

Dialog 2010



SECRET PUNCTUATION MARKS

Okatyev Vladimir

Erehinskaya Tatyana

Ratanova Tatyana

$$a^*(b-(c-(d+e)))$$

Вот два петуха,
Которые будят того пастуха,
Который бранится с коровницей строгою,
Которая доит корову безрогую,
Лягнувшую старого пса без хвоста,
Который за шиворот треплет кота,
Который пугает и ловит синицу,
Которая часто ворует пшеницу,
Которая в темном чулане хранится
В доме,
Который построил Джек,,,,,,,,,,-

С.Маршак

Absorption

Algebra?

, + , = ,

, + . = .

From punctuation to mathematics:

$a*(b-(c-(d+e.$

Virtual punctuation

Principle: One mark – one role.

Consequence: Equality for the number of opening and closing isolating marks.

Conservation law of roles: The role of the punctuation mark being absorbed goes to absorbing mark.

$$,) + .) = .))$$

^Divide and conquer!

Left and right borders of isolations and homogenous members separators:

$$M^{\wedge} = \{ , (- : \wedge \text{и} \}$$
$$M^{\cdot} = \{ ,) - \cdot \}$$
$$M^{|} = \{ , : ; \text{и} \text{'a'} \}$$

Punctuation formula

Contact located marks

_____ , - И _____ .

Role combinations:

) * | ? (*

Consequences:

) (| || () – impossible

) (() | | () (- possible

Operation $M) \times M)$

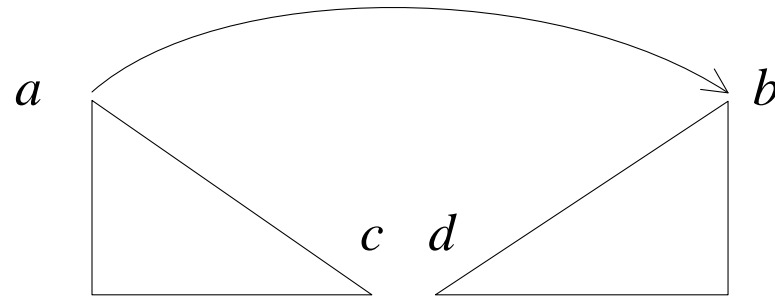
$M) \times M)$,)	-	.
,	,)	, -	.
)),))) -).
-	,)	-	.
.	#	#	#	#

(()) | | () (- in article

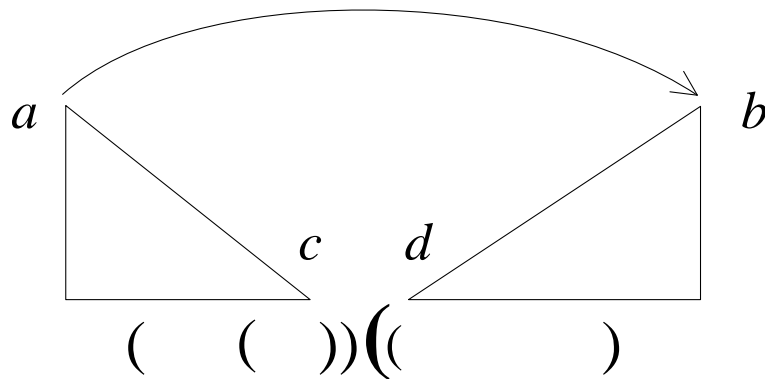
Pair marks

M(\ M)	,)	-	.
,	+			
(+		
-			+	
:			+	
И	+			
^				+

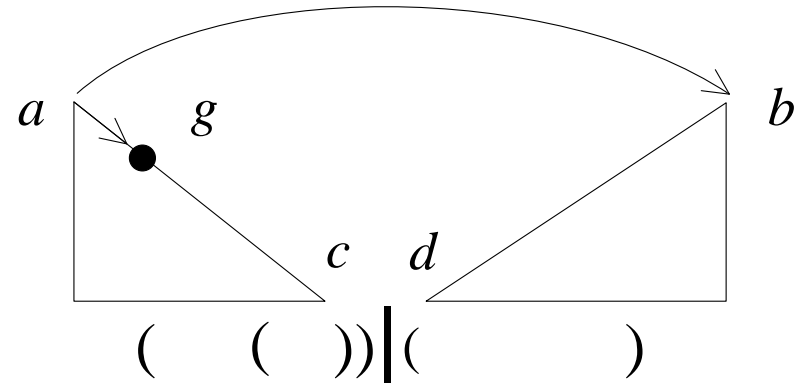
Eisner algorithm + punctuation



Merging of two paths: $p^{\rightarrow} = a, \dots, c$ и $p^{\leftarrow} = b, \dots, d$



Isolations



Coordination

Approach advantages

- Can be applied to other languages
- Possible to apply the approach in classical parsing algorithms
- Preliminary fragmentation isn't required

Contacts

Dictum Ltd.

Address:

603950 Russia, Nizhny Novgorod,
Prospect Gagarina 23, building 7

Phone / fax: +7 (831) 278-67-57

e-mail: oka@dictum.ru

web: www.dictum.ru