Unsupervised Sentence Simplification Using Deep Semantics

Shashi Narayan[1], Claire Gardent[2]

[1]School of Informatics, University of Edinburgh

[2]CNRS, LORIA, Nancy

INLG, September 2016

This talk is about ...

Sentence Simplification

This talk is about ...

Sentence Simplification maps a complex sentence to a simpler sentence(s)

- which is more readable one approximating its content, and

This talk is about ...

Sentence Simplification maps a complex sentence to a simpler sentence(s)

- which is more readable one approximating its content, and
- usually with common and often shorter words, simpler syntactic constructions and fewer modifiers.

This talk is about ...

Sentence Simplification maps a complex sentence to a simpler sentence(s)

- which is more readable one approximating its content, and
- usually with common and often shorter words, simpler syntactic constructions and fewer modifiers.

complex In 1964 Peter Higgs published his second paper in Physical Review Letters describing Higgs mechanism which predicted a new massive spin-zero boson for the first time.

simple Peter Higgs wrote his paper explaining Higgs mechanism in 1964. Higgs mechanism predicted a new elementary particle.

This talk is about ...

Sentence Simplification maps a complex sentence to a simpler sentence(s)

- which is more readable one approximating its content, and
- usually with common and often shorter words, simpler syntactic constructions and fewer modifiers.

complex In 1964 Peter Higgs published his second paper in Physical Review Letters describing Higgs mechanism [which] predicted a new massive spin-zero boson for the first time.

simple Peter Higgs wrote his paper explaining Higgs mechanism in 1964. Higgs mechanism predicted a new elementary particle.

SPLIT

This talk is about ...

Sentence Simplification maps a complex sentence to a simpler sentence(s)

- which is more readable one approximating its content, and
- usually with common and often shorter words, simpler syntactic constructions and fewer modifiers.

complex In 1964 Peter Higgs published his second paper in Physical Review Letters describing Higgs mechanism [which] predicted a new massive spin-zero boson for the first time.

simple Peter Higgs wrote his paper explaining Higgs mechanism in 1964. Higgs mechanism predicted a new elementary particle.

DROP

This talk is about ...

Sentence Simplification maps a complex sentence to a simpler sentence(s)

- which is more readable one approximating its content, and
- usually with common and often shorter words, simpler syntactic constructions and fewer modifiers.

complex In 1964 Peter Higgs published his second paper in Physical Review Letters describing Higgs mechanism [which] predicted a new massive spin-zero boson for the first time.

simple Peter Higgs wrote his paper explaining Higgs mechanism in 1964. Higgs mechanism predicted a new elementary particle.

SUBSTITUTE

This talk is about ...

Sentence Simplification maps a complex sentence to a simpler sentence(s)

- which is more readable one approximating its content, and
- usually with common and often shorter words, simpler syntactic constructions and fewer modifiers.

complex In 1964 Peter Higgs published his second paper in Physical Review Letters describing Higgs mechanism [which] predicted a new massive spin-zero boson for the first time.

simple Peter Higgs wrote his paper explaining Higgs mechanism in 1964. Higgs mechanism predicted a new elementary particle.

SUBSTITUTE

This talk is about ...

Sentence Simplification maps a complex sentence to a simpler sentence(s)

- which is more readable one approximating its content, and
- usually with common and often shorter words, simpler syntactic constructions and fewer modifiers.

complex In 1964 Peter Higgs published his second paper in Physical Review Letters describing Higgs mechanism [which] predicted a new massive spin-zero boson for the first time.

simple Peter Higgs wrote his paper explaining Higgs mechanism in 1964. Higgs mechanism predicted a new elementary particle.

SUBSTITUTE

This talk is about ...

Sentence Simplification maps a complex sentence to a simpler sentence(s)

- which is more readable one approximating its content, and
- usually with common and often shorter words, simpler syntactic constructions and fewer modifiers.

complex In 1964 Peter Higgs published his second paper in Physical Review Letters describing Higgs mechanism [which] predicted a new massive spin-zero boson for the first time.

simple Peter Higgs wrote his paper explaining Higgs mechanism in 1964. Higgs mechanism predicted a new elementary particle.

REORDER

This talk is about ...

Sentence Simplification : Potential applications

- societal applications as
  - reading aid for people with aphasia [Carroll et al., 1999],
  - for low literacy readers [Watanabe et al., 2009] and
  - for non native speakers [Siddharthan, 2002].

This talk is about ...

Sentence Simplification : Potential applications

- societal applications as
  - reading aid for people with aphasia [Carroll et al., 1999],
  - for low literacy readers [Watanabe et al., 2009] and
  - for non native speakers [Siddharthan, 2002].

- preprocessing steps for NLP systems
  - parsers and machine translation systems [Chandrasekar et al., 1996],
  - summarisation [Knight and Marcu, 2000],
  - sentence fusion [Filippova and Strube, 2008] and
  - semantic role labelling [Vickrey and Koller, 2008].

This talk is about ...

Sentence Simplification

This talk is about ...

Sentence Simplification

using **rich linguistic information** in the form of **deep semantic representation**

This talk is about ...

Sentence Simplification

using **rich linguistic information** in the form of **deep semantic representation**

using an **unsupervised framework** relying only on the **comparable wikipedia corpora**

Outline of the talk

Outline of the talk

Sentence Simplification in Literature

- Handcrafted rules capturing syntactic simplification

  [Siddharthan, 2002, Chandrasekar and Srinivas, 1997, Bott et al., 2012, Canning, 2002, Siddharthan, 2011, Siddharthan, 2010]

Sentence Simplification in Literature

- Handcrafted rules capturing syntactic simplification

  [Siddharthan, 2002, Chandrasekar and Srinivas, 1997, Bott et al., 2012, Canning, 2002, Siddharthan, 2011, Siddharthan, 2010]

- Statistical supervised framework for sentence simplification
  - using aligned sentences from **traditional English** and **Simple English Wikipedia**

  [Zhu et al., 2010, Woodsend and Lapata, 2011, Wubben et al., 2012, Coster and Kauchak, 2011, Siddharthan and Mandya, 2014, Narayan and Gardent, 2014, Xu et al., 2016]

Sentence Simplification in Literature

Monolingual tree-based translation model [Zhu et al., 2010]

- constructed a parallel corpus **PWKP** of 108,016/114,924 complex/simple sentences
- **PWKP** has aligned sentences from **English Wikipedia** and **Simple English Wikipedia**

Sentence Simplification in Literature

Monolingual tree-based translation model [Zhu et al., 2010]

- constructed a parallel corpus **PWKP** of 108,016/114,924 complex/simple sentences
- **PWKP** has aligned sentences from **English Wikipedia** and **Simple English Wikipedia**
- simplification models consists of 4 tree rewrite operations: **split**, **reorder**, **delete** and **substitution**.
- based on syntax-based machine translation [Yamada and Knight, 2001], and

Sentence Simplification in Literature

Monolingual tree-based translation model [Zhu et al., 2010]

- constructed a parallel corpus **PWKP** of 108,016/114,924 complex/simple sentences
- **PWKP** has aligned sentences from **English Wikipedia** and **Simple English Wikipedia**
- simplification models consists of 4 tree rewrite operations: **split**, **reorder**, **delete** and **substitution**.
- based on syntax-based machine translation [Yamada and Knight, 2001], and
- evaluated on the test set consisting of 100/131 complex/simple sentences.

