Syntactic and semantic models and algorithms in Question Answering

Alexander Solovyev

Bauman Moscow Sate Technical University

a-soloviev@mail.ru

Agenda

- Question Answering and Answer Validation task
- Answer Validation via Recognizing Text Entailment
  - Bags of Words/Links intersection [Wang 2008]
  - Tree edit distance [Panyakanok, Roth, Yih 2004]
  - Trees alignment [Marsi, Krahmer, Bosma, Theune 2006]
  - Predicates matching [Schlaefer 2007]
  - Parallel traversal [Solovyev 2010]
  - Automatic logic prove for logical forms [Akhmatova 2005]
- Cross-application of syntax and semantic models in various algorithms

Question Answering

ЗАДАЙТЕ ВОПРОС НА РУССКОМ ЯЗЫКЕ:

Где находится город Нягань? Найти ответ

Яндекс нашёл 431 тыс. страниц

1. РОССИЯ, вес: 1.038

Источник: http://www.kod-gorod.ru/index.php?nn=890

Для того, что бы позвонить в Нягань из другого города требуется набрать код страны (в нашем случае, т. к. город находится в России это будит 8 или даже более правильно следует...

2. ХАНТЫ-МАНСИЙСК, вес: 0.533

Источник: http://sp.avs50.ru/view_post.php?id=743

Нягань - город окружного подчинения в Ханты-Мансийском автономном округе. Он находится 230 км (воздушным сообщением) к северу от Ханты-Мансийска.

3. ХАНТЫ-МАНСИЙСК, вес: 0.533

Источник: http://sp.avs50.ru/view_post.php?id=743

Нягань - город окружного подчинения в Ханты-Мансийском автономном округе. Он находится 230 км (воздушным сообщением) к северу от Ханты-Мансийска.

