations. In order to decide which phrases extracted by IntelliText for the verb *выполнять* are collocations an expert evaluation was carried out. According to this evaluation our list of verb-noun collocations was expanded by the following phrases: *выполнять волю, выполнять движение, выполнять действие, выполнять контракт, выполнять маневр, выполнять миссию, выполнять обработку, выполнять операцию, выполнять пожелание, выполнять предписание, выполнять прыжок, выполнять расчёт, выполнять рекомендацию, выполнять соглашение, выполнять трюк.*

To estimate the quality of verb-noun collocation extraction we calculated precision, recall and F-measure using the following formulas:

$$Precision = \frac{|D_{rel} \cap D_{retr}|}{|D_{retr}|} \quad (2)$$

where D_{rel} is a set of relevant expressions from the list of verb-noun collocations; D_{retr} – a set of phrases extracted by IntelliText.

$$Recall = \frac{|D_{rel} \cap D_{retr}|}{|D_{rel}|} \quad (3)$$

where D_{rel} is a set of relevant expressions from the list of verb-noun collocations; D_{retr} – a set of phrases extracted by IntelliText.

$$F = (\beta^2 + 1) \frac{Precision \times Recall}{\beta^2 Precision + Recall} \quad (4)$$

for $\beta = 2$ (the priority is given to recall).

Experiments

In Table 1 one can see results of the first experiment (the impact of POS-filtering on verb-noun collocation extraction) for the verb *выполнять* ('to carry out'). Ranks of collocate candidates are indicated in the first column, other columns contain collocate candidates extracted by IntelliText for different spans, with and without POS-filtering. Words highlighted in white are collocates for the verb *выполнять* ('to carry out') according to our list of verb-noun collocations.

Table 1. Verb-noun collocation extraction by IntelliText (with and without POS-filtering) for the verb *выполнять* ('to carry out')

rank	[-1, 1]		[-2, 2]			[-3, 3]	
	without filtering	POS-noun in the accusative	without filtering	POS-noun in the accusative	without filtering	POS-noun in the accusative	
1.	организационные	приказание	организационные	приказание	перевыполнят ь	приказание	
2.	пунктуально	функция	перевыполнять	функция	беспрекослов но	функция	
3.	беспрекословно	наказ	беспрекословно	поручение	неукоснитель но	поручение	
4.	неукоснительно	поручение	неукоснительно	наказ	функция	наказ	
5.	добросовестно	задание	добросовестно	задание	добросовестн	задание	

					о	
6.	приказание	приказ	приказание	обязанность	приказание	обязанность
7.	функция	упражнение	функция	предписание	поручение	предписание
8.	скрупулёзно	указание	поручение	приказ	задание	обещание
9.	безукоризненно	заказ	наказ	обещание	обязанность	заказ
10.	наказ	предписание	задание	заказ	упражнение	приказ
11.	безропотно	обязанность	упражнение	упражнение	обещание	упражнение
12.	поручение	роль	обещание	обязательство	заказ	обязательство
13.	приказ	директива	обязанность	указание	предписание	миссия
14.	задание	обещание	приказ	директива	приказ	указание
15.	исправно	рейс	заказ	миссия	обязательство	директива
16.	упражнение	инструкция	предписание	трюк	указание	предназначение
17.	возложить	завет	обязательство	роль	миссия	долг
18.	послушно	обязательство	указание	долг	возложить	трюк
19.	интернациональный	распоряжение	возложить	распоряжение	долг	роль
20.	заказ	задача	послушно	предназначение	защитный	распоряжение
21.	указание	маневр	миссия	завет	роль	задача
22.	предписание	трюк	трюк	задача	задача	завет
23.	управленческий	требование	защитный	пожелание	распоряжение	пожелание
24.	обязанность	приветствие	долг	приветствие	рейс	каприз
25.	обещание	миссия	роль	рейс	инструкция	рейс
26.	качественно	пожелание	распоряжение	инструкция	воинский	инструкция
27.	роль	полёт	задача	каприз	обязанный	приветствие
28.	защитный	заповедь	рейс	маневр	требование	просьба
29.	призванный	норма	эффективно	требование	служебный	требование
30.	эффективно	рекомендация	инструкция	рекомендация	успешно	заповедь
31.	обязательство	обряд	воинский	просьба	просьба	рекомендация
32.	успешно	план	обязанный	заповедь	работа	бросок
33.	рейс	воля	требование	обряд	строго	маневр
34.	старательно	просьба	успешно	воля	норма	обряд
35.	инструкция	долг	служебный	полёт	операция	норма
36.	обязанный	блок	рекомендация	норма	полёт	воля
37.	распоряжение	работа	строго	ритуал	боевой	работа
38.	строго	стрельба	просьба	работа	способный	ритуал

