

Technologien für Frage-Antwort-Systeme

Textual evidence gathering and
analysis

(Text 8)

Textual evidence gathering and analysis

Überblick

- Motivation
- Hypothesen
- Supporting Evidence Retrieval und weitere Grundlagen
- Textabsatz-Überprüfungsalgorithmen
 - Algorithmus 1: Passage Term Match
 - Algorithmus 2: Skip-Bigram
 - Algorithmus 3: Textual Alignment
 - Algorithmus 4: Logical Form Answer Candidate Scorer
- Zusammenführung der Bewertungen aus den verschiedenen Algorithmen
- Evaluierung
- Zusammenfassung und Ausblick

Textual evidence gathering and analysis

Motivation

Textual evidence gathering and analysis

Motivation

Absatz:

- wichtige Quelle um Hinweise zu erhalten
- ist der Antwortkandidat im Absatz enthalten?
- ist der Absatz für die betrachtete Frage relevant?



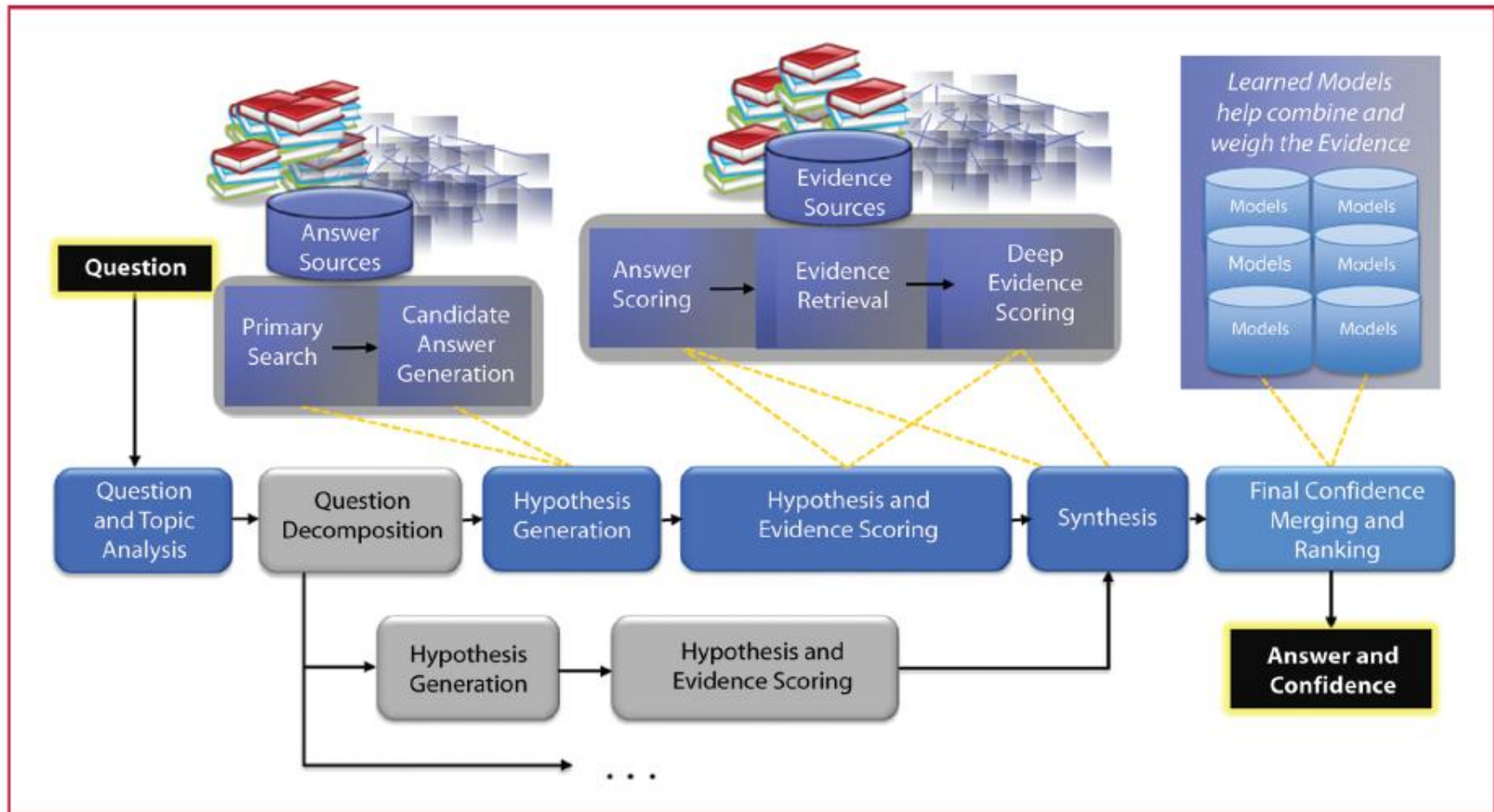
Ja:

Dieser Absatz ist eine wichtige Quelle für die Bewertung zwischen Antwortkandidat und betrachteter Frage