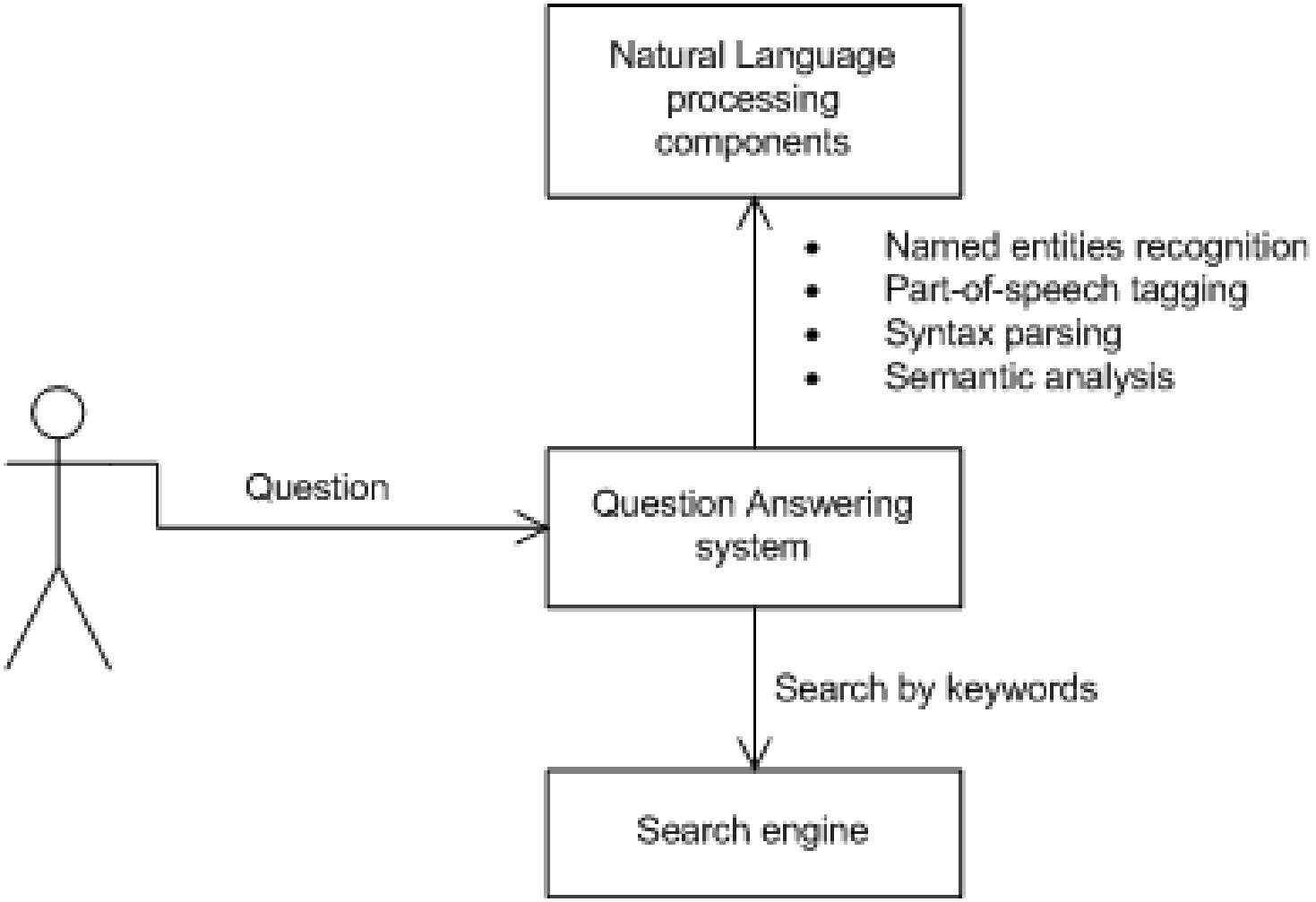
4. ХМАО-ЮГРА, вес: 0.517

Источник: http://iqplus.ru/clients_partners.html

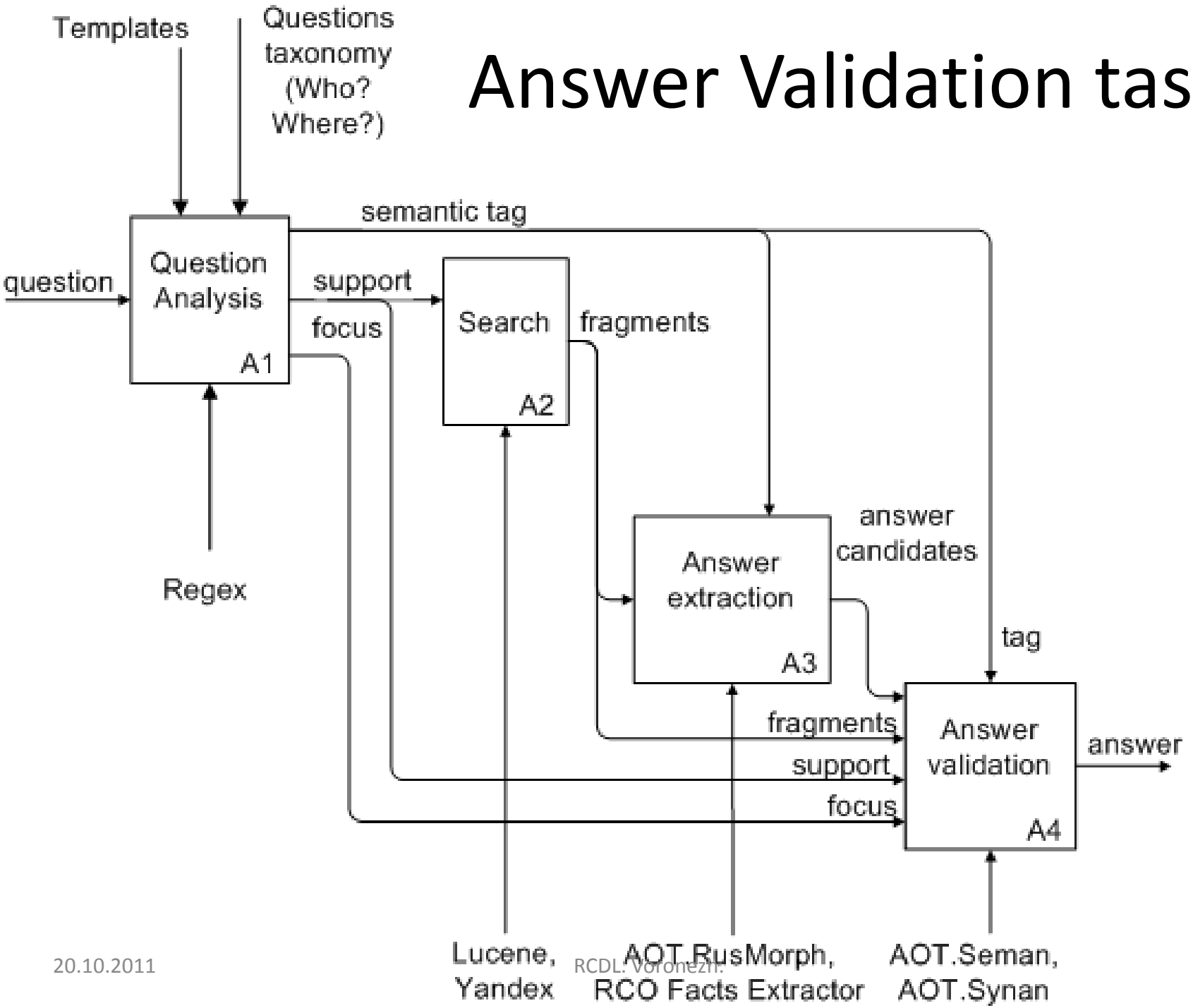
Город окружного значения Нягань, находится на территории ХМАО-Югра. Население города составляет около 60 000 человек. Муниципальное образование Шурышкарский район.

Конфигурация: PatternAnalyzer > Yandex > NerAnswerSelector(MorphNER(AOT)) > InQuestionFilter > Backup(GraphMatch(AotSeman), TEdt(AotSeman), TAlign(AotSeman), PredMatch(AotSeman), BagOfLinksFilter(AotSeman))

Meta-search architecture



Answer Validation task



Bag of words

Backup strategy in [Wang, Neumann. Using Recognizing Textual Entailment as a Core Engine for Answer Validation. 2008]

- Used as baseline method or backup strategy
- Given two sentences – question and snippet
- Replace question focus by **ANS**
- Replace answer in snippet by **ANS**
- Remove stop words and punctuations
- Count sets of distinct words in question and supporting text – *Q* and *P*
- $c=|Q \cap P|/|Q|$
- Answer is supported by snippet if *c > threshold* (e.g. *0.7*)

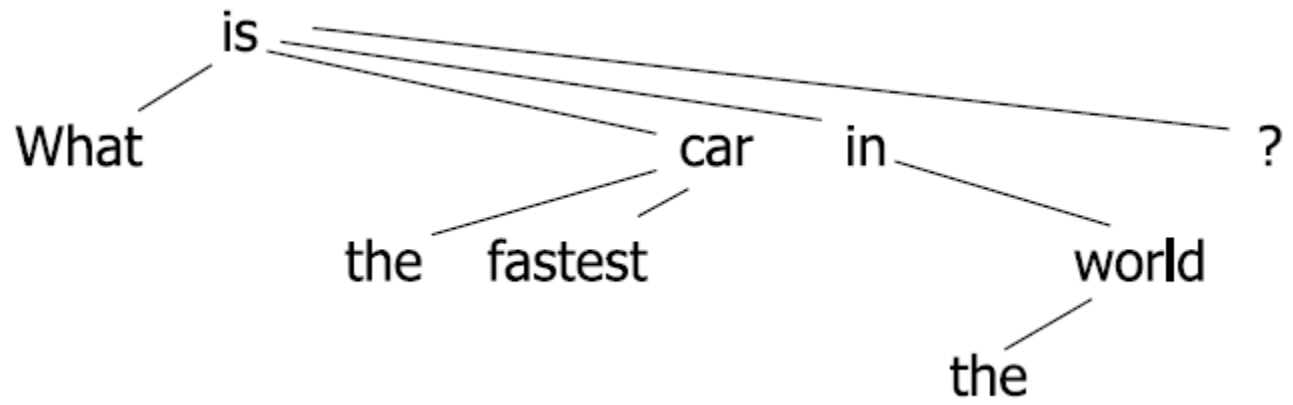
Bag of words example

- What is the fastest car in the world?
- The Jaguar XJ220 is the dearest, fastest and the most sought after car in the world.