39.	задача	посадка	полёт	план	честно	полёт
40.	ежедневно	команда	работа	стрельба	план	план
41.	чётко	операция	способный	перевозка	воля	посадка
42.	полёт	удар	чётко	посадка	команда	перевозка
43.	честно	готовность	норма	операция	объём	монтаж
44.	способный	условие	боевой	команда	позволять	классификация
45.	требование	танец	честно	прыжок	одновременн о	стрельба
46.	самостоятельно	поворот	воля	заявка	удар	операция
47.	боевой	множество	операция	блок	определённы й	назначение
48.	служебный	программа	план	спектр	сложный	команда
49.	важнейший	фигура	позволять	назначение	приходиться	процедура
50.	отказываться	желание	одновременно	программа	правило	прыжок
51.	броситься	решение	команда	процедура	действие	заявка
52.	норма	ряд	объём	ремонт	важный	спектр
53.	одновременно	правило	удар	установка	различный	блок
54.	фактически	способность	определённый	желание	движение	желание
55.	позволять	расчёт	приходиться	удар	предприятие	нагрузка
56.	план	закон	сложный	условие	свой	программа
57.	приходиться	контроль	прямой	фигура	любой	установка
58.	работа	возможность	важный	цикл	программа	действие
59.	определённый	движение	различный	обслуживание	социальный	удар
60.	воля	шаг	свой	правило	организация	ремонт
61.	просьба	действие	правило	действие	должен	обслуживан ие
62.	команда	срок	любой	расчёт	условие	правило
63.	долг	качество	движение	контракт	решение	комплекс
64.	способность	сила	программа	готовность	который	условие
65.	объём	-	должен	решение	лишь	фигура
66.	свой	-	социальный	объём	также	решение
67.	операция	-	действие	танец	всегда	цикл
68.	различный	-	предприятие	поворот	они	расчёт
69.	сложный	-	условие	множество	весь	объём
70.	удар	-	организация	контроль	надо	контракт
71.	любой	-	решение	комплекс	каждый	готовность
72.	специальный	-	основной	приём	по	движение
73.	должен	-	государственный	ряд	этот	контроль
74.	важный	-	закон	услуга	или	танец
75.	поставить	-	также	движение	другой	приём

76.	продолжать	-	всегда	закон	мочь	поворот
77.	условие	-	лишь	соглашение	только	услуга
78.	прийтись	-	который	способность	его	множество
79.	точно	-	она	исследование	тот	способность
80.	следующий	-	надо	разработка	,	закон
81.	лишь	-	весь	шаг	и	ряд
82.	также	-	можно	проект	не	течение
83.	всегда	-	мочь	товар	они	срок
84.	решение	-	только	анализ	же	урок
85.	начать	-	этот	техника)	исследован ие
86.	весь	-	другой	возможность	она	соглашение
87.	они	-	его	очередь	мы	шаг
88.	надо	-	или	течение	то	разработка
89.	можно	-	по	срок	.	реализация
90.	мочь	-	тот	совет	но	проект
91.	который	-	не	число	:	техника
92.	его	-	он	организация	"	очередь
93.	только	-	она	часть	что	счёт
94.	тот	-	мы	счёт	быть	круг
95.	этот	-	и	время	в	возможност ь
96.	не	-	,	качество	а	совет
97.	она	-)	право	как	число
98.	быть	-	быть	положение	"	время
99.	он	-	"	процесс	с	часть
100.	,	-	.	день	на	день
101.	и	-	в	дело	—	дело

In Table 2 and Figures 1, 2 and 3 one can see evaluation of the results of the first experiment for the spans [-1, 1] (see Fig. 1), [-2, 2] (see Fig. 2) and [-3, 3] (see Fig. 3).

