

СПОСОБ РЕКОНСТРУИРОВАНИЯ ПРЕДЛОЖЕНИЙ ЕСТЕСТВЕННОГО ЯЗЫКА ПО ГРАММАТИЧЕСКИМ ПРИЗНАКАМ

Предложен подход к реконструированию предложений естественного языка в Прологе с использованием специализированного словаря, включающего допустимые формы частей речи с набором грамматических признаков. Способ реконструирования ориентирован на системы анализа флективных языков. В частности, приводимый пример демонстрирует обработку предложений для русского языка.

Approach to reconstructing of natural language sentences in Prolog using specialized dictionary that include possible form of words with set of grammatical signs is offered. The reconstructing method is oriented to systems for analysis of inflexional languages. In particular, made an example demonstrates treatment of Russian sentences.

Останов логической регрессии в DCG

В стандартном Прологе реконструирование предложений естественного языка, определяемого формализмом DCG (Definite Clause Grammar), осуществляется простым вызовом обратных действий (далее – *логическая регрессия*) решения задачи контекстного анализа. Такой контекстный анализ предполагает непосредственную проверку словарной формы текущей лексемы на допустимость в предложении, что является приемлемым решением для аналитических языков, имеющих ограниченное количество словарных форм. Иначе обстоит дело с флективными языками, включающих большое разнообразие словарных форм для основных частей речи (существительных, прилагательных и глаголов), что делает неприемлемым эффективный контекстный анализ на основе существующих решений DCG.

В [1] была рассмотрена упрощенная модель контекстов лексем на основе типовой декларативно-процедурной семантики Пролога и продемонстрирован пример обработки согласованных предложений при помощи *смысловых агентов*, внедряемых в систему грамматического разбора DCG для русского языка. Эта контекстная модель содержит *базовые формы* лексем, объединенных в допустимые *смысловые отношения*, а остальные формы тех же лексем вместе со своими грамматическими признаками помещены в специализированный словарь. Применение смысловых отношений позволяет компактно решать задачу контекстного анализа предложений с предусмотренным порядком следования частей речи (как это принято в DCG) на основе единой грамматической сигнатуры. Но, в этом случае возникает проблема реконструирования предложений, для

которой, в силу особенностей данного подхода к контекстному анализу, требуется самостоятельное решение.

Суть проблемы состоит, с одной стороны – в организации специализированного словаря для анализа флективного естественного языка, а с другой – в использовании смысловых отношений, которые хранят лексемы в базовой форме. Тогда, в ходе анализа предложения для проверки допустимости лексемы в текущем контексте она должна быть приведена к своей базовой форме, после чего соответствующий смысловой агент попытается отыскать такую лексему в заданном отношении. И, наоборот, при реструктурировании предложения необходимо найти подходящую лексему в допустимом смысловом отношении, привести к требуемой согласованной форме и подставить в синтезируемый текст.

Таким образом, **целью статьи** является представление результата разработки специального предиката 'возвр_предлож', используемого для эффективного реконструирования предложений русского языка по грамматическим признакам.

Ниже приведен пример из [1], в котором несколько расширен словарь и определен дополнительный предикат 'возвр_предлож' вместе со вспомогательными предикатами. В этом примере задачу преобразования частей речи из согласованной формы в базовую и обратно решают предикаты с функторами 'существит_индекс', 'прилагат_индекс', 'прех_глагол_индекс', 'неперех_глагол_индекс'. В каждом из этих предикатов определено по четыре параметра. Первый из них унифицируется с лексемой в согласованной форме. Этот параметр задается при вызове предиката, когда необходимо выполнить преобразование лек-

семьи к базовой форме. Тогда в качестве второго параметра возвращаются грамматические признаки лексемы, которые извлекаются из словаря, третий параметр игнорируется, а возвращенное значение четвертого параметра представляет собой индекс в списке текущих форм той части речи, к которой относится анализируемая лексема. По этому индексу отыскивается подходящая лексема в списке базовых форм.

%интерпретатор Anzi!Prolog 8.0.6

%проверьте следующие запросы

s1:-'предлож'(['смешная','рыжая','обезьяна','держит','банан'],[]).

s2:-'предлож'(['знаменитый','талантливый','пародист','рассмешил','публику'],[]).

s3:-'предлож'(['легкомысленный','халатный','охранник','прозевал','грабителя'],[]).

s4:-'предлож'(['опасный','вооруженный','гангстер','грабит','банки'],[]).

s5:-'предлож'(['обезьяна','держит','банан'],[]).

s6:-'предлож'(['опасный','вооруженный','гангстер','угрожает','гражданину'],[]).

s7:-'предлож'(['опасный','вооруженный','гангстер','угрожает','гражданин'],[]).

s8:-'предлож'(['опасный','вооруженный','гангстер','грабит','инкассаторов'],[]).

s9:-'предлож'(['легкомысленный','халатный','охранник','читает','книгу'],[]).