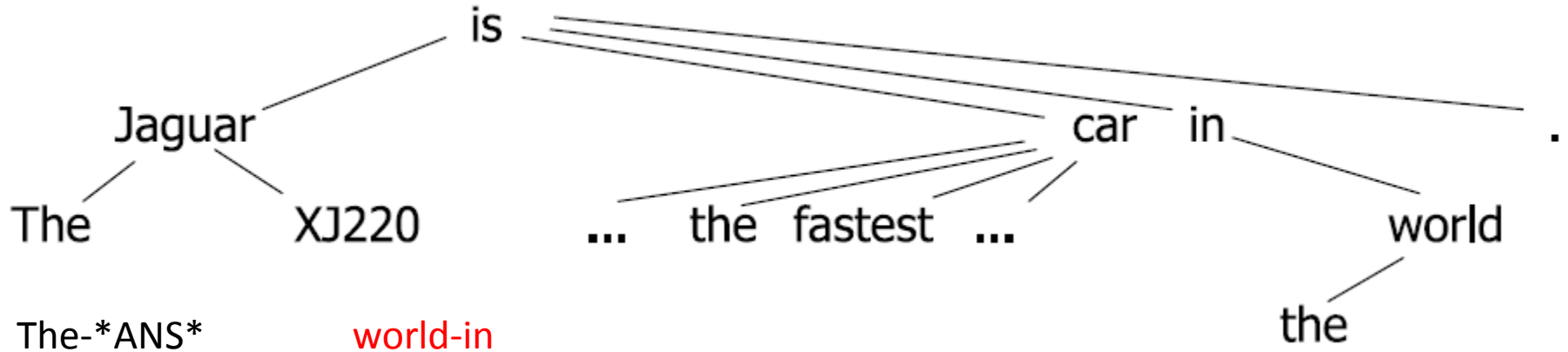
→

- *ANS* is the fastest car in the world?
- The *ANS* is the dearest, fastest and the most sought after car in the world.
- $|Q \cap P| = \{*ANS*, is, the, fastest, car, in, world\}$
- $c = |Q \cap P| / |Q| = 7/7 = 1.0$

Bag of links example



ANS-is
car-is
the-car
fastest-car
in-is
world-in
the-world
?-is



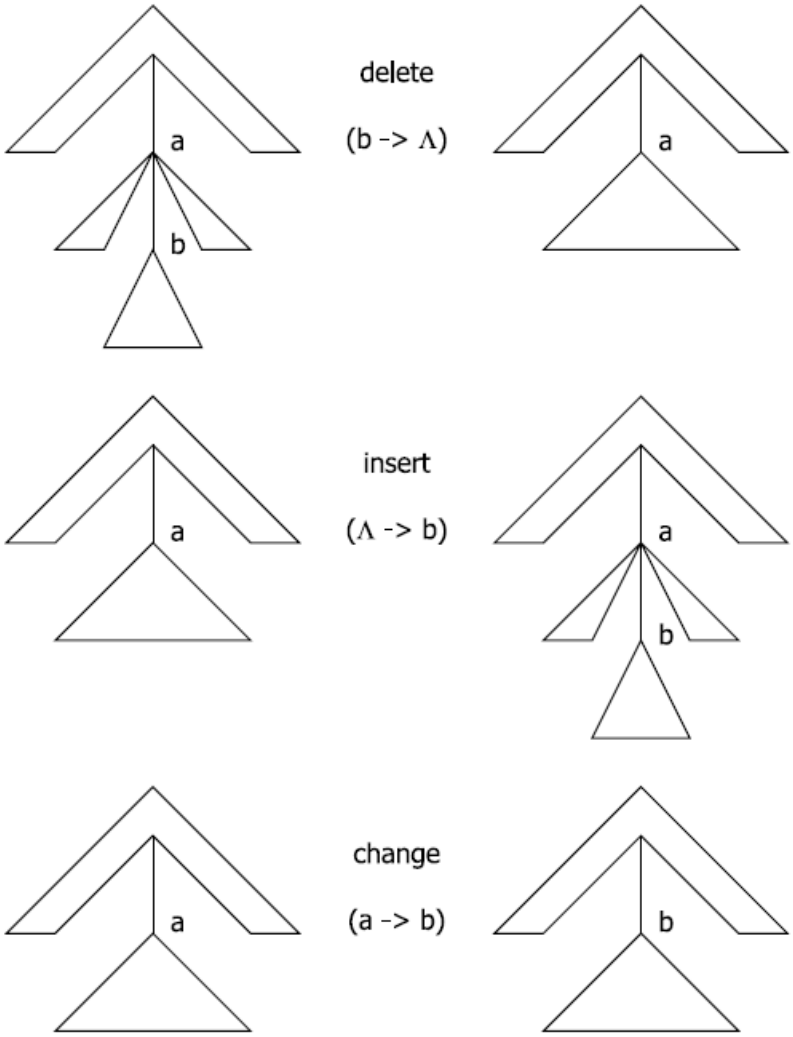
The-*ANS*
ANS-is
car-is
the-car
fastest-car
in-is

world-in
the-world
.-is
…

$$c = |Q \cap P| / |Q| = 7/7 = 1.0$$

Tree edit distance

[Punyakanok et al. Natural Language Inference via Dependency Tree Mapping. An Application to Question Answering. 2004]



- Given two <u>ordered</u> dependency trees representing question statement and snippet: T_q, T_p
- Cost of deleting a node from tree: $\gamma(a \rightarrow \lambda)$
- Cost of inserting a node into tree: $\gamma(\lambda \rightarrow a)$
- Cost of changing a node: $\gamma(a \rightarrow b)$
- Cost of a sequence of operations $S = \langle s_1; s_2; \ldots; s_k \rangle$ is $\gamma(S) = \Sigma\, \gamma(s_i)$
- Find a minimum cost of transformation T_p to T_q:

$$\delta(T_p, T_q) = \min_S \left\{ \gamma(S) \mid S(T_p) = T_q \right\}$$

Tree edit distance with subtree removal

[Zhang, Shasha. Simple fast algorithms for the editing distance between tree and related problems. 1989]