also: Beweis / Hinweis

Textual evidence gathering and analysis

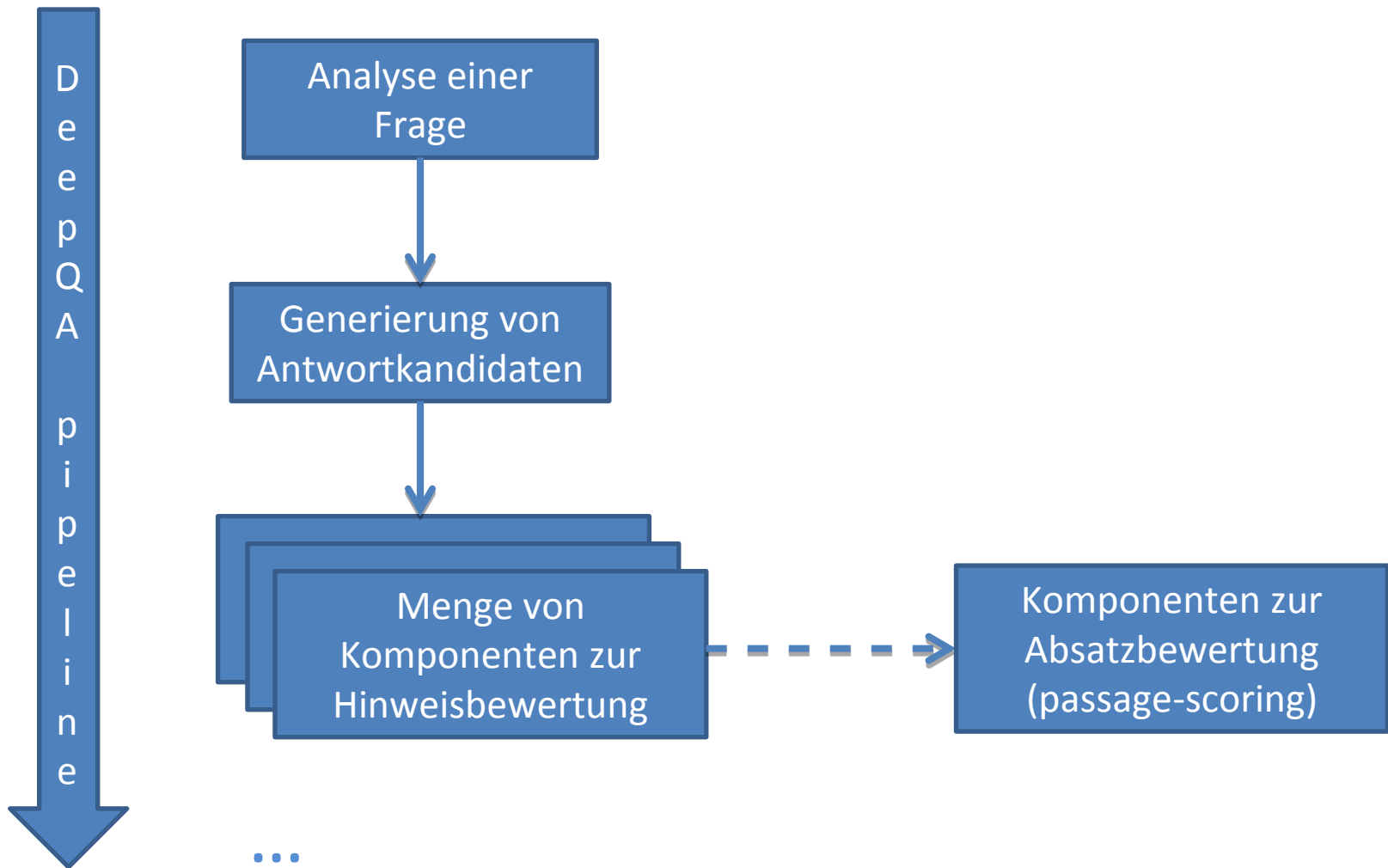
Motivation



Quelle 4 Ferrucci, D. A.: *Introduction to "This is Watson"*

Textual evidence gathering and analysis

Motivation



Textual evidence gathering and analysis

Supporting Evidence Retrieval (SER)

Textual evidence gathering and analysis

Supporting Evidence Retrieval

- Neue Technik
- Ähnlich zum Indri Absatzfindungs-Algorithmus
- Verschiedene Suchanfragen für jeden Antwortkandidaten (parallel)
- Antwortkandidaten (zwingend) in der Suchanfrage => Absätze (mit Schlüsselbeziehungen zwischen Absatz und Frage)
- Vorteil bei Redundanz in den Dokumenten
- Die 20 Absätze mit der höchsten Bewertung werden weiterverwendet

Textual evidence gathering and analysis

Supporting Evidence Retrieval

Beispiel :

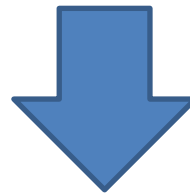
Frage:

„In 1840 this German romantic married Clara Wieck, an outstanding pianist and composer, too.“

Gefundener Antwortkandidat aus versch. Absätzen:

„Robert Schumann“

SER-Anfrage mit „Robert Schumann“ (zwingend enthalten) und weitere Wörtern aus der Frage (optional)



Textual evidence gathering and analysis

Supporting Evidence Retrieval

Beispiel :

SER-Anfrage mit „**Robert Schumann**“ (zwingend enthalten) und weiteren Wörtern aus der Frage (optional)



Neue Absätze die für die weitere Überprüfung verwendet werden
z.B.:

„Although **Robert Schumann** made some ‘symphonic attempts’ in the autumn of 1840, soon after he married his beloved Clara Wieck, he did not compose the symphony until early 1841.“

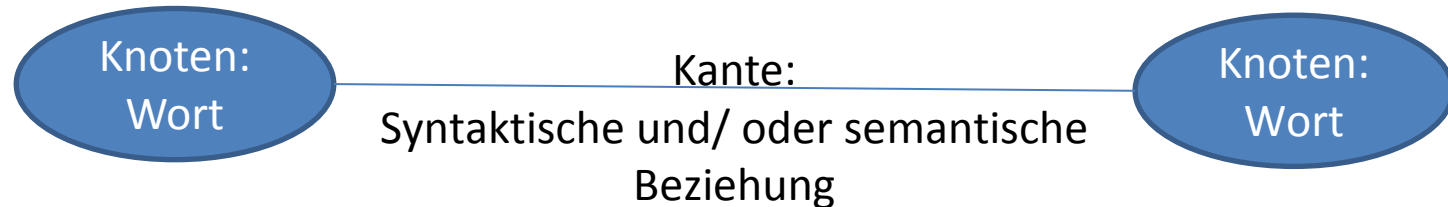
Textual evidence gathering and analysis

Syntaktisch-Semantische-Graphen

Textual evidence gathering and analysis

Syntaktisch-Semantische-Graphen

- Verwendung um strukturelle Eigenschaften von Frage und Absatz zu repräsentieren
- Dient dem Vergleich zwischen Frage und Absatz



Textual evidence gathering and analysis

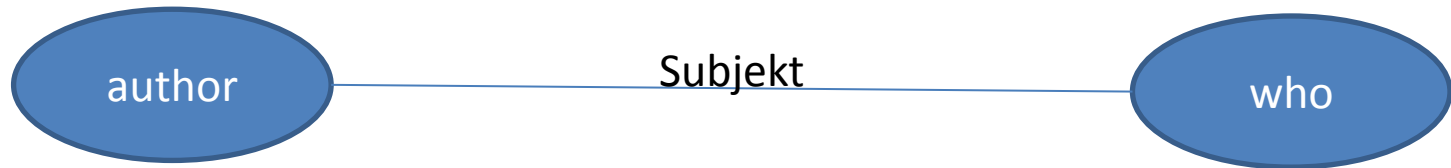
Syntaktisch-Semantische-Graphen

Beispiel :