Sentence Simplification in Literature

Quasi-synchronous grammar and integer linear programming [Woodsend and Lapata, 2011]

- learned a Quasi-synchronous grammar using the **PWKP** corpus [Zhu et al., 2010]
- additional use of the edit history of Simple Wikipedia

Sentence Simplification in Literature

Simplification as monolingual translation
[Coster and Kauchak, 2011, Wubben et al., 2012, Xu et al., 2016]

- learned **statistical machine translation system** considering complex sentences as the source and simple ones as the target

Sentence Simplification in Literature

Simplification as monolingual translation
[Coster and Kauchak, 2011, Wubben et al., 2012, Xu et al., 2016]

- learned **statistical machine translation system** considering complex sentences as the source and simple ones as the target
- phrase alignments generated by Giza++ are modified to accommodate null phrasal alignment to allow **deletion** with **reordering** and **substitution** [Coster and Kauchak, 2011]

Sentence Simplification in Literature

Simplification as monolingual translation
[Coster and Kauchak, 2011, Wubben et al., 2012, Xu et al., 2016]

- learned **statistical machine translation system** considering complex sentences as the source and simple ones as the target
- phrase alignments generated by Giza++ are modified to accommodate null phrasal alignment to allow **deletion** with **reordering** and **substitution** [Coster and Kauchak, 2011]
- post-hoc reranking to rank the n-best output based on their dissimilarity from the source [Wubben et al., 2012]

Sentence Simplification in Literature

Simplification as monolingual translation
[Coster and Kauchak, 2011, Wubben et al., 2012, Xu et al., 2016]

- learned **statistical machine translation system** considering complex sentences as the source and simple ones as the target
- phrase alignments generated by Giza++ are modified to accommodate null phrasal alignment to allow **deletion** with **reordering** and **substitution** [Coster and Kauchak, 2011]
- post-hoc reranking to rank the n-best output based on their dissimilarity from the source [Wubben et al., 2012]
- complete adaptation of a syntax-based machine translation framework to perform simplification [Xu et al., 2016]

Sentence Simplification in Literature

Simplification as monolingual translation
[Coster and Kauchak, 2011, Wubben et al., 2012, Xu et al., 2016]

- learned **statistical machine translation system** considering complex sentences as the source and simple ones as the target
- phrase alignments generated by Giza++ are modified to accommodate null phrasal alignment to allow **deletion** with **reordering** and **substitution** [Coster and Kauchak, 2011]
- post-hoc reranking to rank the n-best output based on their dissimilarity from the source [Wubben et al., 2012]
- complete adaptation of a syntax-based machine translation framework to perform simplification [Xu et al., 2016]
- does not allow **split**s

Sentence Simplification in Literature

Hybrid symbolic/statistical approach
[Siddharthan and Mandya, 2014] combining

- **hand-crafted syntactic simplification rules**, and
- **lexical simplification rules** extracted from aligned English and simple English sentences, and revision histories of Simple Wikipedia.

Sentence Simplification in Literature

Hybrid symbolic/statistical approach
[Siddharthan and Mandya, 2014] combining

- **hand-crafted syntactic simplification rules**, and
- **lexical simplification rules** extracted from aligned English and simple English sentences, and revision histories of Simple Wikipedia.

Hybrid simplification [Narayan and Gardent, 2014] using

- **deep semantic representation** for **splitting** and **deletion**, with
- a **monolingual machine translation** module which handles **reordering** and **substitution**.

Our Simplification Proposal

An **unsupervised framework** relying only on the **comparable wikipedia corpora**

- pipelines three dedicated modules for **lexical simplification**, **sentence splitting** and **sentence compression**

Our Simplification Proposal

An **unsupervised framework** relying only on the **comparable wikipedia corpora**

- pipelines three dedicated modules for **lexical simplification**, **sentence splitting** and **sentence compression**

Exploits **deep semantic representation** for **splitting** and **deletion** probabilities

- Discourse representation structures (**DRS**) assigned by Boxer as input
- facilitates completion (the re-creation of the shared element in the split sentences)
- provide a natural means to avoid deleting obligatory arguments

Outline of the talk

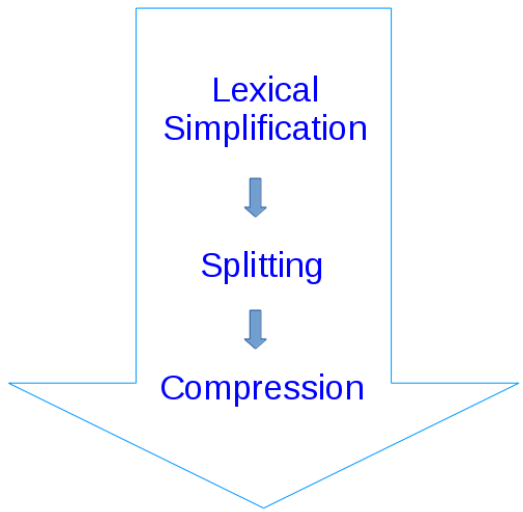
Outline of the talk

Sentence Simplification Framework

In 1964 Peter Higgs published his second paper in Physical Review Letters describing Higgs mechanism which predicted a new massive spin-zero boson for the first time.

Sentence Simplification Framework

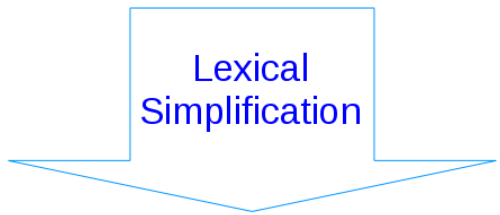
In 1964 Peter Higgs published his second paper in Physical Review Letters describing Higgs mechanism which predicted a new massive spin-zero boson for the first time.



In 1964 Peter Higgs wrote his paper explaining Higgs mechanism. Higgs mechanism predicted a new elementary particle.

