Академия управления

при Президенте Республики Беларусь

Шибут М.С., Яковишин В.С.

РАСПОЗНАВАНИЕ ГРАММАТИЧЕСКОЙ ИНФОРМАЦИИ В ПРОЦЕССЕ ФОРМИРОВАНИЯ ЛИНГВИСТИЧЕСКИХ ЗНАНИЙ

Задача оперативного формирования специализированных баз знаний

Специализированные базы знаний:

- ограниченные некоторой предметной областью
- полученные на основе обработки определенного вида электронных ресурсов – материалов форумов, блогов …

Особенность – индивидуализированные языки участников диалогов:

- использование ненормативной и эрративной ненормативной лексики (языков в стиле эрратива)

В таких условиях использование статичных лингвистических знаний (универсального грамматического словаря) затрудняется

1. Словарь окончаний - 1

Окончания	Типы словоизменения
а	f_1-f_4, m_1–m_7, n_1-n_4
ам	f_1–f_5, m_1–m_7, n_1-n_4
ами	f_1-f_5, m_1–m_7, n_1-n_4
ат	c_7
ах	f_1-f_5, m_1–m_7, n_1-n_4
ая	a_1–a_6
е	f_1-f_4, f_8, f_9, m_1-m_{11}, n_1-n_3, n_5–n_7
его	a_3, a_6, a_7
ее	a_3, a_6, a_7
ей	a_3, a_6, a_7
ем	a_3, a_6, a_7, c_1–c_6, m_5–m_{12}, n_3-n_8
ему	a_3, a_6, a_7
ет	c_1–c_6
…	… …

- f_1-f_7 (f_1,…, f_7) – склонение существительных женского рода;
- m_1–m_{12} (m_1,…, m_{12}) - склонение существительных мужского рода;
- n_1- n_8 (n_1,…, n_8) - склонение существительных среднего рода;
- a_1–a_7 (a_1,…, a_7) – адъективное склонение;
- c_1–c_{12} (c_1,…, c_{12}) – спряжение

2. Алгоритм выявления коррелятивных пар

1. $i := 1$;
2. $j := i+1$;
3. *если* $a_i = s\text{-}P$, *шаг 4; иначе шаг* 5;
4. *если* $s = I(a_j)$, *шаг* 8; *иначе - шаг* 11;
5. *выделение предполагаемой основы* $I(a_i, a_j)$ *и окончаний* $F(a_i)$ *и* $F(a_j)$;
6. *если* $|I(a_i, a_j)| \geq 2$; $|F(a_i)|, |F(a_j)| \leq 3$, *шаг 7; иначе – шаг* 11;
7. *если для* $F(a_i)$ *получено множество парадигм* $P(a_i)$, *шаг* 8; *иначе - шаг* 11;
8. *если для* $F(a_j)$ *получено множество парадигм* $P(a_j)$, *шаг* 9; *иначе - шаг* 11;
9. *если* $P_i \cap P_j \neq \varnothing$, *шаг 10; иначе шаг* 11;
10. $a_i := s\text{-}P$; *удалить элемент* a_j;
11. $j := j+1$;
12. *если* $j \leq n$, *шаг* 3; *иначе шаг 13*;
13. $i := i+1$;
14. *если* $j \leq n$, *шаг* 2; *иначе КОНЕЦ работы алгоритма.*

$a_1, \ldots, a_n$ - массив словоформ

$I(a_i, a_j)$ – предполагаемая общая основа;

$F(a_i)$ и $F(a_j)$ – предполагаемые окончания словоформ

$P(a_i)$ и $P(a_j)$ – множества парадигм словоизменения

$s\text{-}P$ – общая основа s, снабженная множеством общих типов парадигм P

3. Преобразование

a_i = *стола*,
a_j = *столов*,

$P(a_i) = \{f_1\text{-}f_4, m_1\text{–}m_7, n_1\text{-}n_4\}$
$P(a_j) = \{m_1\text{-}m_3\}$;

$P(a_i) \cap P(a_j) = \{m_1, m_3, m_3\}$

стола		*стол-m2*
столбец		*столбец*
столетий		*столетии-*n_7, n_8
столетию	$\Rightarrow$	$\varnothing$
столов		$\varnothing$
столового		*столов-*a_2
столовый		$\varnothing$
столы		$\varnothing$

4. Выявление согласования

***a*(f_1-f_4)→1.113**

окончание –*а*, используемое в типах женского склонения f_1-f_4, заменяется значениями «существительное», «именительный падеж», «единственное число», «женский род» (*карта, книга*);

***a*(m_1, m_6)→1.211/120**

окончание –*а*, используемое в типах мужского склонения m_1, m_6, заменяется значениями «родительный падеж», «единственное число», «мужской род» (*дома, мастера, сторожа*) или значениями «именительный падеж», «множественное число» (*домá, мастерá, сторожá*)

Правила согласования:

3.120.X_1 и 1.120.X_2

большие дома, новые книги

3.51.X_1 и 1.511.X_2

большим домом, новым столом

5. Алгоритм распознавания значений управления

1. морфологический анализ (кодирование значений согласования);
2. выявление и свертывание синтагм, основанных на согласовании;
3. выделение в оставшихся словосочетаниях гипотетических синтагм;
4. верификация гипотетических синтагм и пополнение списка новыми элементами;
5. обработка списка синтагм (кодирование значений управления).