Frage:

„Who authored ‘The Good Earth’?“

syntaktischer Graph:



Jedoch nicht ausreichend, da folgender Absatz kein Verb besitzt:

„Pearl Buck, author of ‘The Good Earth’“

Textual evidence gathering and analysis

Syntaktisch-Semantische-Graphen

Beispiel :

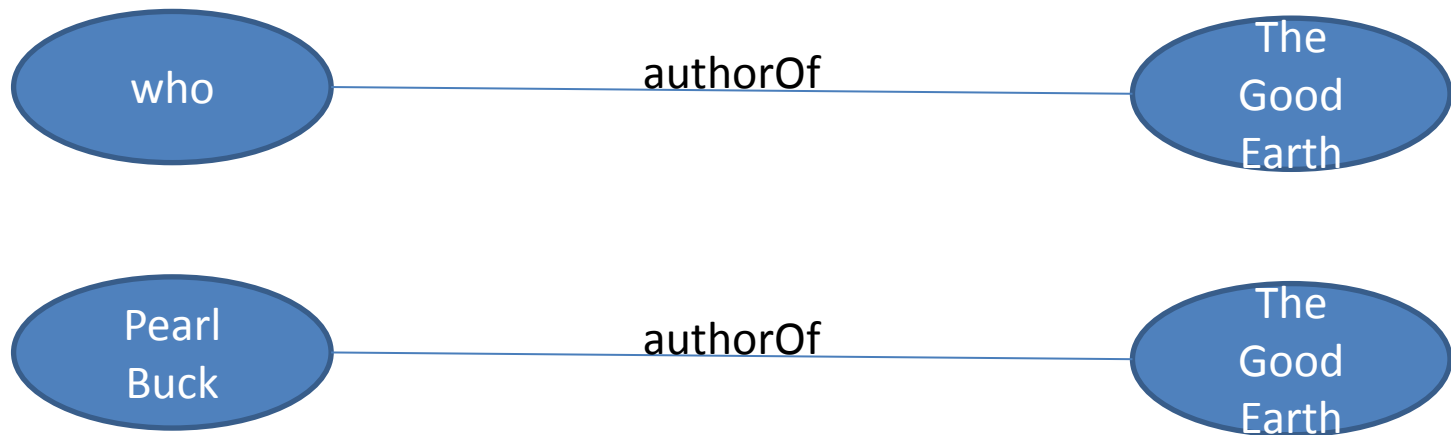
Frage:

„Who authored ‘The Good Earth’?“

Absatz:

„Pearl Buck, author of ‘The Good Earth’“

Erweiterung mit Semantik:



Textual evidence gathering and analysis

Syntaktisch-Semantische-Graphen

Verwendung von 2 Sorten semantischer Beziehungen:

1. Tiefe semantische Beziehungen (deep semantic relations):

- > stark typabhängige ontologische Beziehungen
- > z.B. authorOf Beziehung

2. Flache semantische Beziehungen (shallow semantic relations):

- > für semantische Beziehungen in alternativem syntaktischen Kontext
- > abstrahiert von syntaktischen Unterschieden
- > z.B. relatedTo, instanceOf, sameAs Beziehungen

Textual evidence gathering and analysis

Hypothesen

Textual evidence gathering and analysis

Hypothesen

Hypothese 1:

Die Analyse von Absätzen mit Hilfe einer **Vielfalt von Strategien**, die eine Analyse in **unterschiedlichen Tiefen** durchführen, ist **effektiver** als die Verwendung einer einzelnen Bewertungsstrategie

-> hier: 4 Algorithmen

Textual evidence gathering and analysis

Hypothesen

Hypothese 2:

Supporting Evidence Retrieval (SER) verbessert die **Effektivität** der Absatzbewertung dadurch, dass **mehr Absätze** für jede Antwort zur Verfügung gestellt werden, als dies bei der „Primary Search“ alleine möglich ist

- > SER findet mehr Absätze die eine Schlüsselbeziehung zwischen einem Antwortkandidaten und einer Frage enthalten
- > Einige Absätze enthalten evtl. eine passendere Formulierung für die Antwortbewertungstechniken als andere
- > Redundanz (=> Höher Wahrscheinlichkeit eine richtige Antwort zu finden)

Textual evidence gathering and analysis

Algorithmen

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen

Algorithmus 1:

Passage Term Match

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Passage Term Match

Idee:

Wie oft kommt ein Antwortkandidat zusammen mit den Wörtern
aus der Frage in einem Absatz vor

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Passage Term Match

Beispiel:

Frage:

„In 1840 this German romantic married Clara Wieck, an outstanding pianist and composer, too.“

Antwortkandidat:

Robert Schumann

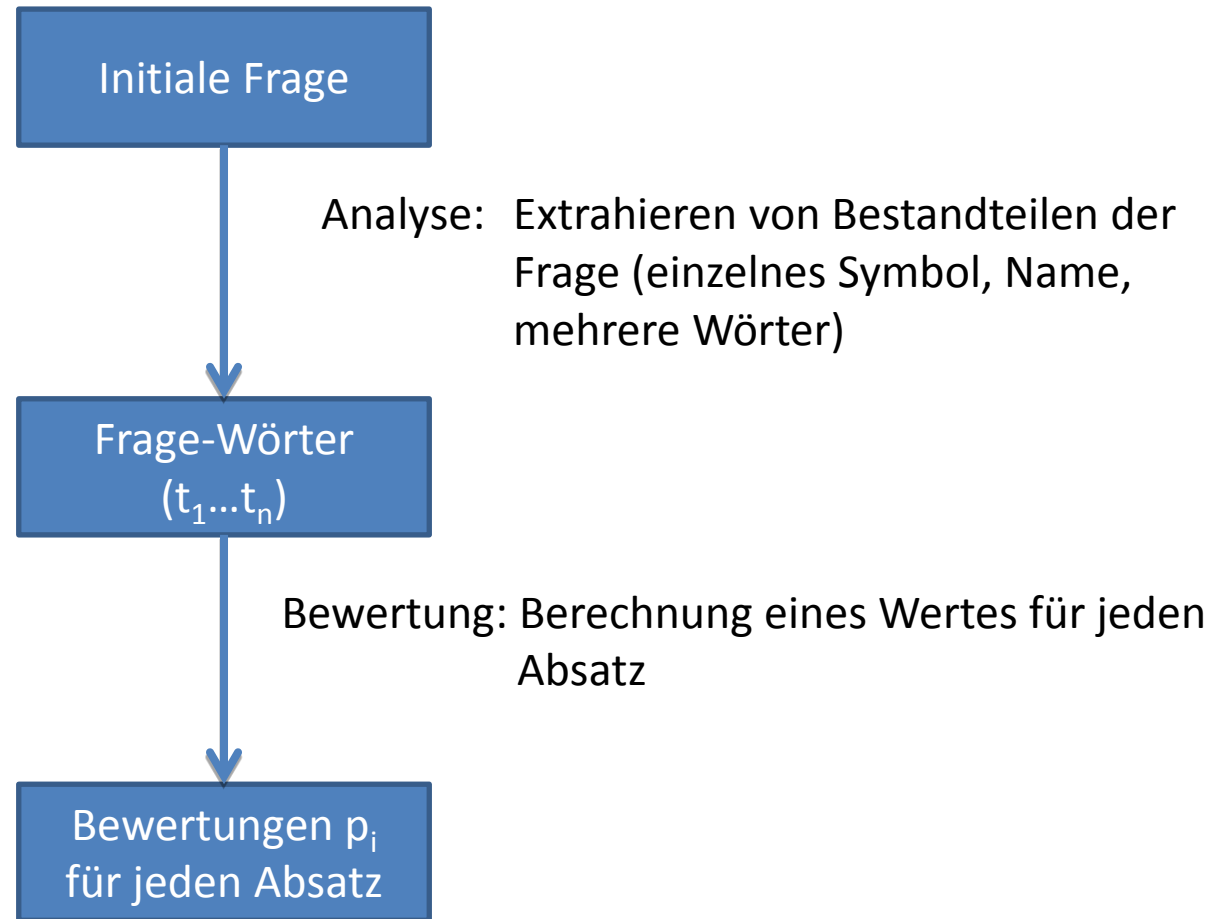
Absatz:

„Although **Robert Schumann** made some ‘symphonic attempts’ in the autumn of **1840**, soon after he **married** his beloved **Clara Wieck**, he did not **compose** the symphony until early 1841.“

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Passage Term Match

Ablauf:



Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Passage Term Match

Berechnung:

$$p_i = \frac{\sum_{j=1}^n w_{ij}}{\sum_{k=1}^n idf(t_k)}$$

$t_1 \dots t_n$

->

Frage Wörter

p_i

->

Bewertung für jeden Absatz

idf

->

Inverse Dokumenten Häufigkeit
(Frequenz)
(Summe der idf-Werte dient
der Normalisierung)

$w_{ij} = idf(t_j)$

wenn Absatz i das Fragewort
 t_j enthält,
sonst

$w_{ij} = 0$

$c(t)$

->

Anzahl der Dokumente, die t
enthalten

$$idf(t) = \log \frac{N}{c(t) + 1}$$

N

->

Gesamtanzahl der Dokumente

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen

Algorithmus 2:

Skip-Bigram

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Skip-Bigram

Idee:

Überprüfung der Übereinstimmung von Wortpaaren in
Frage und Absatz;
hierzu Verwendung/ Vergleich von Skip-Bigrammen

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Skip-Bigram

Definition Skip-Bigram:

Ein Skip-Bigram ist im folgenden Kontext ein Wortpaar, ...

... dass in einem Graphen direkt zusammenhängend ist, oder

... bei dem jedes Wort mit einem gemeinsamen Knoten verbunden ist
(„skipped“)

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Skip-Bigram

Berechnung:

Die Ähnlichkeit zwischen einer Frage und einem Absatz mit einem Wert zwischen 0 und 1 bewertet

$$score_P = \frac{|P_{sb} \cap Q_{sb}|}{|P_{sb}|}$$

P_{sb}, Q_{sb} -> Menge von Skip-Bigrammen; aus Absatz (P) und aus der Frage (Q) extrahiert

$$score_Q = \frac{|P_{sb} \cap Q_{sb}|}{|Q_{sb}|}$$

Bewertungsfaktor für Absatz und Frage ist die Größe des gemeinsamen Skip-Bigram mit Längennormalisierung

$$score = \frac{2 \cdot score_P \cdot score_Q}{score_P + score_Q}$$

Endgültige Bewertungsfaktor:
Harmonisches Mittel

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Skip-Bigram

Beispiel:

“Who invented the motor driven phonograph?”

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Skip-Bigram

Beispiel:

“Who invented the motor driven phonograph?”



Skip-Bigram gibt eine Bewertung
auf Absätze, mit:

„motor driven phonograph“

„motor driving a phonograph“

„phonograph driven by a motor“

...

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Skip-Bigram

Beispiel:

“Who invented the motor driven phonograph?”



Skip-Bigram gibt eine Bewertung auf Absätze, mit:

„motor driven phonograph“

„motor driving a phonograph“

„phonograph driven by a motor“

...

Skip-Bigram gibt jedoch keine hohe Bewertung auf alleinstehende Teile der Frage, also z.B. wenn die Wörter „motor“, „driving“ und „phonographs“ getrennt von einander in verschiedenen Sätzen eines Absatzes vorkommen

(man Merke hier die Abgrenzung zu Passage Term Match!)

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen

Algorithmus 3:

Textual Alignment

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Textual Alignment

Idee:

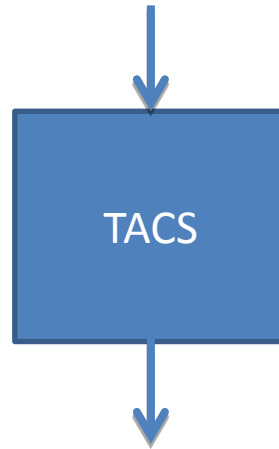
- Textual Alignment Candidate Scorer (TACS) bewertet einen Kandidaten
- Frage und Absatz (Absatz der den Kandidaten enthält) werden auf der Basis einer „oberflächlichen Ähnlichkeit“ (surface similarity) verglichen
- Es werden damit Fälle abgedeckt, in denen es einen Absatz gibt, der die Frage in ähnlicher Art und Weise enthält
- TACS ist ein Algorithmus, der mit der Diskrepanz zwischen Absatz und Frage umgehen kann
 - > TACS ist robust bezüglich Unterschiedlichkeiten
 - > TACS beachtet Bewertungen bezüglich unvollständiger Vergleiche

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Textual Alignment

Idee:

- Input:
 1. Fragestring (angepasst)
 2. Absatzstring
 3. Fokus der Frage
 4. Kandidat der Bewertet wird



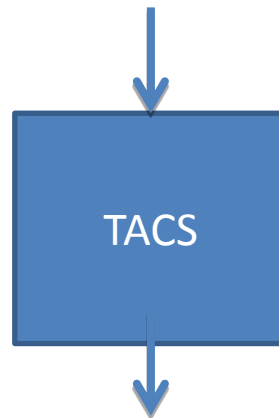
- Output: Ähnlichkeitsmessung

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Textual Alignment

Idee:

- Input:
 1. Fragestring (angepasst)
 2. Absatzstring
 3. Fokus der Frage
 4. Kandidat der Bewertet wird



- Output: Ähnlichkeitsmessung

(hierzu: Verwendung des „Waterman-Smith-Algorithmus“
-> Verwendung im Bereich DNA/Aminosäuresequenz-Abgleich
vgl. auch Quelle 5 Needleman et al.: *A General Method Applicable to the Search for Similarities in the Amino Acid Sequence of Two Proteins*)

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Textual Alignment

Berechnung:

1. Array Initialisierungen (3 Stück) :

P	->	1 Dim Enthält die Token des Absatzes
Q	->	1 Dim Enthält die Token der Frage
score	->	2 Dim Speichert die Vergleichsbewertungen am Ende des Algorithmus Initial: $\text{score}[i][j] = 0$ für alle i und j

2. Berechnung für jeden Wert $\text{score}[i][j]$ für $i > 0$ und $j > 0$

3. Rückgabe des Maximum in score als Ähnlichkeitsbewertung zwischen Eingabefrage und Absatz

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Textual Alignment

Berechnung für jeden Wert in score:

$$\max \begin{pmatrix} score[i-1][j-1] + \text{sim}(P[i], Q[j]), \\ score[i-1][j] + \text{sim}(P[i], \phi), \\ score[i][j-1] + \text{sim}(\phi, Q[j]), \\ 0 \end{pmatrix}$$

Mit der Ähnlichkeitsfunktion
 $\text{sim}(t_1, t_2)$:

$$\text{sim}(t_1, t_2) = \begin{cases} \text{idf}(t_1), & \text{if } t_1 = t_2 \\ -\text{idf}(t_1), & \text{if } t_2 = \phi \\ -\text{idf}(t_2), & \text{if } t_1 = \phi \\ -\text{idf}(t_1), & \text{if otherwise,} \end{cases}$$

$$\text{sim}(FOCUS, CANDIDATE) = \log(N)$$

N ist die Anzahl der Dokumente, die
auch für idf verwendet wurde

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Textual Alignment

Berechnung für jeden Wert in score:

$$idf(t) = \log \frac{N}{c(t) + 1}$$

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Textual Alignment

Berechnung – Interpretation:

- Der Wert jedes Eintrages in $\text{score}[i][j]$ für $i > 0$ und $j > 0$ gibt die maximale Ähnlichkeit zwischen $P[1, \dots, i]$ und $Q[1, \dots, j]$ an
- Diese ist nicht negativ
- Maximalwert von drei Vergleichsmöglichkeiten zwischen $P[1, \dots, i]$ und $Q[1, \dots, j]$:
 1. $P[1, \dots, i-1]$ verglichen mit $Q[1, \dots, j-1]$
 2. $P[1, \dots, i-1]$ verglichen mit $Q[1, \dots, j]$
 3. $P[1, \dots, i]$ verglichen mit $Q[1, \dots, j-1]$
- Ähnlichkeitsfunktion $\text{sim}(t_1, t_2)$ steht für die Ähnlichkeit/ Kosten für das Ersetzen von t_1 durch t_2
- Mit der letzten Definition wird ermöglicht, dass der Vergleich zwischen FOCUS der Frage mit dem Antwortkandidaten CANDIDATE die höchste mögliche Bewertung erhält

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Textual Alignment

Beispiel:

Frage: „Who is the president of France?“

Absatz mit Satz: „Nicolas Sarkozy **is the president of France.**“

=> Hier wird Nicolas Sarkozy mit einer hohen Bewertung als Antwort vermutet

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen

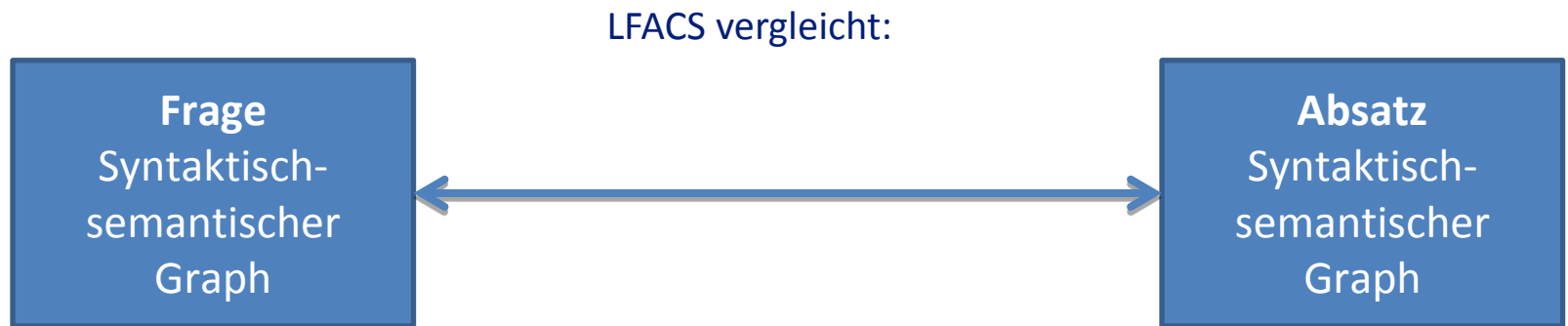
Algorithmus 4:

Logical Form Answer Candidate Scorer (LFACS)

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Logical Form Answer Candidate Scorer (LFACS)

Idee:



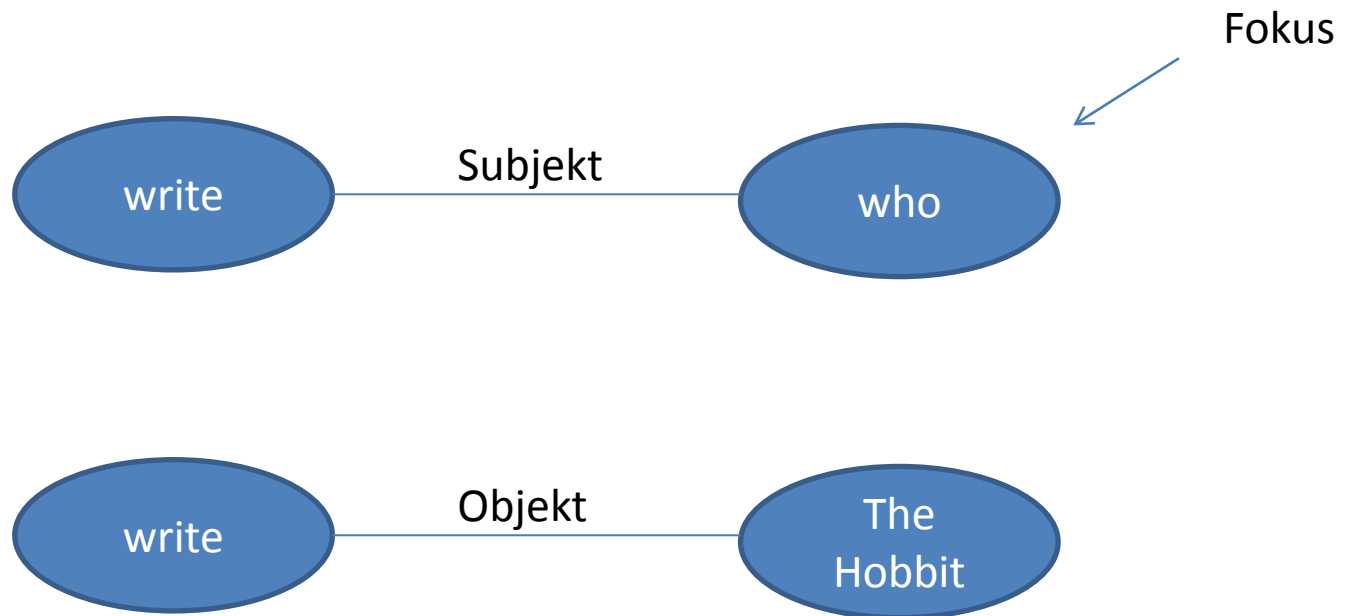
- > Vergleich zwischen Frage-Graph und Absatz-Graph
- > Knoten für den Focus der Frage muss mit dem Knoten für den Antwortkandidaten im Absatz zusammenpassen
- > Der Grad an Übereinstimmung bildet die Bewertung die dem Antwortkandidaten zugewiesen wird

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Logical Form Answer Candidate Scorer (LFACS)

Beispiel:

Who wrote The Hobbit?

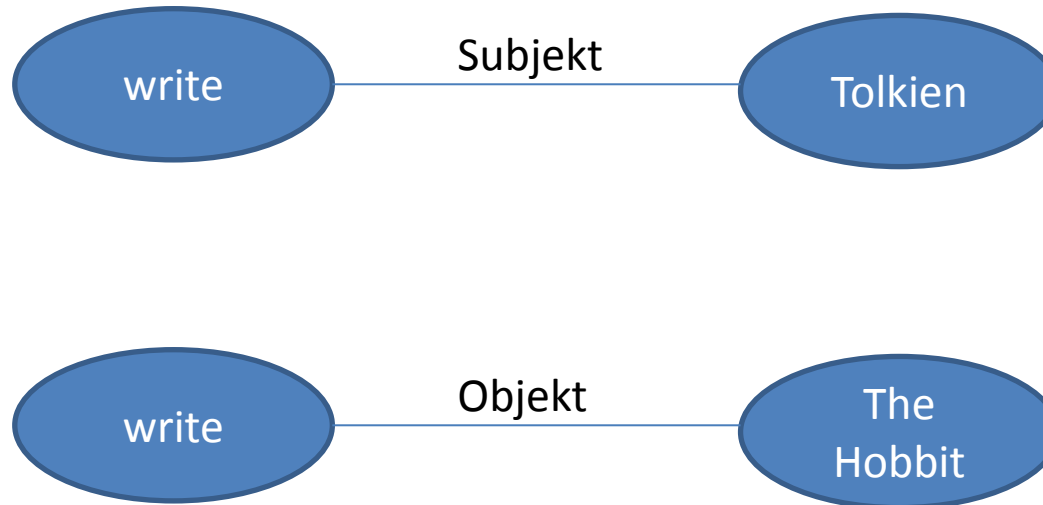


Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Logical Form Answer Candidate Scorer (LFACS)

Beispiel „Who wrote *The Hobbit*?“:

Gleicher Graph für einen Absatz mit: „Tolkien wrote *The Hobbit*“



-> LFACS kann diese Graphen vergleichen

=> Hohe Bewertung für „Tolkien“

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Logical Form Answer Candidate Scorer (LFACS)

Was passiert mit dem Absatz „Dan Brown wrote The Da Vinci Code“??

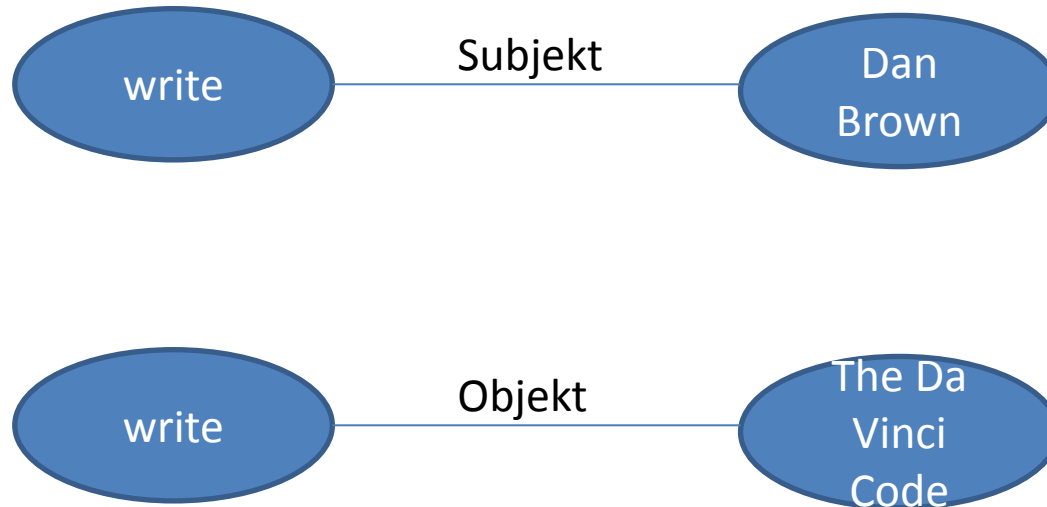
Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Logical Form Answer Candidate Scorer (LFACS)

Beispiel „Who wrote *The Hobbit*?“:

Unterschiede im Graph für einen Absatz mit:

„Dan Brown wrote The Da Vinci Code“



-> LFACS bewertet schlechter (Grund: Objekt „The Da Vinci Code“)

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Logical Form Answer Candidate Scorer (LFACS)

Wo liegt der Unterschied zu Textual Alignment bzw. einfachem Textvergleich?