Table 2. Evaluation of the results of verb-noun collocation extraction for the verb *выполнять* ('to carry out')

Span	Precision	Recall	F-measure
[-1, 1] (without POS-filtering)	0,287129	0,591837	0,488215
[-1, 1] (with POS-filtering)	0,625	0,816327	0,769231
[-2, 2] (without POS-filtering)	0,336634	0,693878	0,572391
[-2, 2] (with POS-filtering)	0,465347	0,959184	0,791246
[-3, 3] (without POS-filtering)	0,316832	0,653061	0,538721
[-3, 3] (with POS-filtering)	0,485149	1	0,824916

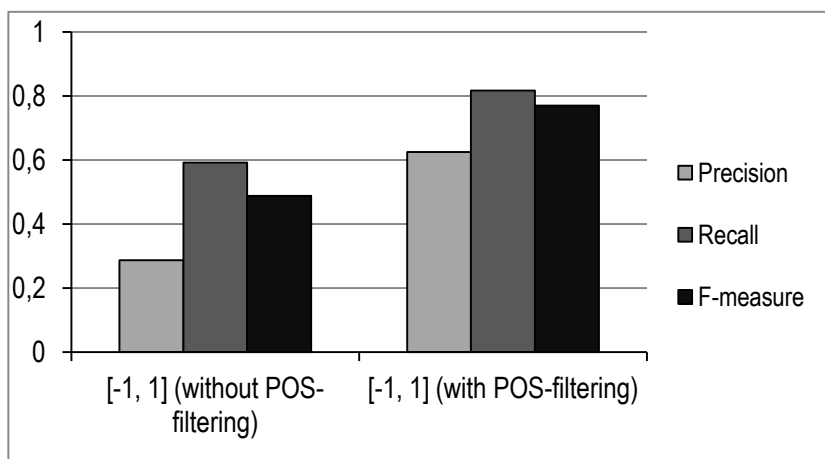


Fig. 1. Evaluation of the results of verb-noun collocation extraction for the verb *выполнять* ('to carry out') (span = [-1, 1])

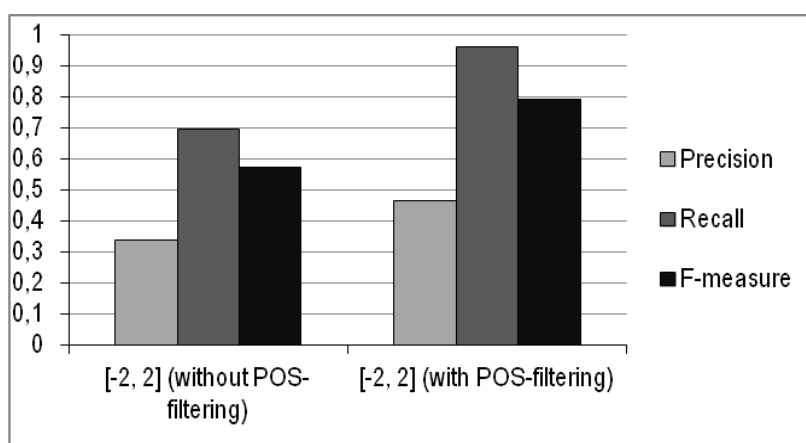


Fig. 2. Evaluation of the results of verb-noun collocation extraction for the verb *выполнять* ('to carry out') (span = [-2, 2])

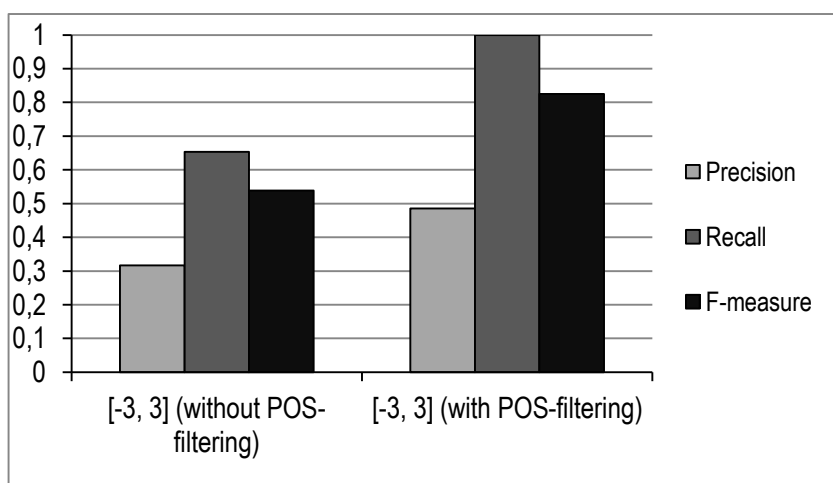


Fig. 3. Evaluation of the results of verb-noun collocation extraction for the verb *выполнять* ('to carry out') (span = [-3, 3])

Figures 1, 2 and 3 show the same tendency: POS-filtering raises precision, recall and F-measure. This observation leads us to conclude that POS-filtering increases the quality of verb-noun collocation extraction. At the same time, the influence of a span size on verb-noun collocation extraction is not essential.