Context-Aware Lexical Simplification

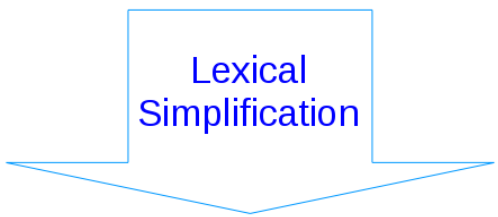
In 1964 Peter Higgs published his second paper in Physical Review Letters describing Higgs mechanism which predicted a new massive spin-zero boson for the first time.



In 1964 Peter Higgs wrote his second paper in Physical Review Letters explaining Higgs mechanism which predicted a new massive elementary particle for the first time.

Context-Aware Lexical Simplification

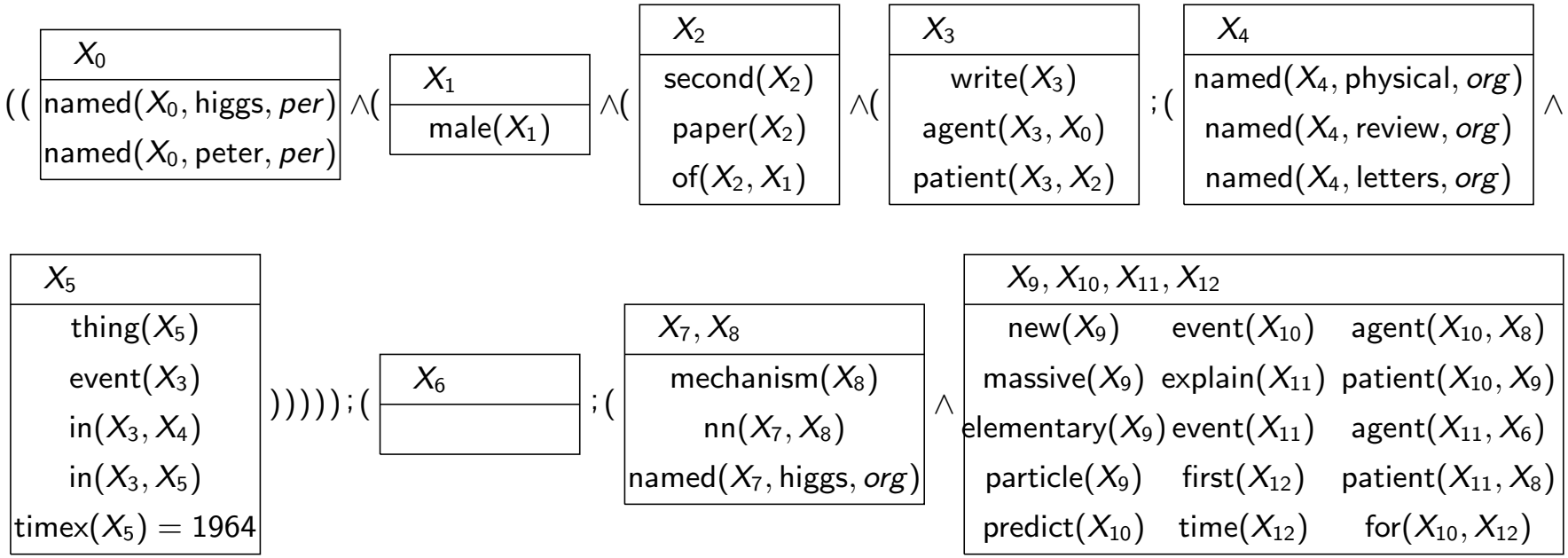
In 1964 Peter Higgs published his second paper in Physical Review Letters describing Higgs mechanism which predicted a new massive spin-zero boson for the first time.



- **Context-aware rules** are learned from **comparable corpora**: English and Simple English Wikipedia [Biran et al., 2011]
- **Viterbi decoding** using highly weighted rules maximising simple sentence probability

In 1964 Peter Higgs wrote his second paper in Physical Review Letters explaining Higgs mechanism which predicted a new massive elementary particle for the first time.

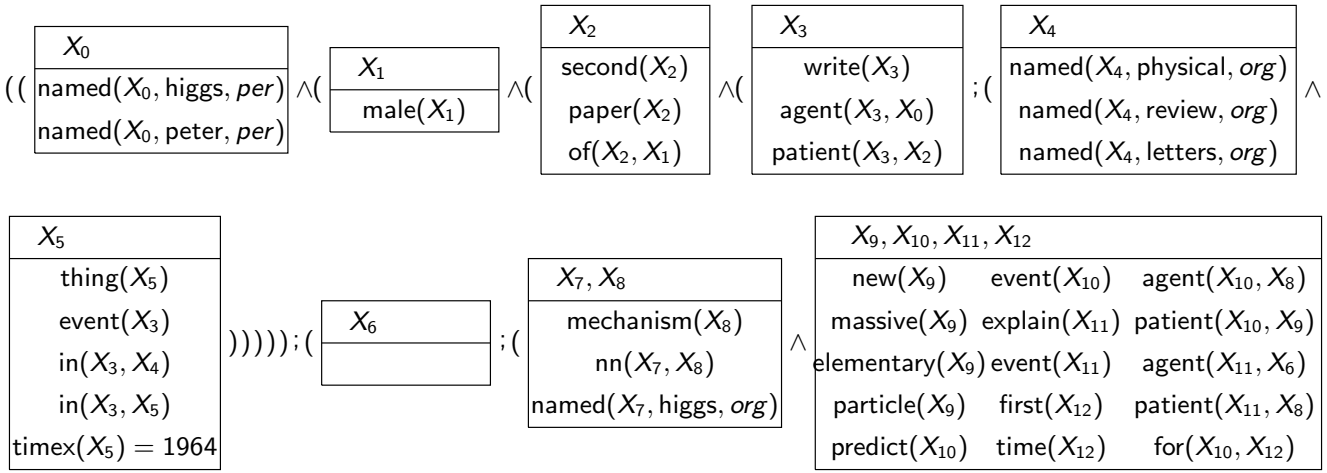
DRS for Semantic Representation



Discourse Representation Structure of

"In 1964 Peter Higgs wrote his second paper in Physical Review Letters explaining Higgs mechanism which predicted a new massive elementary particle for the first time."