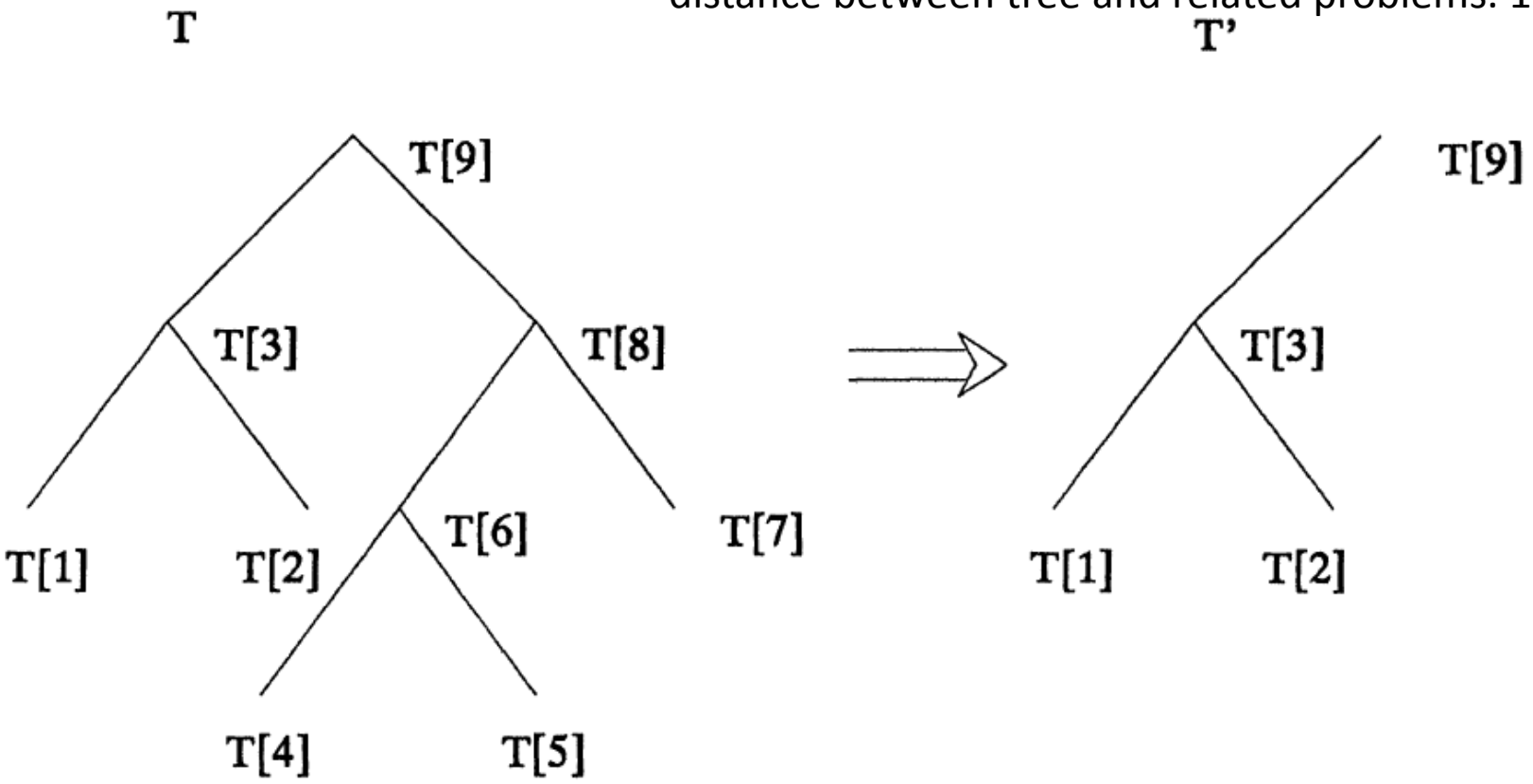


FIG. 9. Remove subtree rooted at T[8].

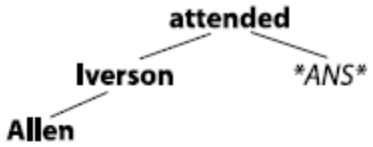
$$DR(T_q, T_p) = \min_{F \in F(T_p)} \delta(T_q, T_p \backslash F)$$

$$\delta(T_q, T_p \backslash F) = \min_S \{\gamma(S) | S(T_q) = T_p \backslash F\}$$

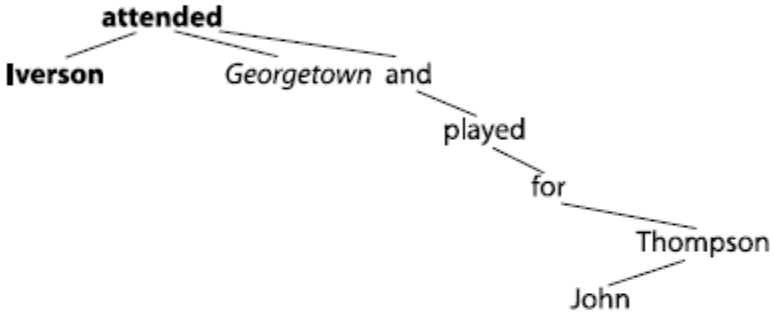
Tree edit distance vs Bag-of-words performance

[Punyakanok et al. Natural Language Inference via Dependency Tree Mapping. An Application to Question Answering. 2004]