-> Syntaktische und semantische Sicht

-> Tiefere Analyse!

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Logical Form Answer Candidate Scorer (LFACS)

Wo liegt der Unterschied zu Textual Alignment bzw. einfachem Textvergleich?

„Dan Brown wrote several books and has read The Hobbit.“

„Tolkien, an English author born in the late nineteenth century, wrote The Hobbit.“

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Logical Form Answer Candidate Scorer (LFACS)

Negativbeispiel (???)

Frage: „In „A Christmas Carol“, we learn this **man** once **apprenticed** with the jolly merchant
Mr. Fezziwig.“

Absatz: „The character Fezziwig owned the business where **Scrooge** was **apprenticed** in
„A Christmas Carol“.“

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Logical Form Answer Candidate Scorer (LFACS)

Berechnung – Bewertung für eine Antwort in einem Absatz:

$$\text{score} = \sum (\text{score for each term})$$

$$\text{score for each question term} = (\text{degree of match}) * (\text{term weight})$$

$$\text{degree of match} = (\text{term match score}) * (\text{structural match score})$$

term match score =

Grad zu dem ein Teil aus der Frage zum entsprechenden Teil aus dem Absatz passt

structural match score =

Maximum aller Wege vom Fragewort zum Fokus ;
dabei das Produkt von allen term match scores und
edge match scores für die Wörter und Kanten auf
diesem Pfad

Textual evidence gathering and analysis

Textabsatz-Überprüfungsalgorithmen - Logical Form Answer Candidate Scorer (LFACS)

Ergebnis:

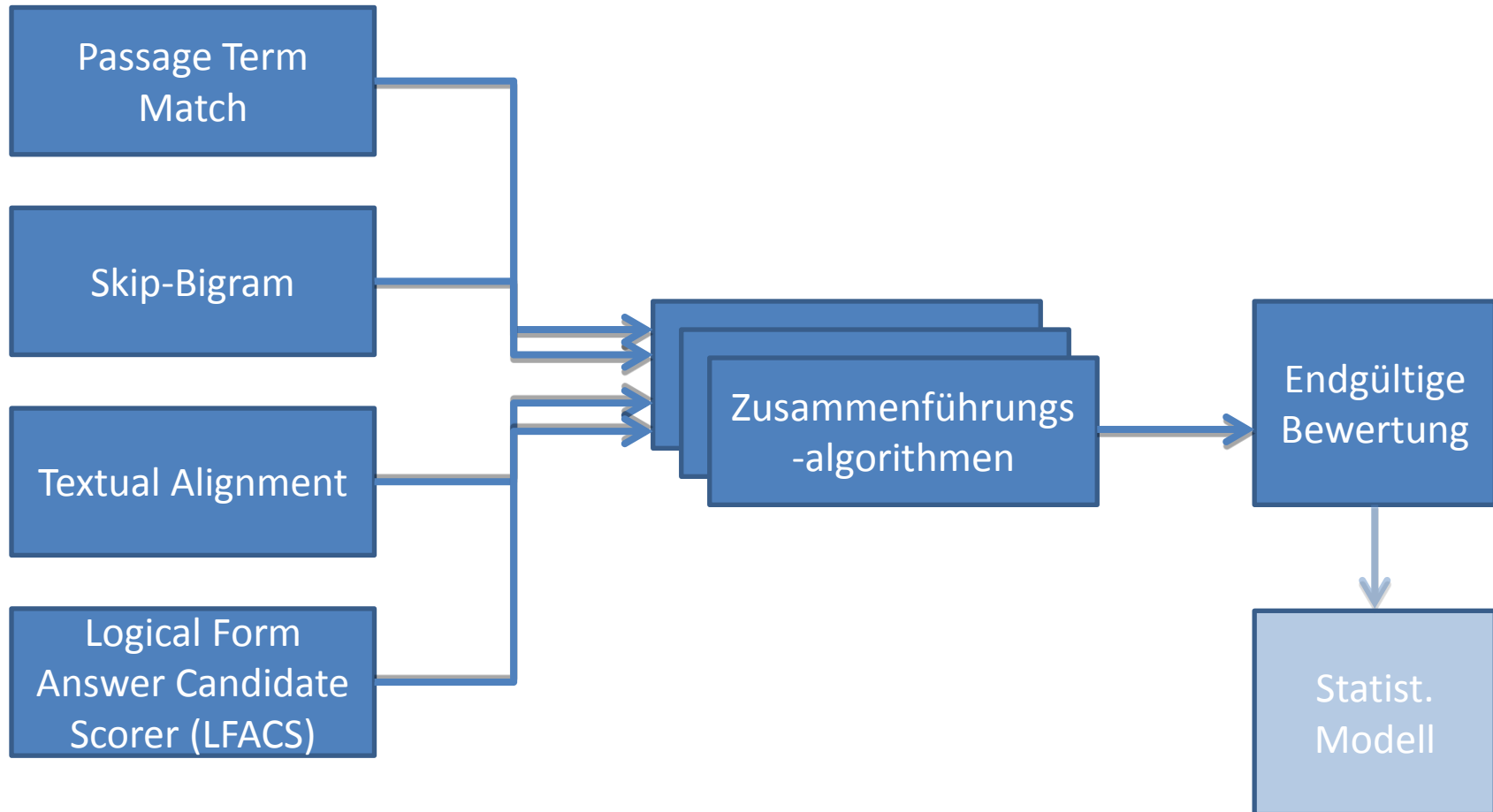
- > Für bestimmte Absatztypen ist LFACS effektiver als die anderen Algorithmen
- > LFACS ist bei unterschiedlichen syntaktischen Strukturen zwischen Absatz und Frage nicht effektiv
- > LFACS ist für das Gesamtsystem wichtig, da Fehleinschätzungen verhindert werden
- > LFACS wichtig für die Zukunft (komplexer Prozess)