DRS for Semantic Representation



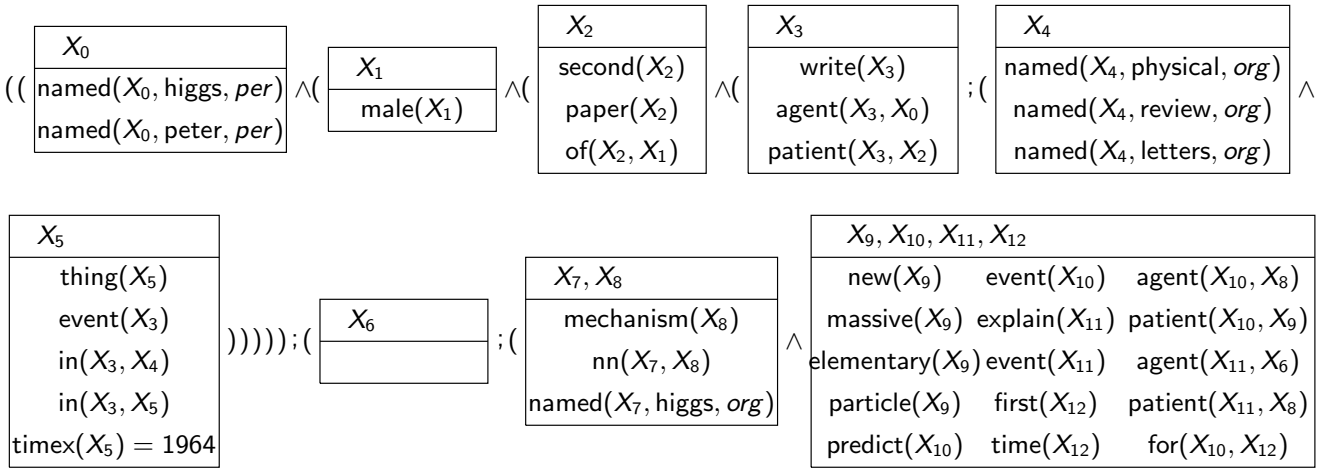
Discourse Representation Structure of

"In 1964 Peter Higgs wrote his second paper in Physical Review Letters explaining Higgs mechanism which predicted a new massive elementary particle for the first time."

node	pos.	in S	predicate/type

rel	pos.	in S	predicate

DRS for Semantic Representation



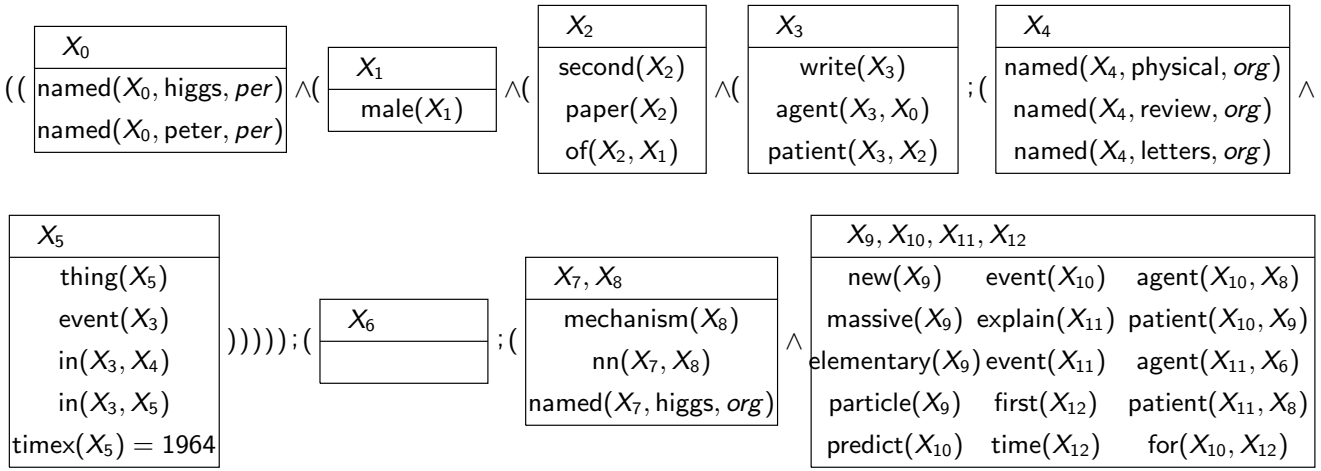
Discourse Representation Structure of

"In 1964 Peter Higgs wrote his second paper in Physical Review Letters explaining Higgs mechanism which predicted a new massive elementary particle for the first time."

node	pos. in S	predicate/type
X_0	3, 4	higgs/per, peter/per
X_1	6	male/a
X_2	6, 7, 8	second/a, paper/a
X_3	5	write/v, **event**
X_4	10, 11, 12	physical/org review/org, letters/org
X_5	2	thing/n, 1964
X_6	6, 7, 8	−−
X_7	14	higgs/org
X_8	14, 15	mechanism/n
X_9	18, 19, 20 21, 22	new/a, elementary/a massive/a, particle/n
X_{10}	17	predict/v, **event**
X_{11}	13	explain/v, **event**
X_{12}	24, 25, 26	first/a, time/n

rel	pos. in S	predicate

DRS for Semantic Representation



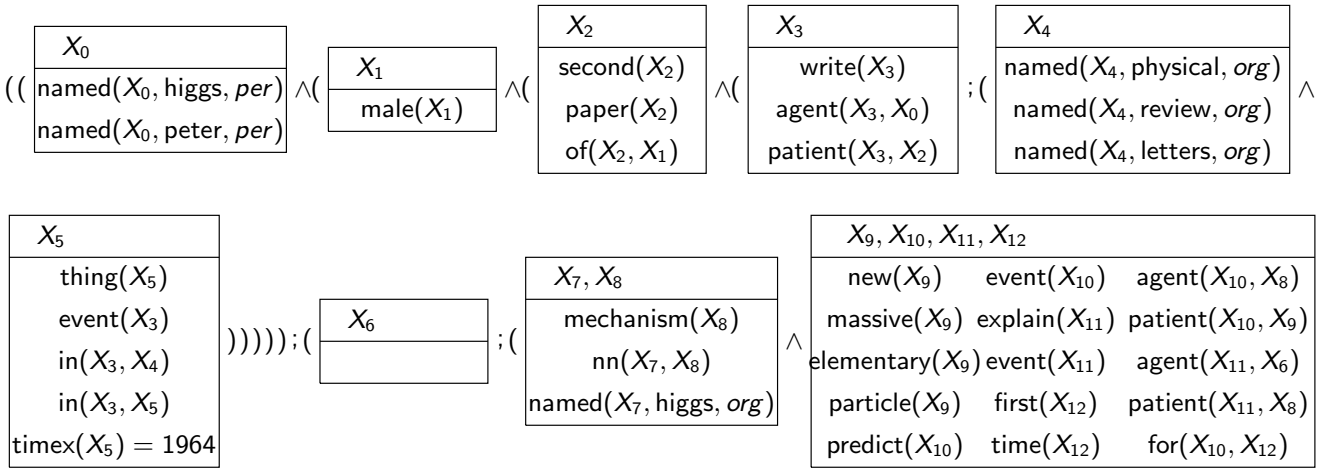
Discourse Representation Structure of

"In 1964 Peter Higgs wrote his second paper in Physical Review Letters explaining Higgs mechanism which predicted a new massive elementary particle for the first time."