%проверьте запросы на реконструирование предложений

rq1(S):-rs(2,'мужск':'единств':'именит'),
'лицо'('третье'):'наст':'несоверш':'действит',
'мужск':'единств':'винит',S).

rq2(S):-rs(2,'мужск':'единств':'именит'),
'лицо'('третье'):'наст':'несоверш':'действит',
'средн':'множеств':'винит',S).

rq3(S):-rs(2,'женск':'единств':'именит'),
'лицо'('третье'):'наст':'несоверш':'действит',
'мужск':'единств':'винит',S).

rq4(S):-rs(1,'женск':'единств':'именит'),
'лицо'('третье'):'наст':'несоверш':'действит',
'мужск':'единств':'винит',S).

rq5(S):-rs(1,'мужск':'единств':'именит'),
'лицо'('третье'):'наст':'несоверш':'действит',
'средн':'множеств':'винит',S).

rq6(S):-rs(2,'мужск':'единств':'именит'),
'лицо'('третье'):'наст':'несоверш':'действит',
'женск':'единств':'винит',S).

rq7(S):-rs(2,'мужск':'единств':'именит'),
'лицо'('третье'):'наст':'несоверш':'действит',
'мужск':'единств':'датель',S).

rq8(S):-rs(2,'мужск':'единств':'именит'),
'лицо'('третье'):'наст':'несоверш':'действит',
'средн':'множеств':'датель',S).

:-op(500,xfu,;).

rs((Dn,G:Q:A),R:T:V:D,Ga:Qa:Aa,S):-Obj=

'объект'(_,_,_),bagof(Obj,Obj,St),

'принадлежать'(Sc,St),

'возвр_предлож'((Dn,G:Q:A),R:T:V:D,Ga:Qa:Aa,S,Sc).

'возвр_предлож'((Nd,G:Q:A),R:T:V:D,Ga:Qa:Aa,S,Sc):-

Sc='объект'(N,'определитель'(Db),'функция'(F)),

'длина_списка'(Db,Dn),(Nd>Dn->fail;

('существит_индекс'(N,_'единств':'именит':C:GBn,_,

In),'существит_индекс'(_,G:Q:A:C:GBn,Kn,In),

'возвр_функц'(F,Fc,Lo),(Aa='винит'->

('перех_глагол_индекс'(Fc,_,_,Iv,CL),

'перех_глагол_индекс'(_,G:Q:R:T:V:D:C,Kv,Iv,CL));

('неперех_глагол_индекс'(Fc,_,_,Iv,CL),

'неперех_глагол_индекс'(_,G:Q:R:T:V:D:C,Kv,Iv,CL))),

'принадлежать'(Ar,Lo),

'существит_индекс'(Ar,_'единств':'именит':Ca:GBa,_,

Ia),'существит_индекс'(_,Ga:Qa:Aa:Ca:GBa,Ka,Ia),

(Nd==0->S=[Kn,Kv,Ka];

('возвр_опред'(Db,G:Q:A:C,[],Ld),(Nd==Dn->

'конкатенация_списков'(Ld,[Kn,Kv,Ka],S);

('принадлежать'(Ld1,Ld),

'конкатенация_списков'([Ld1],[Kn,Kv,Ka],S))))).

'возвр_функц'(F,Fc,Lo):-

'принадлежать'(Fs,F),Fs=..[Fc,Lo].

'возвр_опред'([],_:::_,Ld,Ld):-!.

'возвр_опред'([D|Dt],G:Q:A:C,Dc,Ld):-

'прилагат_индекс'(D,_,_,I),

'прилагат_индекс'(_,G:Q:A,Kd,I),Dg=[Kd|Dc],

'возвр_опред'(Dt,G:Q:A:C,Dg,Ld).

'предлож'->'фраза_существит'(G,Q,'именит',C,F),

'фраза_глагол'(G,Q,_,_,_,C,F).

'фраза_существит'(G,Q,A,C,F)-->[W1],[W2],[W3],

{'прилагат_индекс'(W1,G:Q:A,_,I1),

p>'прилагат_индекс'(W2,G:Q:A,_,I2),

'существит_индекс'(W3,G:Q:A:C:GB,_,I3),

'прилагат_индекс'(_, 'мужск':'единств':'именит',K1,I1),

'прилагат_индекс'(_, 'мужск':'единств':'именит',K2,I2),

'существит_индекс'(_,GB:'единств':'именит':C:_,K3,I3),

'агент_о'([K1,K2],S),'агент_с'(K3,S,F)}.

'фраза_существит'(G,Q,A,C,F)-->[W],

{'существит_индекс'(W,G:Q:A:C:GB,_,I),

'существит_индекс'(_,GB:'единств':'именит':C:_,K,I),

((is_list(F)->'агент_фо'(K,F);Obj='объект'(K,_,_),

bagof(Obj,Obj,S),'агент_с'(K,S,F)))}

'фраза_глагол'(G,Q,_,_,_,C,F)-->[W],

{'перех_глагол_индекс'(W,G:Q:'лицо'('третье')):_:V:

'действит':C,_,I,Lc),

'перех_глагол_индекс'(_, 'неопфм':'неопфм':'неопфм':

'неопфм':V:'неопфм':C,K,I,Lc),'агент_ф'(K,F,Fo)},

'фраза_существит'(_,_, 'винит',_,Fo).