The analysis of the phrases extracted by IntelliText in the first experiment has shown that the left context doesn't provide relevant collocates for the verbs within our research. It can be explained by the fact that in the Russian language the direct object (a noun in the accusative) comes after the verb. Therefore, in the second experiment we confined ourselves to the right context, limiting the span up to 5 words.

For the indication of a span size we use the signs [0, 1], [0, 2], [0, 3], [0, 4] and [0, 5], which mean a zero left context and a right context equal to one, two, three, four and five words respectively.

Results of the second experiment (the impact of a span size on verb-noun collocation extraction for the verb *выполнять* ('to carry out')) are presented in Table 3. Ranks of collocate candidates are indicated in the first column, other columns contain nouns in the accusative extracted by IntelliText for different right contexts of the verb. Nouns highlighted in white are collocates for the verb *выполнять* ('to carry out') according to our list of verb-noun collocations.

Table 3. Verb-noun collocation extraction for the verb *выполнять* ('to carry out') by IntelliText (span size changed)

rank	[0, 1]	[0, 2]	[0, 3]	[0, 4]	[0, 5]
1.	приказание	приказание	приказание	приказание	функция
2.	наказ	функция	функция	функция	приказание
3.	функция	поручение	поручение	поручение	поручение
4.	поручение	наказ	задание	задание	задание
5.	задание	задание	наказ	наказ	наказ
6.	приказ	предписание	обязанность	гидроизоляция	гидроизоляция
7.	упражнение	обязанность	предписание	обязанность	обязанность
8.	заказ	заказ	заказ	предписание	предписание
9.	указание	приказ	обещание	заказ	заказ
10.	предписание	обещание	приказ	обязательство	обязательство
11.	роль	указание	обязательство	обещание	обещание
12.	директива	обязательство	упражнение	приказ	приказ
13.	обязанность	упражнение	указание	упражнение	упражнение
14.	рейс	директива	миссия	указание	указание
15.	обещание	миссия	предназначение	миссия	миссия
16.	инструкция	трюк	долг	предназначение	предназначение

17.	завет	распоряжение	директива	директива	долг
18.	распоряжение	роль	трюк	долг	директива
19.	маневр	долг	роль	трюк	трюк
20.	обязательство	завет	распоряжение	каприз	распоряжение
21.	требование	предназначение	пожелание	распоряжение	роль
22.	задача	пожелание	каприз	роль	каприз
23.	пожелание	задача	завет	задача	задача
24.	полёт	рейс	задача	пожелание	пожелание
25.	заповедь	инструкция	рейс	завет	завет
26.	норма	каприз	инструкция	рейс	рейс
27.	просьба	требование	просьба	инструкция	инструкция
28.	воля	заповедь	требование	требование	присяга
29.	рекомендация	просьба	бросок	просьба	просьба
30.	план	маневр	рекомендация	бросок	требование
31.	блок	рекомендация	заповедь	рекомендация	бросок
32.	посадка	воля	маневр	заповедь	рекомендация
33.	работа	полёт	обряд	маневр	заповедь
34.	долг	обряд	ритуал	обряд	маневр
35.	удар	ритуал	воля	ритуал	обряд
36.	операция	норма	полёт	воля	ритуал
37.	команда	работа	норма	монтаж	воля
38.	условие	перевозка	работа	полёт	норма
39.	танец	план	монтаж	норма	монтаж
40.	поворот	стрельба	посадка	работа	полёт
41.	множество	прыжок	перевозка	перевозка	работа
42.	фигура	заявка	план	посадка	перевозка
43.	программа	посадка	операция	план	посадка
44.	ряд	операция	назначение	операция	план
45.	решение	блок	классификация	классификация	операция