node	pos. in S	predicate/type
X_0	3, 4	higgs/per, peter/per
X_1	6	male/a
X_2	6, 7, 8	second/a, paper/a
X_3	5	write/v, **event**
X_4	10, 11, 12	physical/org review/org, letters/org
X_5	2	thing/n, 1964
X_6	6, 7, 8	−−
X_7	14	higgs/org
X_8	14, 15	mechanism/n
X_9	18, 19, 20 21, 22	new/a, elementary/a massive/a, particle/n
X_{10}	17	predict/v, **event**
X_{11}	13	explain/v, **event**
X_{12}	24, 25, 26	first/a, time/n

rel	pos. in S	predicate
R_1	5	*agent*, $X_3 \rightarrow X_0$
R_2	5	*patient*, $X_3 \rightarrow X_2$
R_3	6	*of*, $X_2 \rightarrow X_1$
R_4	9	*in*, $X_3 \rightarrow X_4$
R_5	1	*in*, $X_3 \rightarrow X_5$
R_6	13	*agent*, $X_{11} \rightarrow X_6$
R_7	13	*patient*, $X_{11} \rightarrow X_8$
R_8	−−	*nn*, $X_8 \rightarrow X_7$
R_9	17	*agent*, $X_{10} \rightarrow X_8$
R_{10}	17	*patient*, $X_{10} \rightarrow X_9$
R_{11}	23	*for*, $X_{10} \rightarrow X_{12}$

DRS for Semantic Representation



Discourse Representation Structure of

"In 1964 Peter Higgs wrote his second paper in Physical Review Letters explaining Higgs mechanism which predicted a new massive elementary particle for the first time."

node	pos. in S	predicate/type
X_0	3, 4	higgs/per, peter/per
X_1	6	male/a
X_2	6, 7, 8	second/a, paper/a
X_3	5	write/v, **event**
X_4	10, 11, 12	physical/org review/org, letters/org
X_5	2	thing/n, 1964
X_6	6, 7, 8	−−
X_7	14	higgs/org
X_8	14, 15	mechanism/n
X_9	18, 19, 20 21, 22	new/a, elementary/a massive/a, particle/n
X_{10}	17	predict/v, **event**
X_{11}	13	explain/v, **event**
X_{12}	24, 25, 26	first/a, time/n
O_1	16	which/WDT

rel	pos. in S	predicate
R_1	5	*agent*, $X_3 \rightarrow X_0$
R_2	5	*patient*, $X_3 \rightarrow X_2$
R_3	6	*of*, $X_2 \rightarrow X_1$
R_4	9	*in*, $X_3 \rightarrow X_4$
R_5	1	*in*, $X_3 \rightarrow X_5$
R_6	13	*agent*, $X_{11} \rightarrow X_6$
R_7	13	*patient*, $X_{11} \rightarrow X_8$
R_8	−−	*nn*, $X_8 \rightarrow X_7$
R_9	17	*agent*, $X_{10} \rightarrow X_8$
R_{10}	17	*patient*, $X_{10} \rightarrow X_9$
R_{11}	23	*for*, $X_{10} \rightarrow X_{12}$

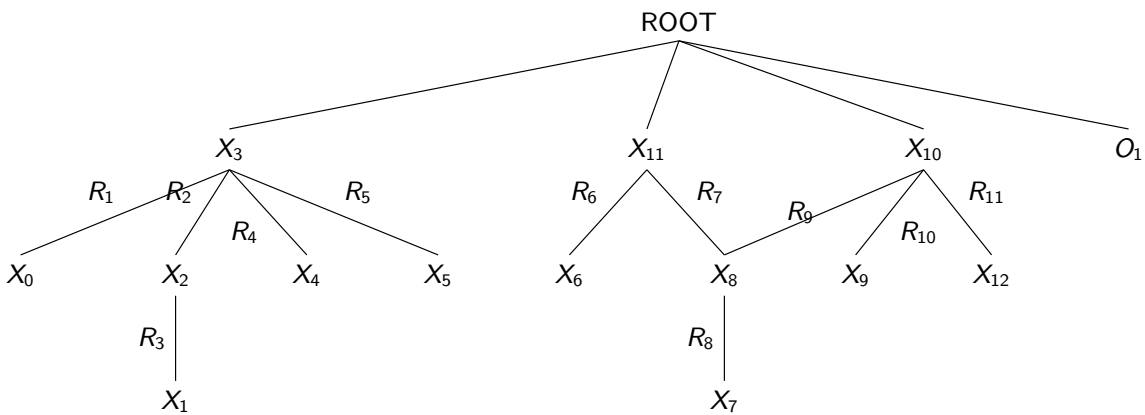
DRS for Semantic Representation

$$((\ \boxed{\begin{array}{|l}\hline X_0 \\ \hline \text{named}(X_0, \text{higgs}, \textit{per}) \\ \text{named}(X_0, \text{peter}, \textit{per}) \\ \hline\end{array}}\ \wedge(\ \boxed{\begin{array}{|l}\hline X_1 \\ \hline \text{male}(X_1) \\ \hline\end{array}}\ \wedge(\ \boxed{\begin{array}{|l}\hline X_2 \\ \hline \text{second}(X_2) \\ \text{paper}(X_2) \\ \text{of}(X_2, X_1) \\ \hline\end{array}}\ \wedge(\ \boxed{\begin{array}{|l}\hline X_3 \\ \hline \text{write}(X_3) \\ \text{agent}(X_3, X_0) \\ \text{patient}(X_3, X_2) \\ \hline\end{array}}\ ;(\ \boxed{\begin{array}{|l}\hline X_4 \\ \hline \text{named}(X_4, \text{physical}, \textit{org}) \\ \text{named}(X_4, \text{review}, \textit{org}) \\ \text{named}(X_4, \text{letters}, \textit{org}) \\ \hline\end{array}}\ \wedge$$

$$\boxed{\begin{array}{|l}\hline X_5 \\ \hline \text{thing}(X_5) \\ \text{event}(X_3) \\ \text{in}(X_3, X_4) \\ \text{in}(X_3, X_5) \\ \text{timex}(X_5) = 1964 \\ \hline\end{array}}\ ))))));(\ \boxed{\begin{array}{|l}\hline X_6 \\ \hline \\ \hline\end{array}}\ ;(\ \boxed{\begin{array}{|l}\hline X_7, X_8 \\ \hline \text{mechanism}(X_8) \\ \text{nn}(X_7, X_8) \\ \text{named}(X_7, \text{higgs}, \textit{org}) \\ \hline\end{array}}\ \wedge\ \boxed{\begin{array}{|lll}\hline X_9, X_{10}, X_{11}, X_{12} & & \\ \hline \text{new}(X_9) & \text{event}(X_{10}) & \text{agent}(X_{10}, X_8) \\ \text{massive}(X_9) & \text{explain}(X_{11}) & \text{patient}(X_{10}, X_9) \\ \text{elementary}(X_9) & \text{event}(X_{11}) & \text{agent}(X_{11}, X_6) \\ \text{particle}(X_9) & \text{first}(X_{12}) & \text{patient}(X_{11}, X_8) \\ \text{predict}(X_{10}) & \text{time}(X_{12}) & \text{for}(X_{10}, X_{12}) \\ \hline\end{array}}$$

Discourse Representation Structure of

"In 1964 Peter Higgs wrote his second paper in Physical Review Letters explaining Higgs mechanism which predicted a new massive elementary particle for the first time."