'фраза_глагол'(G,Q,_,_,_,C,F)-->[W],

{'неперех_глагол_индекс'(W,G:Q:'лицо'('третье')):_:V:

'действит':C,_,I,Lc),

'неперех_глагол_индекс'(_, 'неопфм':'неопфм':'неопфм':

'неопфм':V:'неопфм':C,K,I,Lc),'агент_ф'(K,F,Fo)},

'фраза_существит'(_,_, 'датель',_,Fo).

'существит_индекс'(W,G:Q:A:C:GB,K,I):-

atomic(W),atomic(I),!,fail.

'существит_индекс'(W,G:Q:A:C:GB,K,I):-

(atomic(W)->('словарь_существит'(G:Q:A:C:GB,L),

'элемент'('принадлежать'(W,L),'индекс'(W,L,I));

(atomic(I),atomic(G),atomic(Q),atomic(A),

'словарь_существит'(G:Q:A:C:GB,L),

'искать_слово'(K,L,I)),!,

'прилагат_индекс'(W,G:Q:A,K,I):-atomic(W),atomic(I),

!,fail.

'прилагат_индекс'(W,G:Q:A,K,I):-atomic(W)->

('словарь_прилагат'(G:Q:A,L),

'элемент'('принадлежать'(W,L),'индекс'(W,L,I));

(atomic(I),atomic(G),atomic(Q),atomic(A),

'словарь_прилагат'(G:Q:A,L),'искать_слово'(K,L,I)),!,

'местоим_индекс'(W,G:Q:R:A,K,I):-atomic(W),

atomic(I),!,fail.

'местоим_индекс'(W,G:Q:R:A,K,I):-atomic(W)->

('словарь_местоим'(G:Q:R:A,L),

'элемент'('принадлежать'(W,L),'индекс'(W,L,I));

(atomic(I),atomic(G),atomic(Q),atomic(R),atomic(A),

'словарь_местоим'(G:Q:R:A,L),'искать_слово'(K,L,I))),

!.