What college did Allen Iverson attend ?
Allen Iverson attended *ANS*



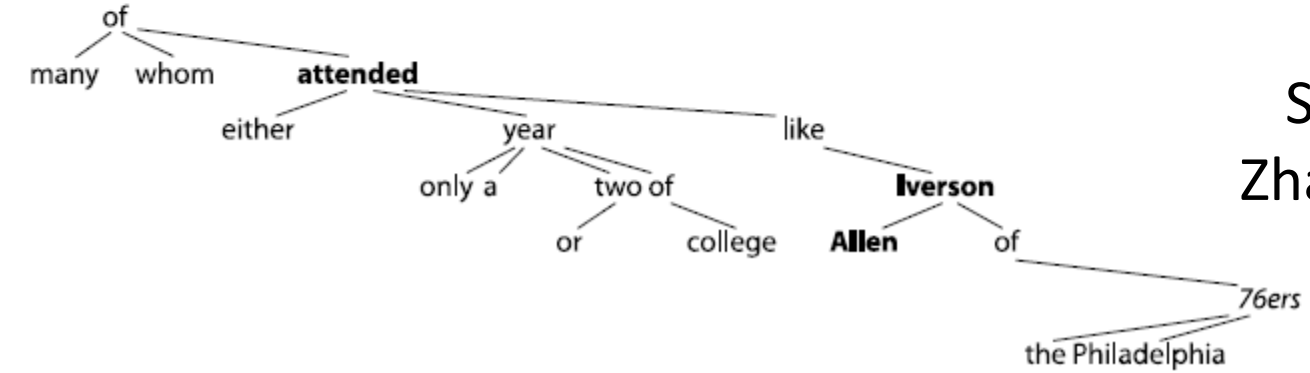
Iverson attended Georgetown and played for John Thompson ...



TREC 2002 QA

Method	Correct #	Correct %	Correct %(454)
Tree Matching	183	36.60	40.31
Bag-of-Word	151	30.20	33.26

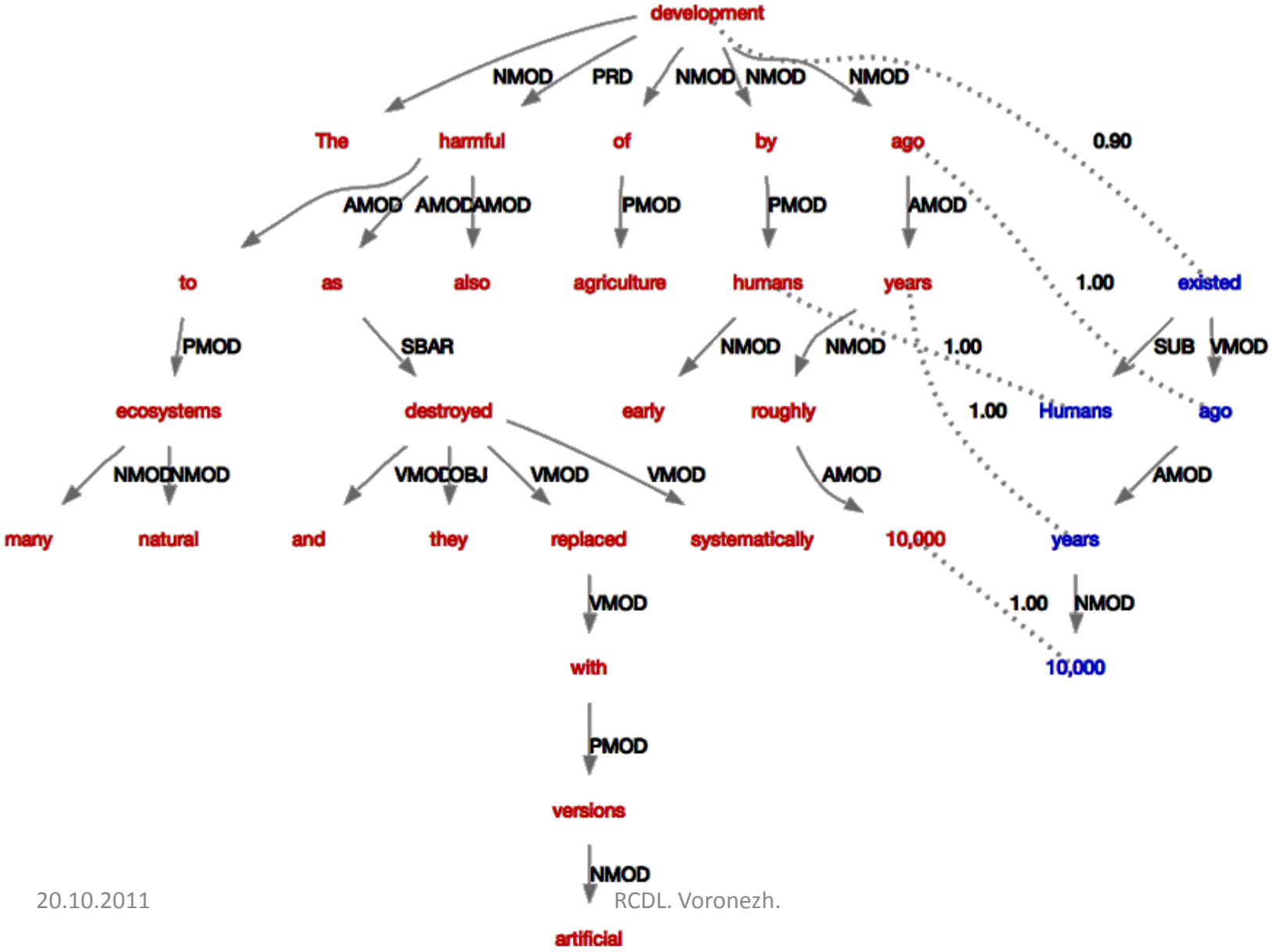
... many of whom either attended only a year or two of college, like Allen Iverson of the Philadelphia 76ers ...



Significant limitation of Zhang-Shasha algorithm: ordered trees only!

Trees alignment

[Krahmer, Bosma. Normalized alignment of dependency trees for detecting textual entailment. 2006]



Trees alignment

[Krahmer, Bosma. Normalized alignment of dependency trees for detecting textual entailment. 2006]

- Given two dependency trees representing question statement and snippet: T_q, T_p
- Skip penalty *SP,* Parent weight *PW*
- Calculate sub-trees match matrix S=|T_q|x|T_p|
- Every element $s=<v_q, v_p>$ to be calculated recursively
- Trees similarity is a score of predicates similarity

Modification:

- To replace question focus by **ANS**
- To replace answer in snippet by **ANS**
- to rotate trees to have *ANS* in roots, and use similarities of these roots.