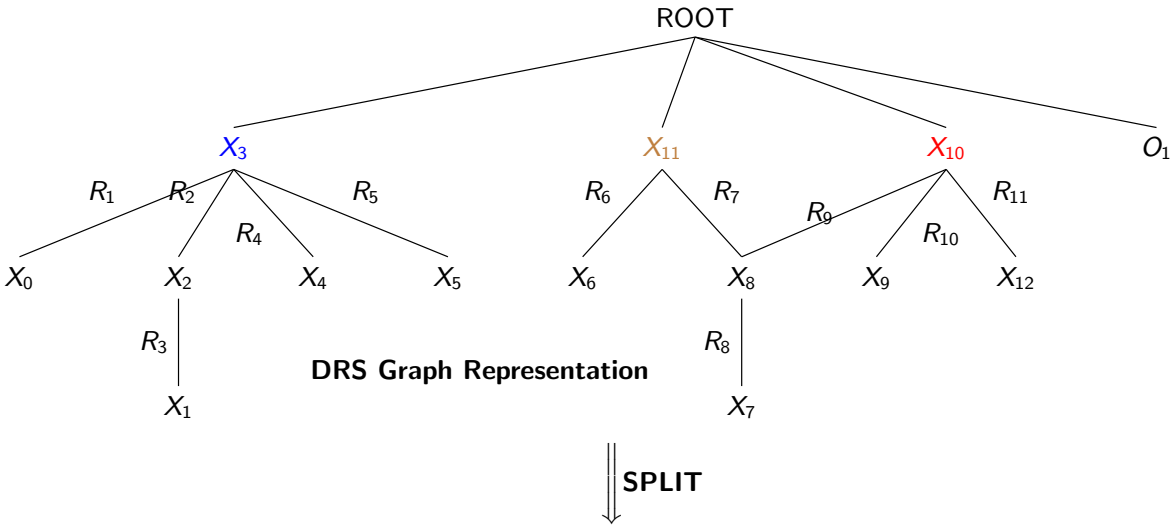


DRS Graph Representation

node	pos. in S	predicate/type
X_0	3, 4	higgs/per, peter/per
X_1	6	male/a
X_2	6, 7, 8	second/a, paper/a
X_3	5	write/v, **event**
X_4	10, 11, 12	physical/org review/org, letters/org
X_5	2	thing/n, 1964
X_6	6, 7, 8	−−
X_7	14	higgs/org
X_8	14, 15	mechanism/n
X_9	18, 19, 20 21, 22	new/a, elementary/a massive/a, particle/n
X_{10}	17	predict/v, **event**
X_{11}	13	explain/v, **event**
X_{12}	24, 25, 26	first/a, time/n
O_1	16	which/WDT

rel	pos. in S	predicate
R_1	5	$\textit{agent}, X_3 \rightarrow X_0$
R_2	5	$\textit{patient}, X_3 \rightarrow X_2$
R_3	6	$\textit{of}, X_2 \rightarrow X_1$
R_4	9	$\textit{in}, X_3 \rightarrow X_4$
R_5	1	$\textit{in}, X_3 \rightarrow X_5$
R_6	13	$\textit{agent}, X_{11} \rightarrow X_6$
R_7	13	$\textit{patient}, X_{11} \rightarrow X_8$
R_8	−−	$\textit{nn}, X_8 \rightarrow X_7$
R_9	17	$\textit{agent}, X_{10} \rightarrow X_8$
R_{10}	17	$\textit{patient}, X_{10} \rightarrow X_9$
R_{11}	23	$\textit{for}, X_{10} \rightarrow X_{12}$

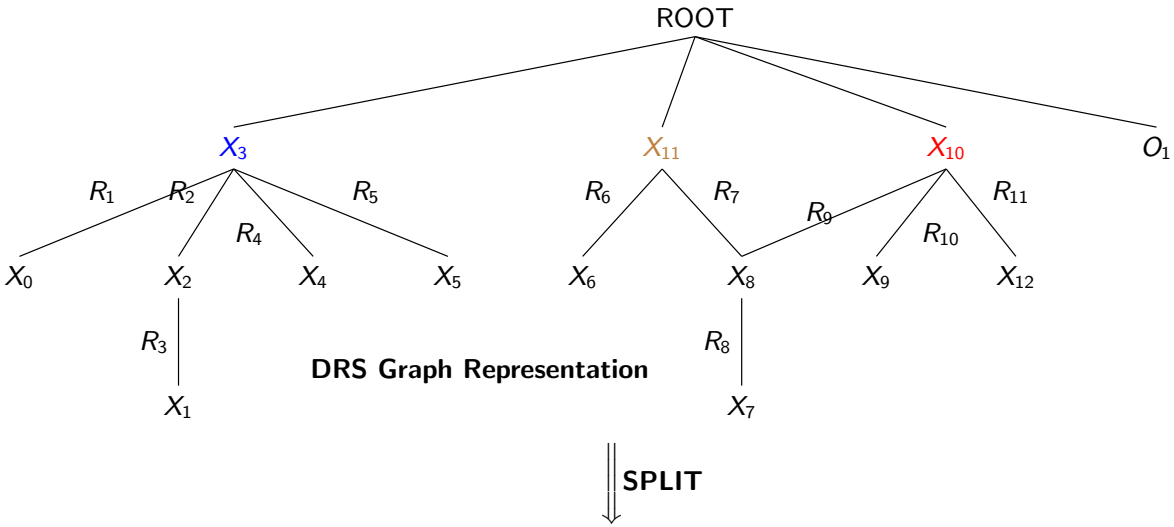
Sentence Splitting



node	pos. in S	predicate/type
X_0	3, 4	higgs/per, peter/per
X_1	6	male/a
X_2	6, 7, 8	second/a, paper/a
X_3	5	write/v, **event**
X_4	10, 11, 12	physical/org review/org, letters/org
X_5	2	thing/n, 1964
X_6	6, 7, 8	−−
X_7	14	higgs/org
X_8	14, 15	mechanism/n
X_9	18, 19, 20 21, 22	new/a, elementary/a massive/a, particle/n
X_{10}	17	predict/v, **event**
X_{11}	13	explain/v, **event**
X_{12}	24, 25, 26	first/a, time/n
O_1	16	which/WDT

rel	pos. in S	predicate
R_1	5	$agent, X_3 \rightarrow X_0$
R_2	5	$patient, X_3 \rightarrow X_2$
R_3	6	$of, X_2 \rightarrow X_1$
R_4	9	$in, X_3 \rightarrow X_4$
R_5	1	$in, X_3 \rightarrow X_5$
R_6	13	$agent, X_{11} \rightarrow X_6$
R_7	13	$patient, X_{11} \rightarrow X_8$
R_8	−−	$nn, X_8 \rightarrow X_7$
R_9	17	$agent, X_{10} \rightarrow X_8$
R_{10}	17	$patient, X_{10} \rightarrow X_9$
R_{11}	23	$for, X_{10} \rightarrow X_{12}$