'перех_глагол_индекс'(W,G:Q:R:T:V:D:C,K,I,Lc):-

atomic(W),atomic(I),!,fail.

```

'перех_глагол_индекс'(W,G:Q:R:T:V:D:C,K,I,Lc):-
(atomic(W)->('перех_глагол'(G:Q:R:T:V:D:C,L,Lc),
'элемент'('принадлежать'(W,L)), 'индекс'(W,L,I));
(atomic(I),atomic(G),atomic(Q),(R=='неопфм'->
true;(R=..[P,Vp],atomic(P),atomic(Vp))),
atomic(T),atomic(V),atomic(D),atomic(C),
'перех_глагол'(G:Q:R:T:V:D:C,L,Lc),
'искать_слово'(K,L,I)),!.
'неперех_глагол_индекс'(W,G:Q:R:T:V:D:C,K,I,Lc):-
atomic(W),atomic(I),!,fail.
'неперех_глагол_индекс'(W,G:Q:R:T:V:D:C,K,I,Lc):-
(atomic(W)->('неперех_глагол'(G:Q:R:T:V:D:C,L,Lc),
'элемент'('принадлежать'(W,L)), 'индекс'(W,L,I));
(atomic(I),atomic(G),atomic(Q),(R=='неопфм'-> true;
(R=..[P,Vp],atomic(P),atomic(Vp))),
atomic(T),atomic(V),atomic(D),atomic(C),
'неперех_глагол'(G:Q:R:T:V:D:C,L,Lc),
'искать_слово'(K,L,I)),!.
'длина_списка'([],0):-!.
'длина_списка'([X|Xs],Ln):- 'длина_списка'(Xs,L1),
Ln is L1+1.
'конкатенация_списков'([],L,L).
'конкатенация_списков'([X|L1],L2,[X|L3]):-
'конкатенация_списков'(L1,L2,L3).
'элемент'(P):-P,!.
'подмножество'(L,[]).
'подмножество'(L,[X|T]):-
'элемент'('принадлежать'(X,L)),
'подмножество'(L,T).
'принадлежать'(X,[X|_]).
'принадлежать'(X,[_|T]):-'принадлежать'(X,T).
'индекс'(X,[ ],I):-I is 0.
'индекс'(X,[X|T],I):-!,I is 1.
'индекс'(X,[_|T],I):- 'индекс'(X,T,Ic),
(Ic==0->I is Ic;I is Ic+1).
'искать_слово'(X,[_|_],0):-!,fail.
'искать_слово'(X,[X|_],1).
'искать_слово'(X,[_|T],I):-Ic is I-1,
'искать_слово'(X,T,Ic).
'агент_о'([K1,K2],S):-
Obj='объект'(N,'определитель'(L),F),
bagof(Obj,(Obj,'подмножество'(L,[K1,K2])),S).
'агент_с'(Ls,[],):-!,fail.
'агент_с'(Ls,S,F):-'функция_субъект'(Ls,S,F).
'функция_субъект'(Ls,[],[]).
'функция_субъект'(Ls,['объект'(Ls,_,'функция'(F)|_),
F).
'функция_субъект'(_,['объект'(_,_, 'функция'(_)|St],
_):-'функция_субъект'(Ls,St,F).
'агент_ф'(Lv,F,Fo):-'аргум_функция_субъект'(Lv,F,Fo).
'аргум_функция_субъект'(Lv,[],):-!,fail.
'аргум_функция_субъект'(Lv,Lf,Fo):-Lf=[Lf|Lft],
Lfn=..[Fn,Fco],(Lv=Fn->Fo=Fco;
'аргум_функция_субъект'(Lv,Lft,Fo)).
'агент_фо'(Lo,Fo):-'элемент'('принадлежать'(Lo,Fo)).
'объект'('гангстер','определитель'(['вооруженный',
'опасный']), 'функция'(['грабить'(['банк','инкассатор',
'казино','склад']), 'угрожать'(['гражданин'])]).
'объект'('обезьяна','определитель'(['смешной',
'рыжий']), 'функция'(['держать'(['банан','палка'])]).
'объект'('пародист','определитель'(['знаменитый',
'талантливый'])),
'функция'(['рассмешить'(['публика'])]).
'объект'('охранник','определитель'(['легкомысленный',
'халатный'])), 'функция'(['прозевать'(['грабитель'])),
'читать'(['книга'])).
'словарь_существит'('мужск': 'единств': 'именит': 'люди':
'мужск', ['гангстер', 'грабитель', 'гражданин',
'инкассатор', 'охранник', 'пародист']).
'словарь_существит'('мужск': 'единств': 'родит': 'люди':
'мужск', ['гангстера', 'грабителя', 'гражданина',
'инкассатора', 'охранника', 'пародиста']).
'словарь_существит'('мужск': 'единств': 'датель': 'люди':
'мужск', ['гангстеру', 'грабителю', 'гражданину',
'инкассатору', 'охраннику', 'пародисту']).
'словарь_существит'('мужск': 'единств': 'винит': 'люди':
'мужск', ['гангстера', 'грабителя', 'гражданина',
'инкассатора', 'охранника', 'пародиста']).
'словарь_существит'('мужск': 'единств': 'творит': 'люди':
'мужск', ['гангстером', 'грабителем', 'гражданином',
'инкассатором', 'охранником', 'пародистом']).
'словарь_существит'('мужск': 'единств': 'предлож':
'люди': 'мужск', ['гангстере', 'грабителе', 'гражданине',
'инкассаторе', 'охраннике', 'пародисте']).
'словарь_существит'('средн': 'множеств': 'именит':
'люди': 'мужск', ['гангстеры', 'грабители', 'граждане',
'инкассаторы', 'охранники', 'пародисты']).
'словарь_существит'('средн': 'множеств': 'родит': 'люди':
'мужск', ['гангстеров', 'грабителей', 'граждан',
'инкассаторов', 'охранников', 'пародистов']).
'словарь_существит'('средн': 'множеств': 'датель': 'люди':
'мужск', ['гангстерам', 'грабителям', 'гражданам',
'инкассаторам', 'охранникам', 'пародистам']).
'словарь_существит'('средн': 'множеств': 'винит': 'люди':
'мужск', ['гангстеров', 'грабителей', 'граждан',
'инкассаторов', 'охранников', 'пародистов']).
'словарь_существит'('средн': 'множеств': 'творит':
'люди': 'мужск', ['гангстерами', 'грабителями',
'гражданами', 'инкассаторами', 'охранниками',
'пародистами']).
'словарь_существит'('средн': 'множеств': 'предлож':
'люди': 'мужск', ['гангстерах', 'грабителях', 'гражданах',
'инкассаторах', 'охранниках', 'пародистах']).
'словарь_существит'('женск': 'единств': 'именит': 'люди':
'женск', ['публика']).
'словарь_существит'('женск': 'единств': 'родит': 'люди':
'женск', ['публики']).
'словарь_существит'('женск': 'единств': 'датель': 'люди':
'женск', ['публике']).
'словарь_существит'('женск': 'единств': 'винит': 'люди':
'женск', ['публику']).
'словарь_существит'('женск': 'единств': 'творит': 'люди':
'женск', ['публикой']).
'словарь_существит'('женск': 'единств': 'предлож':
'люди': 'женск', ['публике']).
'словарь_существит'('женск': 'единств': 'именит':
'животные': 'женск', ['обезьяна']).
'словарь_существит'('женск': 'единств': 'родит':
'животные': 'женск', ['обезьяны']).
'словарь_существит'('женск': 'единств': 'датель':
'животные': 'женск', ['обезьяне']).
'словарь_существит'('женск': 'единств': 'винит':
'животные': 'женск', ['обезьяну']).
'словарь_существит'('женск': 'единств': 'творит':
'животные': 'женск', ['обезьяной']).
'словарь_существит'('женск': 'единств': 'предлож':
'животные': 'женск', ['обезьяне']).
'словарь_существит'('мужск': 'единств': 'именит':
'растения': 'мужск', ['банан']).
'словарь_существит'('мужск': 'единств': 'родит':
'растения': 'мужск', ['банана']).
'словарь_существит'('мужск': 'единств': 'датель':
'растения': 'мужск', ['банану']).
'словарь_существит'('мужск': 'единств': 'винит':
'растения': 'мужск', ['банан']).
'словарь_существит'('мужск': 'единств': 'творит':
'растения': 'мужск', ['бананом']).
'словарь_существит'('мужск': 'единств':