Trees alignment

[Krahmer, Bosma. Normalized alignment of dependency trees for detecting textual entailment. 2006]

$$S(v,v') = \max \begin{pmatrix} TreeMatch(v,v') \\ \max\limits_{i} S(v_i,v') \\ \max\limits_{j} S(v,v'_j) - SP \end{pmatrix}$$

- root node *v* can be directly aligned to root node *v'*
- any of the children of v can be aligned to v'
- v can be aligned to any of the children of v' with skip penalty

$$TreeMatch(v,v') = PW \cdot ParentMatch(v,v') + (1-PW) \cdot ChildMatch(v,v')$$

$$ParentMatch(v,v') = \begin{cases} 1 & if\ word(v) = word(v') \\ 1 & if\ lemma(v) = lemma(v') \\ 1 & if\ synonym(v,v') \\ 1 & if\ hypernym(v,v') \\ sim(v,v') & if\ sim(v,v') > 0.1 \\ 0 & otherwise \end{cases}$$

$$ChildMatch(v,v') = \max_{p \in P(v,v')} \left[\sum_{(i,j) \in p} \frac{|v'_j|}{|v'|} \cdot S(v_i, v'_j) \right]$$

- P(v, v') is the set of all possible pairings of the n children of v against the m children of v', which amounts to the power set of {1…n}×{1…m}
- $|v'_j|/|v'|$ represent the number of tokens dominated by the j-th child node of node v' in the question divided by the total number of tokens dominated by node v'.

Trees alignment performance in RTE-2

[Krahmer, Bosma. Normalized alignment of dependency trees for detecting textual entailment. 2006]

parameters

Task	SP	PW	TH
IE	0.6	0.2	0.6
IR	0.8	0.1	0.6
QA	0.9	0.2	0.6
SUM	0.9	0.1	0.4

accuracy

Task	Dev_o	Dev_s	Test
IE	56.0	53.0	52.0
IR	61.0	58.0	58.5
QA	60.0	57.5	62.5
SUM	72.0	72.0	69.0
Overall	62.25	60.1	60.5

- BUT, For the RTE-2 test set, Zanzotto et al. found that simple lexical overlapping (sophisticated bag-of-words) achieves accuracy of 60%, better than any other sophisticated lexical methods they tested

Predicates matching

OpenEphyra: [Schlaefer. A Semantic Approach to Question Answering. 2007]

Semantic Role Labeling:

- Terms labeled either as predicates or arguments
- Every term fills some predicate's argument position
- Predicate-argument relationship is labeled by type of argument: ARG0, ARG1, ARGM-LOC, ARGM-TMP etc.
- Schlaefer's method ignores labels and not uses deep syntax dependencies. SRL gives two-level hierarchy: predicates and arguments. Dependencies between arguments are not considered – they all depends on predicate.

<ARGM_TMP>In what year was</ARGM_TMP>
<ARG1>the Carnegie Mellon campus</ARG1>
<ARGM_LOC>at the west coast</ARGM_LOC>
<TARGET>established</TARGET>?

<ARG1>The CMU campus</ARG1>
<ARGM_LOC>at the US west cost</ARGM_LOC>
was <TARGET>founded</TARGET>
<ARGM_TMP>in the year 2002</ARGM_TMP>

Predicates matching

[Schlaefer. A Semantic Approach to Question Answering. 2007]

- Given two Semantic-Role-Labeled statements: question and snippet
- Calculate similarity between all possible predicate-predicate pairs
- Score of the best match to consider as answer confidence

$$Sim_{Args}(p_a, p_q) := \frac{\sum_{t_a \in T_a} \max_{t_q \in T_q} \left(Sim_{ExpTerm}(t_a, t_q) \right)}{|T_q| + \left| \left\{ t_a \in T_a \middle| \max_{t_q \in T_q} \left(Sim_{ExpTerm}(t_a, t_q) \right) = 0 \right\} \right|}$$

$Sim_{ExpTerm}(t_a, t_q)$

-wordnet-based lexical similarity of terms

$$sim_{pred} = sim_{verb} \times sim_{\arg s}$$

Predicates matching performance

[Schlaefer. A Semantic Approach to Question Answering. 2007]

Technique	Questions Answered	Questions Correct	Precision	Recall
Answer type analysis	361	173	0.479	0.387
Pattern learning	293	104	0.355	0.233
Semantic parsing	154	90	0.584	0.201

Precision and recall on TREC 11 questions **with correct answers**
(500 -53=447 factoid questions)

Parallel traversal

[Solovyev. Who is to blame and Where the dog is buried? Method of answers validations based on fuzzy matching of semantic graphs in Question answering system. Romip 2010]

- Given two directed graphs representing semantic relations in question statement and in snippet
- Replace focus by *ANS* in question and answer by *ANS* in snippet
- Shortcut every node in snippet: for every pair of incoming and outgoing edge (e_i,e_o) create a new edge (source(e_i),target(e_o))

(continued..)

Parallel traversal

[Solovyev. Who is to blame and Where the dog is buried? Method of answers validations based on fuzzy matching of semantic graphs in Question answering system. Romip 2010]

- Calculate similarity of nodes *ANS* and *ANS* by recursive formula:

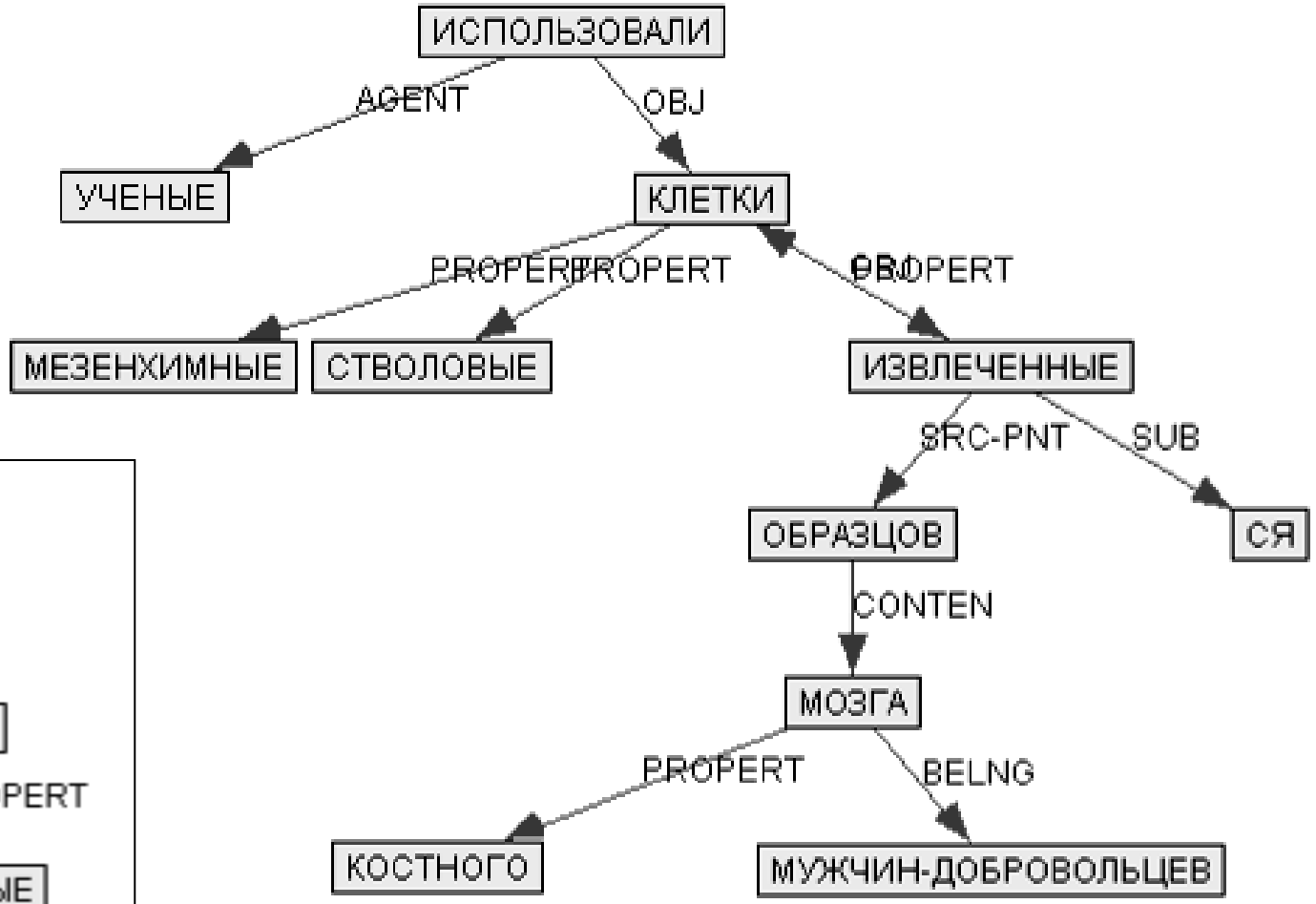
$$s(n_q, n_p) = \begin{cases} 1 + sim_{inc}(n_q, n_p) + sim_{out}(n_q, n_p) & if \ sim(n_q, n_p) > 0 \\ 0 & otherwise \end{cases}$$

$$sim_{inc}(n_q, n_p) = \sum_{e_q \in inc(n_q)} \max_{e_p \in inc(n_p)} s(src(e_q), src(e_p))$$

$$sim_{out}(n_q, n_p) = \sum_{e_q \in out(n_q)} \max_{e_p \in out(n_p)} s(trg(e_q), trg(e_p))$$

Parallel traversal

[Solovyev. Who is to blame and Where the dog is buried? Method of answers validations based on fuzzy matching of semantic graphs in Question answering system. Romip 2010]