46.	правило	команда	стрельба	прыжок	классификация
47.	расчёт	спектр	команда	назначение	прыжок
48.	желание	назначение	прыжок	стрельба	стрельба
49.	контроль	программа	заявка	команда	назначение
50.	закон	процедура	спектр	заявка	команда
51.	движение	ремонт	блок	спектр	заявка
52.	действие	удар	процедура	блок	процедура
53.	-	установка	желание	процедура	спектр
54.	-	условие	программа	действие	блок
55.	-	желание	удар	желание	программа
56.	-	фигура	действие	программа	действие
57.	-	контракт	ремонт	удар	желание
58.	-	действие	установка	ремонт	удар
59.	-	объём	условие	услуга	ремонт
60.	-	танец	правило	условие	установка
61.	-	правило	комплекс	установка	услуга
62.	-	расчёт	фигура	обслуживание	условие
63.	-	поворот	объём	приём	срок
64.	-	множество	расчёт	правило	правило
65.	-	контроль	контракт	комплекс	обслуживание
66.	-	решение	решение	фигура	приём
67.	-	комплекс	танец	объём	комплекс
68.	-	приём	приём	расчёт	фигура
69.	-	ряд	поворот	контракт	объём
70.	-	исследование	множество	решение	движение
71.	-	закон	контроль	контроль	расчёт
72.	-	движение	движение	движение	контракт
73.	-	проект	ряд	танец	решение
74.	-	разработка	урок	поворот	обработка

75.	-	анализ	закон	множество	контроль
76.	-	шаг	исследование	разработка	исследование
77.	-	совет	срок	срок	танец
78.	-	часть	проект	ряд	поворот
79.	-	число	услуга	исследование	мероприятие
80.	-	вид	разработка	мероприятие	множество
81.	-	дело	анализ	урок	разработка
82.	-	-	шаг	закон	ряд
83.	-	-	совет	соглашение	течение
84.	-	-	договор	течение	урок
85.	-	-	течение	проект	закон
86.	-	-	счёт	договор	соглашение
87.	-	-	часть	шаг	шаг
88.	-	-	процесс	счёт	проект
89.	-	-	число	анализ	десяток
90.	-	-	дело	предложение	предложение
91.	-	-	вид	совет	счёт
92.	-	-	-	очередь	договор
93.	-	-	-	деятельность	совет
94.	-	-	-	часть	анализ
95.	-	-	-	момент	очередь
96.	-	-	-	организация	деятельность
97.	-	-	-	цель	положение
98.	-	-	-	прогресс	момент
99.	-	-	-	дело	часть
100.	-	-	-	сторона	сторона
101.	-	-	-	время	время

Results of the second experiment show that the extension of a span size leads to the decrease of precision and F-measure whereas recall increases and reaches its maximum value for the spans [0, 4] and [0, 5] (see Fig. 4).

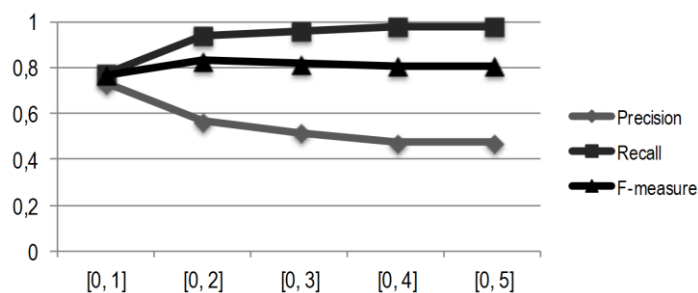


Fig. 4. Evaluation of the results of verb-noun collocation extraction for the verb *выполнять* ('to carry out')

The experimental data we have received allow us to make the following observations. On the one hand, the extension of a span size enables the extraction of distant collocations (collocations in which components are separated by other words). Below there are some examples of distant collocations for the verb *выполнять* ('to carry out') extracted from the corpus RNC2010-MOCKY for different spans.

Span = [0, 2].

...он продолжает работать, **выполняя основные функции**, которые не требуют большого потребления...

...выяснение, насколько Америка способна **выполнять эту роль** гипердержавы. И там нет уверенности...

Span = [0, 3].

...лётчики и подводники, **выполняющие особые правительственные задания**, офицеры военной разведки...

...Его ум позволяет радоваться, **выполняя приказы и указания**. Наверно, это круто...

Span = [0, 4].

...в приоритетном порядке **выполнять взятые по конвенции обязательства**. Президент США...

...постоянно **выполняло и перевыполняло государственный план**, занимая первое место...

Span = [0, 5].