```

'предлож': 'растения': 'мужск', ['банане']).	'словарь_прилагат' ('мужск': 'единств': 'дател',
'словарь_существит' ('мужск': 'единств': 'именит':	['вооруженному', 'знаменитому', 'легкомысленному',
'учереждения': 'мужск', ['банк', 'склад']).	'опасному', 'рыжему', 'смешному', 'талантливому',
'словарь_существит' ('мужск': 'единств': 'родит':	'халатному']).
'учереждения': 'мужск', ['банка', 'склада']).	'словарь_прилагат' ('мужск': 'единств': 'винит',
'словарь_существит' ('мужск': 'единств': 'дател':	['вооруженного', 'знаменитого', 'легкомысленного',
'учереждения': 'мужск', ['банку', 'складу']).	'опасного', 'рыжего', 'смешного', 'талантливое',
'словарь_существит' ('мужск': 'единств': 'винит':	'халатного']).
'учереждения': 'мужск', ['банк', 'склад']).	'словарь_прилагат' ('мужск': 'единств': 'творит',
'словарь_существит' ('мужск': 'единств': 'творит':	['вооруженным', 'знаменитым', 'легкомысленным',
'учереждения': 'мужск', ['банком', 'складом']).	'опасным', 'рыжим', 'смешным', 'талантливым',
'словарь_существит' ('мужск': 'единств': 'предлож':	'халатным']).
'учереждения': 'мужск', ['банке', 'складе']).	'словарь_прилагат' ('мужск': 'единств': 'предлож',
'словарь_существит' ('средн': 'множеств': 'именит':	['вооруженном', 'знаменитом', 'легкомысленном',
'учереждения': 'мужск', ['банки', 'склады']).	'опасном', 'рыжем', 'смешном', 'талантливым',
'словарь_существит' ('средн': 'множеств': 'родит':	'халатном']).
'учереждения': 'мужск', ['банков', 'складов']).	'словарь_прилагат' ('женск': 'единств': 'именит',
'словарь_существит' ('средн': 'множеств': 'дател':	['вооруженная', 'знаменитая', 'легкомысленная',
'учереждения': 'мужск', ['банкам', 'складам']).	'опасная', 'рыжая', 'смешная', 'талантливая',
'словарь_существит' ('средн': 'множеств': 'винит':	'халатная']).
'учереждения': 'мужск', ['банки', 'склады']).	'словарь_прилагат' ('женск': 'единств': 'родит',
'словарь_существит' ('средн': 'множеств': 'творит':	['вооруженную', 'знаменитую', 'легкомысленную',
'учереждения': 'мужск', ['банками', 'складами']).	'опасную', 'рыжую', 'смешную', 'талантливую',
'словарь_существит' ('средн': 'множеств': 'предлож':	'халатную']).
'учереждения': 'мужск', ['банках', 'складах']).	'словарь_прилагат' ('женск': 'единств': 'дател',
'словарь_существит' ('средн': 'единств': 'именит':	['вооруженной', 'знаменитой', 'легкомысленной',
'учереждения': 'средн', ['казино']).	'опасной', 'рыжей', 'смешной', 'талантливой',
'словарь_существит' ('средн': 'единств': 'родит':	'халатной']).
'учереждения': 'средн', ['казино']).	'словарь_прилагат' ('женск': 'единств': 'винит',
'словарь_существит' ('средн': 'единств': 'дател':	['вооруженную', 'знаменитую', 'легкомысленную',
'учереждения': 'средн', ['казино']).	'опасную', 'рыжую', 'смешную', 'талантливую',
'словарь_существит' ('средн': 'единств': 'винит':	'халатную']).
'учереждения': 'средн', ['казино']).	'словарь_прилагат' ('женск': 'единств': 'творит',
'словарь_существит' ('средн': 'единств': 'творит':	['вооруженной', 'знаменитой', 'легкомысленной',
'учереждения': 'средн', ['казино']).	'опасной', 'рыжей', 'смешной', 'талантливой',
'словарь_существит' ('средн': 'единств': 'предлож':	'халатной']).
'учереждения': 'средн', ['казино']).	'словарь_прилагат' ('женск': 'единств': 'предлож',
'словарь_существит' ('средн': 'множеств': 'именит':	['вооруженной', 'знаменитой', 'легкомысленной',
'учереждения': 'средн', ['казино']).	'опасной', 'рыжей', 'смешной', 'талантливой',
'словарь_существит' ('средн': 'множеств': 'родит':	'халатной']).
'учереждения': 'средн', ['казино']).	'перех_глагол' ('неопфм': 'неопфм': 'неопфм': 'неопфм':
'словарь_существит' ('средн': 'множеств': 'дател':	'несоверш': 'неопфм': 'люди', ['грабить', 'читать'],
'учереждения': 'средн', ['казино']).	['люди']).
'словарь_существит' ('средн': 'множеств': 'винит':	'перех_глагол' ('неопфм': 'неопфм': 'неопфм': 'неопфм':
'учереждения': 'средн', ['казино']).	'несоверш': 'неопфм': 'C', ['держать', 'зевать', 'смешить'],
'словарь_существит' ('средн': 'множеств': 'творит':	Lc) :- ['животные', 'люди'] = Lc,
'учереждения': 'средн', ['казино']).	'элемент' ('принадлежать' (C, Lc)).
'словарь_существит' ('средн': 'множеств': 'предлож':	'перех_глагол' ('неопфм': 'неопфм': 'неопфм': 'неопфм':
'учереждения': 'средн', ['казино']).	'соверш': 'неопфм': 'C', ['продержать', 'прозевать',
'словарь_существит' ('женск': 'единств': 'именит':	'рассмешить'], Lc) :- ['животные', 'люди'] = Lc,
'литература': 'женск', ['книга']).	'элемент' ('принадлежать' (C, Lc)).
'словарь_существит' ('женск': 'единств': 'родит':	'неперех_глагол' ('неопфм': 'неопфм': 'неопфм': 'неопфм':
'литература': 'женск', ['книги']).	'несоверш': 'неопфм': 'C', ['утрожать'], Lc) :-
'словарь_существит' ('женск': 'единств': 'дател':	['животные', 'люди'] = Lc,
'литература': 'женск', ['книге']).	'элемент' ('принадлежать' (C, Lc)).
'словарь_существит' ('женск': 'единств': 'винит':	'неперех_глагол' ('неопфм': 'неопфм': 'неопфм': 'неопфм':
'литература': 'женск', ['книгу']).	'соверш': 'неопфм': 'C', ['проутрожать'], Lc) :-
'словарь_существит' ('женск': 'единств': 'творит':	['животные', 'люди'] = Lc,
'литература': 'женск', ['книгой']).	'элемент' ('принадлежать' (C, Lc)).
'словарь_существит' ('женск': 'единств': 'предлож':	'перех_глагол' ('мужск': 'единств': 'лицо' ('третье'):
'литература': 'женск', ['книге']).	'наст': 'несоверш': 'действит': 'люди', ['грабит',
'словарь_прилагат' ('мужск': 'единств': 'именит',	'читает'], ['люди']).
['вооруженный', 'знаменитый', 'легкомысленный',	'перех_глагол' ('мужск': 'единств': 'лицо' ('третье'):
'опасный', 'рыжий', 'смешной', 'талантливый',	'наст': 'несоверш': 'действит': 'C', ['держит', 'зевает',
'халатный']).	'смешит'], Lc) :- ['животные', 'люди'] = Lc,
'словарь_прилагат' ('мужск': 'единств': 'родит',	'элемент' ('принадлежать' (C, Lc)).
['вооруженного', 'знаменитого', 'легкомысленного',	'перех_глагол' ('женск': 'единств': 'лицо' ('третье'):
'опасного', 'рыжего', 'смешного', 'талантливое',	'наст': 'несоверш': 'действит': 'C', ['держит', 'зевает',
'халатного']).	'смешит'], Lc) :- ['животные', 'люди'] = Lc,