Sentence Splitting

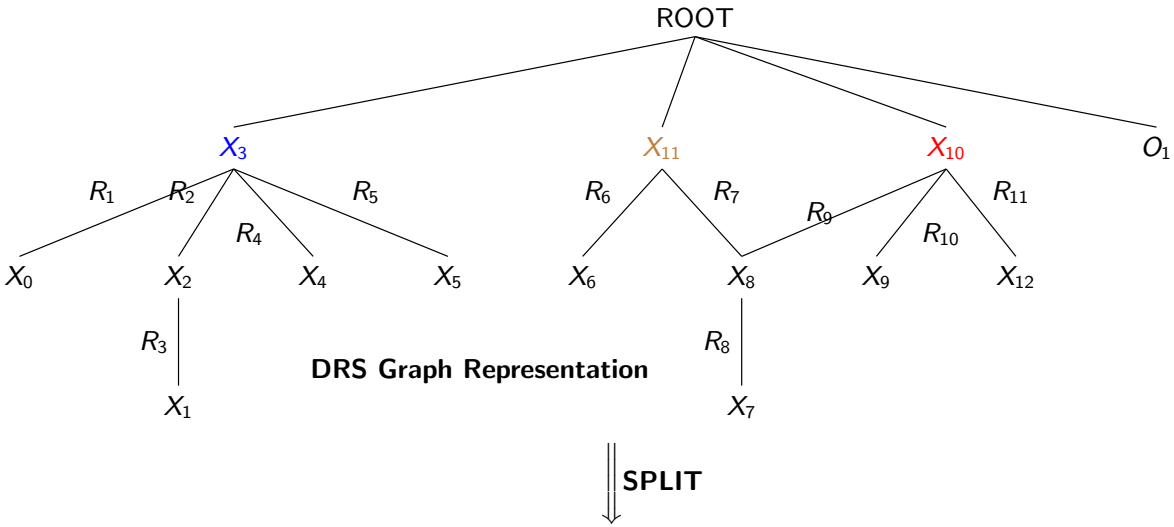


node	pos. in S	predicate/type
X_0	3, 4	higgs/per, peter/per
X_1	6	male/a
X_2	6, 7, 8	second/a, paper/a
X_3	5	write/v, **event**
X_4	10, 11, 12	physical/org review/org, letters/org
X_5	2	thing/n, 1964
X_6	6, 7, 8	−−
X_7	14	higgs/org
X_8	14, 15	mechanism/n
X_9	18, 19, 20 21, 22	new/a, elementary/a massive/a, particle/n
X_{10}	17	predict/v, **event**
X_{11}	13	explain/v, **event**
X_{12}	24, 25, 26	first/a, time/n
O_1	16	which/WDT

rel	pos. in S	predicate
R_1	5	$agent, X_3 \rightarrow X_0$
R_2	5	$patient, X_3 \rightarrow X_2$
R_3	6	$of, X_2 \rightarrow X_1$
R_4	9	$in, X_3 \rightarrow X_4$
R_5	1	$in, X_3 \rightarrow X_5$
R_6	13	$agent, X_{11} \rightarrow X_6$
R_7	13	$patient, X_{11} \rightarrow X_8$
R_8	−−	$nn, X_8 \rightarrow X_7$
R_9	17	$agent, X_{10} \rightarrow X_8$
R_{10}	17	$patient, X_{10} \rightarrow X_9$
R_{11}	23	$for, X_{10} \rightarrow X_{12}$

Semantic Patterns in Simple Wikipedia	**prob.**
$\langle$(*agent*, *patient*)$\rangle$	0.059
$\langle$(*agent*, *patient*), (*agent*, *in*, *in*, *patient*)$\rangle$	0.023
$\langle$(*agent*, *patient*), (*agent*, *in*, *in*, *patient*),(*agent*, *for*, *patient*)$\rangle$	0.003

Sentence Splitting



$$S \rightarrow \{s_1, s_2, \ldots, s_n\}$$

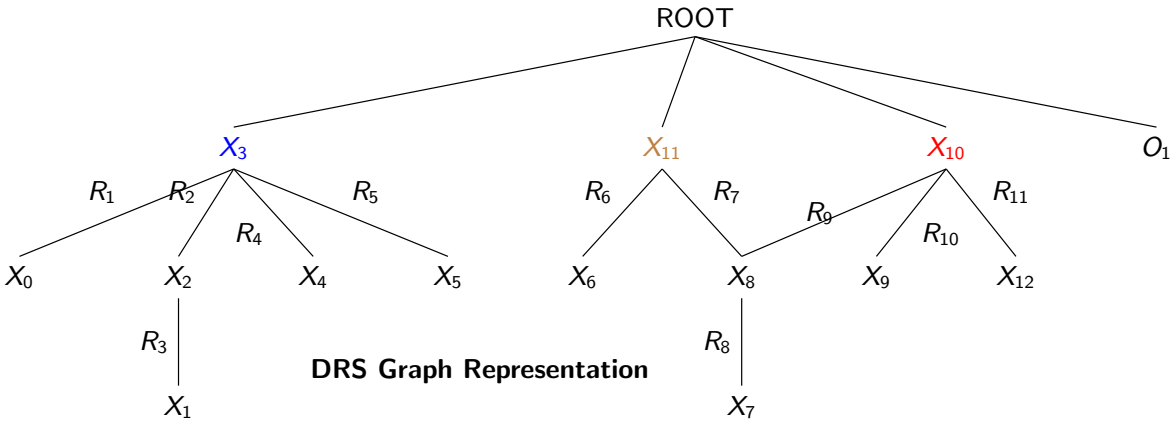
$$P_{split} = \frac{1}{n} \sum_{s_i} \alpha_{s_i} \times lm_{s_i} \times SFT_{s_i}$$