5. Алгоритм распознавания значений управления

Писатель подарил ученикам новую книгу с дарственной надписью

1.111.ПИСАТЕЛ- 2.111.ПОДАРИ- 1.320.УЧЕНИК- 3.413.НОВ- 1.413.КНИГ- С 3.513.ДАРСТВЕНН-1.513.НАДПИС-

ПОДАРИ- 3.УЧЕНИК- 4.КНИГ- С5.НАДПИС-

1.111.ПИСАТЕЛ- 2.111.ПОДАРИ,
3.413.НОВ- 1.413.КНИГ,
3.513.ДАРСТВЕНН-1.513.НАДПИС-;

ПОДАРИ- 3.УЧЕНИК-,
ПОДАРИ- 4.КНИГ-,
ПОДАРИ- С5.НАДПИС-,
УЧЕНИК- 4.КНИГ-,
КНИГ- С5.НАДПИС-

ПОДАРИ- 4.КНИГ- **ПОДАРИ- 3.УЧЕНИК-** **КНИГ- С5.НАДПИС-**	⇒	**ПОДАРИ- 3 / 4** **КНИГ- С5,**

ПОДАРИ- 4.КНИГ-,
КНИГ- С5.НАДПИС-

Академия управления

при Президенте Республики Беларусь

Шибут М.С., Яковишин В.С.

РАСПОЗНАВАНИЕ ГРАММАТИЧЕСКОЙ ИНФОРМАЦИИ В ПРОЦЕССЕ ФОРМИРОВАНИЯ ЛИНГВИСТИЧЕСКИХ ЗНАНИЙ

Спасибо за внимание

(+375 17) 220-48-46

m_shi@tut.by

7. Типы управления прилагательного

Тип управ-ления	Объекты		Примеры
	o_1	o_2	
1	2	3	4
*q*1	1.*G*□		*Достойный уважения*
*q*2	1.*D*□	в 1.*P*□	*Покорный судьбе*
*q*3	1.*I*□		*Сильный духом*
*q*4		на 1.*A*□	*Лицом похож на мать*
*q*5	в 1.*P*□		*Сильный в арифметике*
*q*6	для 1.*G*□		*Интересный для всех*
*q*7	до 1.*G*□		*Жадный до денег*
*q*8	к 1.*D*□		*Склонный к размышлениям*
*q*9	перед 1.*G*□		*Беззащитный перед судьбой*
*q*10	с 1.*I*□		*Почтительный со всеми*

7. Типы управления существительного

Тип управ-ления	Объекты		Примеры
	o_1	o_2	
*s*1	в 1.*A*□		*Превращение в сахар*
*s*2	1.*I*□		*Заражение энтузиазмом*
*s*3	в 1.*P*□		*Превосходство в силе*
*s*4	за 1.*A*□		*Оплата за труд*
*s*5	к 1.*D*□		*Готовность к подвигу*
*s*6	на 1.*A*□		*Реакция на свет*
*s*7	перед 1.*I*□		*Восхищение перед картиной*
*s*8	против1.*G*□		*Борьба против войны*
*s*9	с 1.*I*□		*Интервью с писателем*
*s*10	1.*D*□		*Гимн труду; памятник герою*
*s*11		в 1.*P*□	*Помощь отцу в работе*
*s*12	между 1.*I*□		*Дружба между народами*
*s*13	на 1.*A*□		*Материал на костюм*
*s*14	над 1.*I*□		*Победа над противником*
*s*15	о 1.*P*□		*Отзыв о работе*

7. Типы управления глагола

Тип управ-ления	Объекты		Примеры
	o_1	o_2	
*v*1	1.*G*□		*Добиваться успеха*
*v*2	1.*D*□		*Радоваться удаче*
*v*3		1.*I*□	*Грозить врагу войной*
*v*4		в 1.*P*□	*Подражать мастеру в работе*
*v*5		за 1.*A*□	*Заплатить автору за статью*
*v*6	1.*A*□		*Видеть дом*
*v*7		1.*G*□	*Удостоить работу внимания*
*v*8		1.*D*□	*Вручить награды героям*
*v*9		1.*I*□	*Наполнить стакан водой*
*v*10		в 1.*A*□	*Ввернуть лампочку в патрон*
*v*11		в 1.*P*□	*Изобличить свидетеля во лжи*
*v*12		за 1.*A*□	*Благодарить коллег за помощь*
*v*13		из 1.*G*□	*Извлечь корень из числа*
*v*14		к 1.*G*□	*Прикрепить ручку к двери*

7. Типы управления глагола-2

Тип управ-ления	Объекты		Примеры
	o_1	o_2	
*v*15	1.*A*□	на 1.*A*□	*Вдохновить народ на подвиг*
*v*16		о 1.*P*□	*Знать правду о событиях*
*v*17		от 1.*G*□	*Беречь имущество от пожара*
*v*18		с 1.*G*□	*Взыскать налоги с населения*
*v*19		с 1.*I*□	*Поздравлять воинов с наградой*
*v*20		у 1.*G*□	*Просить внимания у слушателя*
*v*21	1.*I*□		*Руководить работой*
*v*22	без 1.*G*□		*Обходиться без помощи*
*v*23	в 1.*A*□		*Верить в удачу*
*v*24	за 1.*A*□		*Радоваться за товарища*
*v*25	за 1.*I*□		*Наблюдать за событиями*
*v*26	к 1.*D*□		*Готовиться к подвигу*
*v*27	на 1.*A*□		*Сердиться на соседей*
*v*28	над 1.*I*□		*Властвовать над миром*
*v*29	о 1.*A*□		*Споткнуться о порог*

7. Типы управления глагола-3

Тип управ-ления	Объекты		Примеры
	o_1	o_2	
*v*30	о 1.*P*□		*Заботиться о детях*
*v*31	от 1.*G*□		*Воздержаться от суждений*
*v*32	перед 1.*I*□		*Благоговеть перед красотой*
*v*33	по 1.*D*□		*Тосковать по родине*
*v*34	против 1.*G*□		*Восстать против зла*
*v*35	с 1.*P*□		*Воевать с пережитками*