'элемент' ('принадлежать' (C, Lc)).
 'перех_глагол' ('мужск': 'единств': 'лицо' ('третье')):
 'прош': 'соверш': 'действит': 'люди', ['ограбил'],
 ['люди']).
 'перех_глагол' ('мужск': 'единств': 'лицо' ('третье')):
 'прош': 'соверш': 'действит': C, ['продержал', 'прозевал',
 'рассмешил'], Lc):- ['животные', 'люди']= Lc,
 'элемент' ('принадлежать' (C, Lc)).
 'перех_глагол' ('женск': 'единств': 'лицо' ('третье')):
 'прош': 'соверш': 'действит': C, ['продержала',
 'прозевала', 'рассмешила', 'пригрозила'], Lc):-
 ['животные', 'люди']= Lc,
 'элемент' ('принадлежать' (C, Lc)).
 'неперех_глагол' ('мужск': 'единств': 'лицо' ('третье')):
 'наст': 'несоверш': 'действит': C, ['угрожает'], Lc):-
 ['животные', 'люди']= Lc,
 'элемент' ('принадлежать' (C, Lc)).
 'неперех_глагол' ('мужск': 'единств': 'лицо' ('третье')):
 'прош': 'соверш': 'действит': C, ['пригрозил'], Lc):-
 ['животные', 'люди']= Lc,
 'элемент' ('принадлежать' (C, Lc)).
 'неперех_глагол' ('женск': 'единств': 'лицо' ('третье')):
 'прош': 'соверш': 'действит': C, ['пригрозила'], Lc):-
 ['животные', 'люди']= Lc,
 'элемент' ('принадлежать' (C, Lc)).