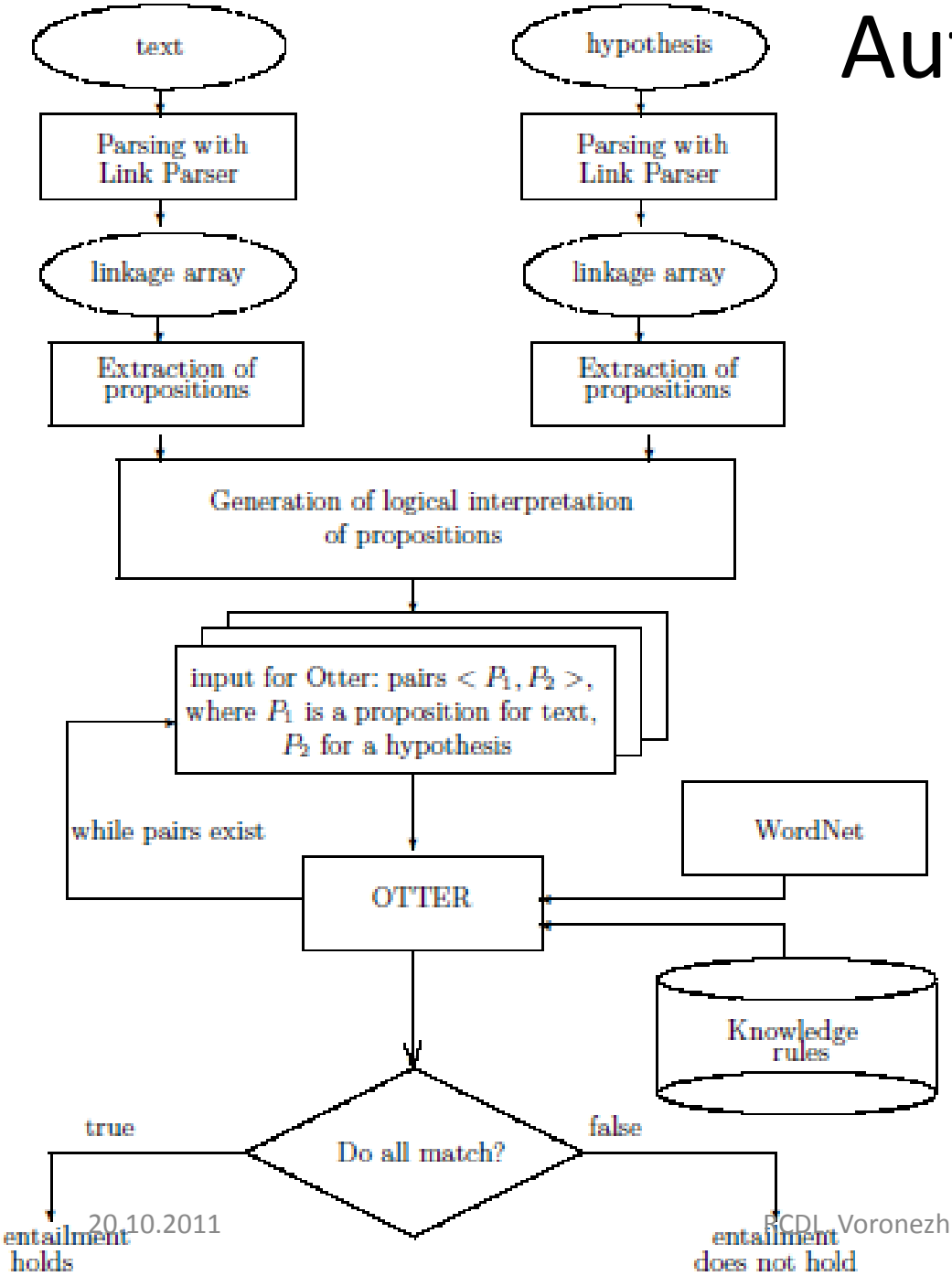


Parallel traversal performance on ROMIP

[Solovyev. Who is to blame and Where the dog is buried? Method of answers validations based on fuzzy matching of semantic graphs in Question answering system. Romip 2010]

	Recall	Error
myrtle-lucene (bag-of-words)	0.083	0.598
myrtle-seman (parallel traversal)	0.050	0.264

Automatic logic prove for logical forms



[Akhmatova et al. Recognizing Textual Entailment Via Atomic Propositions. 2006]

text. *A boy bought a desk.*
hypothesis. *A boy bought a table.*

Axiom extracted from WordNet:
X is a desk→X is a table

Input for Otter:

exists *x exists y exists e (boy(x) & bought(e, x, y) & desk(y)).*

all *x (desk(x) → table(x)).*

-(exists *x, y, e (boy(x) & bought(e, x, y) & table(y))).*

Automatic logic prove. Example.

Coffee boosts energy and provides health benefits

Coffee boosts energy.

Coffee provides health benefits.

Coffee gives health benefits

Coffee gives health benefits.

for (*Coffee gives health benefits*)

Coffee boosts energy $\nRightarrow$ *Coffee gives health benefits*

Coffee provides health benefits $\Rightarrow$ *Coffee gives health benefits.*

as

$$health\ benefits = health\ benefits$$

$$provides \Rightarrow gives,\ coffee = coffee,$$

match == 1 $\Rightarrow$entailment holds

Cross-application of models and algorithms

	Bag of words	Syntax dependencies	Semantic dependencies	Logic forms
Sets intersection	Wang 2008 Zanzotto 2006	**A**	Wang 2008	
Predicates matching		**B**	Schlaefer 2007	
Trees alignment		Marsi, Krahmer, Bosma, Theune 2006	**C**	
Tree-edit distance		Panyakanok, Roth, Yih 2004	**D**	
Parallel traversal		**E**	Solovyev 2010	
Automatic theorem prove				Akhmatova 2005

Conclusion

- There are many works on RTE and AVE in QA use Dependency trees
- Authors usually stick to single parsing model from the very beginning till the end
- There is an opportunity for replacing underlying model in existing algorithms (experiments A, B, C, D, E)
- Looks like in most works syntax relations are enough, authors ignores sophisticated semantic attributes
- Almost nothing done in Russian language
- In current project all algorithms above implemented
- Russian AVE collection is being developed…

References

QA overview:

- Ittycheriah, Abraham. A Statistical Approach for Open Domain Question Answering // Advances in Open Domain Question Answering. Springer Netherlands, 2006. Part 1. Vol.32.
- Prager, John. Open-Domain Question-Answering // Foundation and Trends in Information Retrieval, vol 1, no 2, pp 91-231, 2006.
- Rodrigo, Á., Peñas, A., and Verdejo, F. 2009. Overview of the answer validation exercise 2008. In Proceedings of the 9th CLEF 2008. Lecture Notes In Computer Science. Springer-Verlag, Berlin, Heidelberg, 296-313.

RTE in QA:

- Wang, Neumann. Using Recognizing Textual Entailment as a Core Engine for Answer Validation // Working Notes for the CLEF 2008 Workshop.
- Panyakanok, V., Roth, D. and Yih, W. Natural language interface via dependency tree mapping: An application to question answering // AI and Math.- January 2004.
- Zhang, Shasha. Simple fast algorithms for the editing distance between tree and related problems. 1989
- Marsi, E. and Krahmer, E. and Bosma, W.E. and Theune, M. (2006) Normalized Alignment of Dependency Trees for Detecting Textual Entailment. In: Second PASCAL Recognising Textual Entailment Challenge, 10-12 April 2006, Venice, Italy.
- Schlaefer, Nico. A Semantic Approach to Question Answering: Saarbrücken 2007.
- Akhmatova, E. Textual Entailment Resolution via Atomic Propositions // Proceedings of the PASCAL Challenges Workshop on Recognising Textual Entailment, Southampton, UK (2005) 61–64.
- Solovyev. Who is to blame and Where the dog is buried? Method of answers validation based on fuzzy matching of semantic graphs in Question answering system. Romip 2010

Tools:

- http://aot.ru http://rco.ru http://mu.lti.cs.cmu.edu/trac/Ephyra/wiki

Thanks

- Questions?

- a-soloviev@mail.ru
- http://qa.lib.bmstu.ru