Textual evidence gathering and analysis

Zusammenführung der
Bewertungen aus den verschiedenen
Algorithmen

Textual evidence gathering and analysis

Zusammenführung der Bewertungen aus den verschiedenen Algorithmen



Textual evidence gathering and analysis

Zusammenführung der Bewertungen aus den verschiedenen Algorithmen

3 Zusammenführungsalgorithmen:

1. Maximum

2. Summe (Sum)

3. Abklingende Summe (Decaying sum)

Textual evidence gathering and analysis

Zusammenführung der Bewertungen aus den verschiedenen Algorithmen

Maximum

-> Maximale Bewertung wählen

-> Verwendung, wenn eine sehr gute Bewertung für einen Absatz wertvoll ist

-> gut für LFACS

Textual evidence gathering and analysis

Zusammenführung der Bewertungen aus den verschiedenen Algorithmen

Summe (Sum)

-> Summe aller Bewertungen

-> Verwendung, wenn viele Ergebnisse relevanter sind als ein Ergebnis, dass sehr gut passt

-> gut bei Skip-Bigram

Textual evidence gathering and analysis

Zusammenführung der Bewertungen aus den verschiedenen Algorithmen

Abklingende Summe (Decaying sum)

$$\text{decay}(p_0, \dots, p_K) = \sum_{i=0}^K \frac{p_i}{2^i}$$

Quelle 3 Gondek, D. C. et al.: *A framework for merging and ranking of answers in DeepQA*

-> Kompromiss zwischen Maximum und Sum

-> gut bezüglich **Textual Alignment** und **Passage Term Match**

Textual evidence gathering and analysis

Evaluierung

Textual evidence gathering and analysis

Evaluierung

- Testen der Hypothesen
- Experiment zur Evaluierung: 3 508 Fragen (vorher ungesehen)
- Standardkonfiguration von Watson
- Baselinekonfiguration von Watson
- Vergleich dieser beiden Konfigurationen

Textual evidence gathering and analysis

Evaluierung

Table 1 Accuracy with and without SER and passage scoring.

	<i>No passage scoring</i>	<i>All passage scoring without SER</i>	<i>All passage scoring with SER</i>
<i>Watson answer-scoring baseline configuration</i>			
<i>Accuracy</i>	54.9%	58.8%	61.7%
<i>Difference vs. no passage scoring</i>		+4.0%	+6.9%
<i>Full configuration</i>			
<i>Accuracy</i>	67.1%	68.3%	70.4%
<i>Difference vs. no passage scoring</i>		+1.3%	+3.3%

Quelle 1 Murdock, J. W. et al.: *Textual evidence gathering and analysis*

Textual evidence gathering and analysis

Evaluierung

Table 2 Accuracy (with SER) with no passage scoring versus each passage scorer separately.

	<i>No passage scoring</i>	<i>Passage Term Match</i>	<i>Skip-Bigram</i>	<i>Textual Alignment</i>	<i>LFACS</i>
<i>Watson answer-scoring baseline configuration</i>					
<i>Accuracy</i>	54.9%	58.4%	57.6%	58.1%	57.0%
<i>Difference vs. no passage scoring</i>		+3.6%	+2.8%	+3.2%	+2.1%
<i>Full configuration</i>					
<i>Accuracy</i>	67.1%	68.9%	68.7%	68.6%	67.5%
<i>Difference vs. no passage scoring</i>		+1.8%	+1.7%	+2.1%	+0.4% ^a
^a Difference is not statistically significant.					

Quelle 1 Murdock, J. W. et al.: *Textual evidence gathering and analysis*

Textual evidence gathering and analysis

Evaluierung

Table 3 Accuracy (with SER) with all four passage scoring versus all except each passage scorer.

	<i>All passage scoring</i>	<i>Ablating Passage Term Match</i>	<i>Ablating Skip-Bigram</i>	<i>Ablating Textual Alignment</i>	<i>Ablating LFACS</i>	<i>No passage scoring</i>
<i>Watson answer-scoring baseline configuration</i>						
<i>Accuracy</i>	61.7%	60.7%	60.7%	60.0%	58.3%	54.9%
<i>Difference vs. all passage scoring</i>		-1.0%	-1.1%	-1.8%	-3.4%	-6.9%
<i>Full configuration</i>						
<i>Accuracy</i>	70.4%	70.0%	69.9%	70.5%	70.0%	67.1%
<i>Difference vs. All passage scoring</i>		-0.3% ^a	-0.4%	+0.2% ^a	-0.4%	-3.2%

^aDifference is not statistically significant.

Quelle 1 Murdock, J. W. et al.: *Textual evidence gathering and analysis*

Textual evidence gathering and analysis

Evaluierung

Table 4 Precision and recall of the components relative to the global mean score from each component, with a correcting factor for class imbalance.

	<i>Passage Term Match</i>	<i>Skip-Bigram</i>	<i>Textual Alignment</i>	<i>LFACS</i>
<i>Relative Precision</i>	73.7%	81.4%	75.3%	86.2%
<i>Relative recall</i>	91.8%	90.4%	84.9%	57.5%
<i>Relative F measure</i>	81.7%	85.7%	79.8%	69.0%

Quelle 1 Murdock, J. W. et al.: *Textual evidence gathering and analysis*

Textual evidence gathering and analysis

Zusammenfassung und Ausblick

Textual evidence gathering and analysis

Literatur / Quellen

1. Murdock, J. W. ; Fan, J. ; Lally, A. ; Shima, H. ; Boguraev, B. K.:
Textual evidence gathering and analysis, 2012
2. Murdock, J. W.:
Structure Mapping for Jeopardy! Clues, 2011
3. Gondek, D. C. ; Lally, A. ; Kalyanpur, A. ; Murdock, J. W. ; Duboue, P. A. ; Zhang, L. ; Pan, Y. ; Qiu, Z. M. ; Welty, C.:
A framework for merging and ranking of answers in DeepQA, 2012
4. Ferrucci, D. A.:
Introduction to "This is Watson", 2012
5. Needleman, S. B. ; Wunsch C. D.:
A General Method Applicable to the Search for Similarities in the Amino Acid Sequence of Two Proteins, 1970

Textual evidence gathering and analysis

Diskussion

Fragen??