В свою очередь, для преобразования лексем из базовой формы, извлекаемой из смыслового отношения, в необходимую согласованную форму, при вызове упомянутых предикатов первый параметр игнорируется. Во втором параметре задаются грамматические признаки, которым должна соответствовать требуемая части речи, а в четвертом – индекс, который подсчитывается по списку базовых форм. Тогда в качестве третьего параметра возвращается форма части речи, запрошенная в соответствии с грамматическими признаками. Таким образом, в данном случае работа со словарем основана на доступе к требуемой форме части речи по индексу, значение которого равно номеру позиции лексемы в соответствующем списке форм словаря.

Заметим, что при ином подходе к работе со словарем флективного языка, перебор огромного количества сочетаний из различных форм лексем в ходе работы механизма DCG в направлении логической регрессии, неизбежно ведет к крайне неэффективному решению задачи реконструирования предложений. По этой причине в разработанных предикатах преобразования форм предусмотрена проверка наличия значений параметров, содержащих либо форму лексемы, либо индекс. В случае их отсутствия (если запросить реконструирование предложений по правилам DCG в виде свободной переменной и пустого списка), ход логической

регрессии блокируется предикатом анализа типа `atomic`.

Работа примера в соответствии с заданными контекстными моделями дает следующие результаты:

?- rq1(S).

S = [опасный, вооруженный, гангстер, ограбил, банк];
 S = [опасный, вооруженный, гангстер, ограбил, инкассатора];
 S = [опасный, вооруженный, гангстер, ограбил, склад];
 S = [халатный, легкомысленный, охранник, прозевал, грабителя];
 no

?- rq2(S).

S = [опасный, вооруженный, гангстер, грабит, банки];
 S = [опасный, вооруженный, гангстер, грабит, инкассаторов];
 S = [опасный, вооруженный, гангстер, грабит, казино];
 S = [опасный, вооруженный, гангстер, грабит, склады];
 S = [халатный, легкомысленный, охранник, зевает, грабителей];
 no

?- rq3(S).

S = [рыжая, смешная, обезьяна, держит, банан];
 no

?- rq4(S).

S = [рыжая, обезьяна, держит, банан];
 S = [смешная, обезьяна, держит, банан];
 no
 ?- rq5(S).

S = [опасный, гангстер, грабит, банки];
 S = [вооруженный, гангстер, грабит, банки];
 S = [опасный, гангстер, грабит, инкассаторов];
 S = [вооруженный, гангстер, грабит, инкассаторов];
 S = [опасный, гангстер, грабит, казино];
 S = [вооруженный, гангстер, грабит, казино];
 S = [опасный, гангстер, грабит, склады];
 S = [вооруженный, гангстер, грабит, склады];
 S = [халатный, охранник, зевает, грабителей];
 S = [легкомысленный, охранник, зевает, грабителей];
 no

?- rq6(S).

S = [талантливый, знаменитый, пародист, смешит, публику];
 S = [халатный, легкомысленный, охранник, читает, книгу];
 no

?- rq7(S).

S = [опасный, вооруженный, гангстер, угрожает, гражданину];

no
 ?- rq8(S).

S = [опасный, вооруженный, гангстер, угрожает, гражданам];
 no
 ?- quit.

Здесь вспомогательные предикаты с функтором типа `rq1` используются для упрощения запросов. Эти предикаты обеспечивают вызов предиката с функтором `rs`, в котором в качестве параметров принимаются значения грамматических признаков частей

речи. Затем, при помощи встроенного предиката *bagof* в базе знаний собираются все допустимые смысловые отношения, после чего предикат 'возвр_предлож' осуществляет их обработку и синтезирует согласованные предложения.

Особенности применения грамматических признаков

Для вызова предиката 'возвр_предлож' используются пять параметров, первые три из которых представляют собой группы. Т.е. этот предикат задан определением формулы:

'возвр_предлож' ((Dn,G:Q:A),R:T:V:D,Ga:Qa:Aa,S,Sc).

Первая группа параметров в виде структуры без функтора включает следующие признаки:

Dn – количество определителей (прилагательных) в группе существительного DCG;

G:Q:A – род, число и падеж для согласования группы существительного в DCG;

R:T:V:D – лицо, время, вид и залог глагола в DCG;

Ga:Qa:Aa – род, число и падеж существительного в глагольной группе DCG;

S – свободная переменная, используемая для возврата реконструированного предложения;

Sc – переменная, куда при вызове предиката помещен список допустимых смысловых отношений, на основе которых выполняется реконструирование.