...Ваша задача научиться **выполнять любые атакующие и защитные движения** так, будто они уже достигли...

...посетителей на стенде «Линии График», **выполнявших какие-то только им понятные действия**, привлекало...

On the other hand, the extension of a span size without taking into consideration syntactic relations between words leads to the extraction of erroneous collocates. As examples we have cited the main "false" collocations with the verb *выполнять* ('to carry out') extracted from the corpus RNC2010-MOCKY.

Span = [0, 3].

...Старайтесь с первых же шагов в обучении **выполнять предоставленные методы действия** без предварительной подготовки...

*In this example the collocation **выполнять действия** is wrong: the noun **действия** is related to the noun **методы**, not to the verb **выполнять**.*

Span = [0, 4].

...Свою работу эксперты наши **выполняют добросовестно**, а дальнейшие **действия** - кто в дальнейшем будет манипулировать...

For the verb **выполняют** IntelliText extracts the noun **действия** as a collocate, which is a mistake because this noun refers to the second simple sentence in the compound sentence.

...Можно заставить подчинённых **выполнять работу**, отдав соответствующее **распоряжение**, но такой процесс...

The collocation **выполнять распоряжение** is erroneous because the noun **распоряжение** is a part of the dangling participle **отдав соответствующее распоряжение** and is syntactically related to the adverbial participle **отдав**.

Experiments on verb-noun collocation extraction for the verbs *нарушать* ('to violate') and *принимать* ('to accept') have shown similar results.

In order to solve the problem of erroneous collocates we plan to make a program for collocation extraction which would combine statistical approach (association measures) and syntactic means (taking into account syntactic relations between collocation components).

Conclusion

Our experiments enable us to draw the following conclusions:

1. The quality of verb-noun collocation extraction with POS-filtering is higher than without POS-filtering.
2. In general, the extension of a span size with POS-filtering raises recall, but lowers precision and F-measure of verb-noun collocation extraction.
3. The extension of a span size with POS-filtering has an ambiguous effect. On the one hand, this method makes it possible to extract distant collocates, but on the other hand, it leads to errors in verb-noun collocation extraction, which pays therefore to consider not only the part of speech of a collocate, but syntactic relations between components of a collocation as well.
4. In both experiments one can see that MI is an effective association and ranking measure for verb-noun collocation extraction. Ranks of collocate candidates speak for their relation to collocates in real speech: if a collocate candidate has a high rank, it occurs more often as a component of widely used collocates; if a collocate candidate has a low rank, it is usually a part of free word phrases.

In conclusion, our analysis and comparison of methods of automatic verb-noun collocation extraction allow us to state that a solution to this problem requires a complex syntax-based approach, in other words, it demands a combination of statistical and syntactic methods of verb-noun collocation extraction.

Bibliography

- [Pečina, 2009] Pečina P. Lexical Association Measures: Collocation Extraction. Praha, 2009.
- [Church, 1991] Church, K., Gale, W., Hanks, P. and Hindle, D. 1991. Using Statistics in Lexical Analysis. Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon. New Jersey, Lawrence Erlbaum. P. 115-164.
- [Evert, 2004] Evert, S. 2004. The Statistics of Word Cooccurrences: Word Pairs and Collocations. Ph.D. thesis, University of Stuttgart.
- [Slovar' sochetaemosti slov russkogo jazyka, 1983] Slovar' sochetaemosti slov russkogo jazyka. P.N. Denisov, V.V. Morkovkin (ed.). M., 1983.
- [Deribas, 1983] Deribas V.M. Ustojchivye glagol'no-imennye slovosochetaniya russkogo jazyka. M., 1983.

[Slovar' russkogo jazyka, 1981 – 1984] Slovar' russkogo jazyka: V 4 t. (MAS). A.P. Evgen'eva (ed.). M.: Russkij jazyk, 1981–1984.

Authors' Information



Svetlana Koshcheeva – Student, Saint-Petersburg State University, Universitetskaya emb., 11, Saint-Petersburg 199034 Russia; e-mail: swetik-1995@mail.ru

Major Fields of Scientific Research: Corpus linguistics



Victor Zakharov – Associate Professor, Saint-Petersburg State University, Universitetskaya emb., 11, Saint-Petersburg 199034 Russia; e-mail: vz1311@yandex.ru

Major Fields of Scientific Research: Corpus linguistics, Computational Lexicography