По приведенным в предыдущем разделе результатам, в соответствии с запросами, можно наблюдать изменение грамматических признаков в синтезированных и согласованных предложениях. Так, в запросе κ_1 группа существительного выводится с двумя прилагательными единственного числа, в именительном падеже мужского рода, а глагол – в третьем лице, прошедшем времени, совершенном виде и действительном залоге. Для существительного глагольной группы заданы: мужской род, единственное число и винительный падеж, так как глагол 'грабить' применяется в качестве переходного и по правилам русского языка требует существительного в винительном падеже. В запросе κ_2 изменены грамматические признаки глагольной группы. Для глагола используются: третье лицо, настоящее время, несовершенный вид, действительный залог, а для существительного глагольной группы – средний род, множественное число и винительный падеж.

Аналогичные изменения и согласования частей речи можно наблюдать в синтезиро-

ванных предложениях, соответствующих остальным запросам, где заданы другие грамматические признаки. Кроме того, в запросах κ_4 и κ_5 задан вывод одного, а не двух определителей-прилагательных группы существительного, т.е. параметр Dn унифицирован с единицей. В результате – в ходе реконструирования были выведены предложения с одним из возможных определителей группы существительного. Такие предложения синтезированы с учетом однократного предшествования прилагательного во всех вариантах, предусмотренных в смысловых отношениях, что обеспечено работой механизма перебора с возвратом.

Отметим также, что предикат 'возвр_предлож' легко модифицируется. Так, в запросах κ_7 и κ_8 существительное глагольной группы выводится в дательном падеже в соответствии с логикой применения непереходного глагола 'угрожать'. Вывод дательного падежа предусмотрен в правиле, соответствующем предикату 'возвр_предлож'. Подобная гибкая модификация правил позволяет учитывать тонкости языковой грамматики для синтеза предложений на основе грамматических признаков.

Роль смысловых отношений

Смысловые отношения представляют собой основу организации логической связи слов при реконструировании предложений, в отличие от правил грамматики естественного языка, которые обеспечивают только грамматическое согласование. Прежде всего, обратим внимание на различие в содержаниях выводимых предложений, хотя все они могут обрабатываться едиными правилами DCG и в своих смысловых отношениях используют один и тот же шаблон для организации структур. Другими словами, использование ограниченного набора правил DCG и небольшого количества смысловых отношений позволяет существенно влиять на содержание как анализируемых, так и реконструируемых предложений в пределах одной и той же грамматической сигнатуры. Для этого достаточно использовать специализированный словарь с грамматическими признаками частей речи и обеспечить обработку лексем естественного языка с учетом значений этих признаков, как при анализе, так и при реконструировании предложений.

Для лучшего представления о назначении смысловых отношений рассмотрим работу предиката 'возвр_предлож' более детально.

Все смысловые отношения определены в виде структур, которые имеют общий функ-

тор 'объект', одинаковую арность и назначение аргументов [1]. Перед обращением к предикату 'возвр_предлож' в предикате `rs` вводится переменная `obj`, которая унифицируется со структурой 'объект', что обеспечивает создание списка всех таких структур предикатом `bagof` и унификацию этого списка с переменной `st`. Далее, предикатом 'принадлежать' из списка `st` извлекается текущая обрабатываемая структура и унифицируется с переменной `sc`, которая вместе с заданными грамматическими признаками передается предикату 'возвр_предлож'.

Работа предиката 'возвр_предлож' начинается с того, что из структуры `sc` извлекаются элементы реконструируемого предложения, которые представлены в базовой форме:

`n` – существительное (субъект действия);

`db` – список определителей группы существительного;

`f` – список структур, представляющих глагольную группу, в которых функтор соответствует глаголу, а в качестве аргумента представлен список существительных (объектов действия).

Тогда, дальнейшая обработка смыслового отношения сводится к преобразованию лексем из базовой формы в согласованную форму по заданным грамматическими признакам и объединение в готовое предложение.

Заметим, что порядок обработки лексем в данном случае не имеет значения. Так, в

приведенном примере сначала обрабатывается объект-существительное, затем лексемы глагольной группы и после этого идет перебор возможных определителей-прилагательных. Результат представляет собой реконструированное и согласованное предложение в виде списка лексем, который получен предикатом 'конкатенация_списков'.

Выводы

1. Эффективность работы механизма DCG, при стандартном подходе к реконструированию предложений флективного естественного языка, резко снижается из-за большого разнообразия форм основных частей речи.

2. Решение рассмотренной задачи упрощается за счет работы со специализированным словарем, который содержит образцы форм частей речи с грамматическими признаками, а также за счет введения допустимых смысловых отношений.

3. Использование грамматических признаков позволяет гибко изменять согласующиеся формы частей речи, учитывая особенности грамматики естественного языка.

4. Применение смысловых отношений дает возможность управлять содержанием реконструируемых предложений, которые представлены единой грамматической сигнатурой в правилах DCG.

Список литературы

1. Кузнецов А.В. Смысловая аморфность в грамматических сигнатурах предложений естественного языка //Вісник НТУУ «КПІ». Інформатика, управління та обчислювальна техніка: Зб. наук. пр. - К.: Век+. - 2007. - №47.- С.242-254.).

Поступила в редакцию 17.12.2009