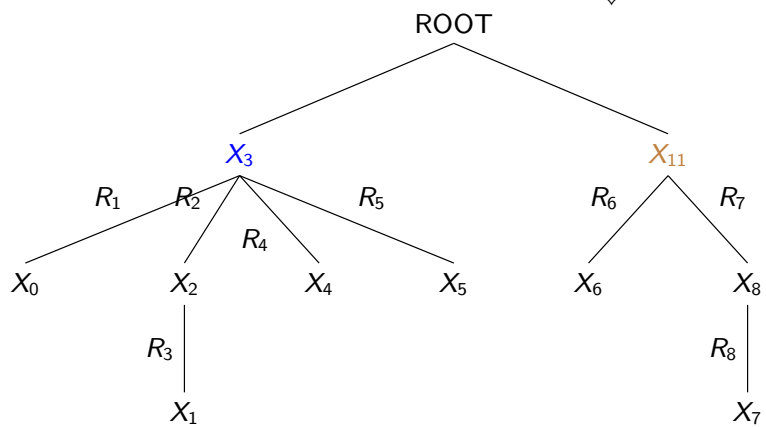
Semantic Patterns in Simple Wikipedia	**prob.**
⟨(*agent*, *patient*)⟩	0.059
⟨(*agent*, *patient*), (*agent*, *in*, *in*, *patient*)⟩	0.023
⟨(*agent*, *patient*), (*agent*, *in*, *in*, *patient*),(*agent*, *for*, *patient*)⟩	0.003

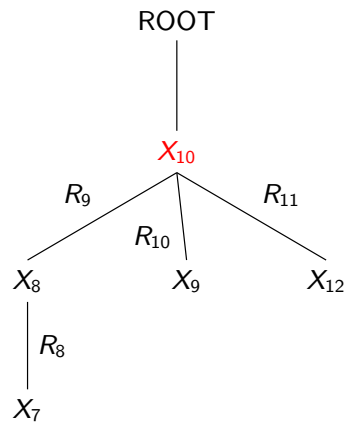
node	pos. in S	predicate/type
X_0	3, 4	higgs/per, peter/per
X_1	6	male/a
X_2	6, 7, 8	second/a, paper/a
X_3	5	write/v, **event**
X_4	10, 11, 12	physical/org review/org, letters/org
X_5	2	thing/n, 1964
X_6	6, 7, 8	−−
X_7	14	higgs/org
X_8	14, 15	mechanism/n
X_9	18, 19, 20 21, 22	new/a, elementary/a massive/a, particle/n
X_{10}	17	predict/v, **event**
X_{11}	13	explain/v, **event**
X_{12}	24, 25, 26	first/a, time/n
O_1	16	which/WDT

rel	pos. in S	predicate
R_1	5	$agent, X_3 \rightarrow X_0$
R_2	5	$patient, X_3 \rightarrow X_2$
R_3	6	$of, X_2 \rightarrow X_1$
R_4	9	$in, X_3 \rightarrow X_4$
R_5	1	$in, X_3 \rightarrow X_5$
R_6	13	$agent, X_{11} \rightarrow X_6$
R_7	13	$patient, X_{11} \rightarrow X_8$
R_8	−−	$nn, X_8 \rightarrow X_7$
R_9	17	$agent, X_{10} \rightarrow X_8$
R_{10}	17	$patient, X_{10} \rightarrow X_9$
R_{11}	23	$for, X_{10} \rightarrow X_{12}$

Sentence Splitting







"In 1964 Peter Higgs wrote his second paper in Physical Review Letters explaining Higgs mechanism "

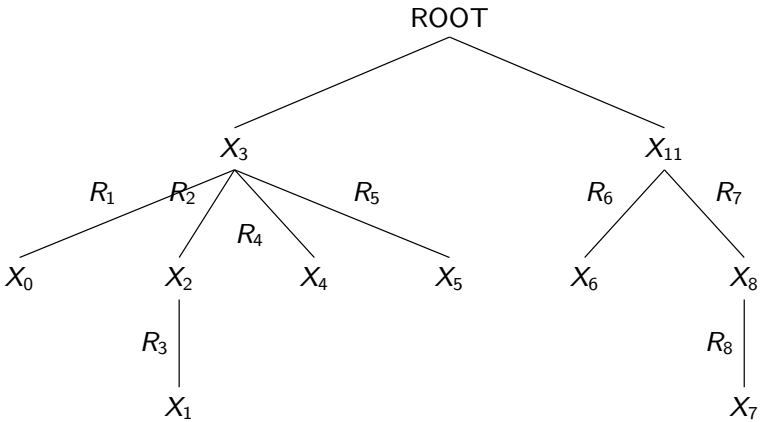
"Higgs mechanism predicted a new massive elementary particle for the first time"

node	pos. in S	predicate/type
X_0	3, 4	higgs/per, peter/per
X_1	6	male/a
X_2	6, 7, 8	second/a, paper/a
X_3	5	write/v, **event**
X_4	10, 11, 12	physical/org review/org, letters/org
X_5	2	thing/n, 1964
X_6	6, 7, 8	−−
X_7	14	higgs/org
X_8	14, 15	mechanism/n
X_9	18, 19, 20 21, 22	new/a, elementary/a massive/a, particle/n
X_{10}	17	predict/v, **event**
X_{11}	13	explain/v, **event**
X_{12}	24, 25, 26	first/a, time/n
O_1	16	which/WDT

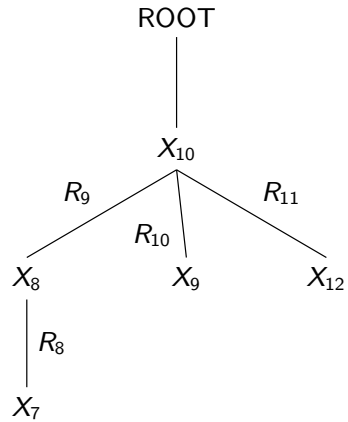
rel	pos. in S	predicate
R_1	5	$agent, X_3 \rightarrow X_0$
R_2	5	$patient, X_3 \rightarrow X_2$
R_3	6	$of, X_2 \rightarrow X_1$
R_4	9	$in, X_3 \rightarrow X_4$
R_5	1	$in, X_3 \rightarrow X_5$
R_6	13	$agent, X_{11} \rightarrow X_6$
R_7	13	$patient, X_{11} \rightarrow X_8$
R_8	−−	$nn, X_8 \rightarrow X_7$
R_9	17	$agent, X_{10} \rightarrow X_8$
R_{10}	17	$patient, X_{10} \rightarrow X_9$
R_{11}	23	$for, X_{10} \rightarrow X_{12}$

Semantic Patterns in Simple Wikipedia	**prob.**
⟨(*agent*, *patient*)⟩	0.059
⟨(*agent*, *patient*), (*agent*, *in*, *in*, *patient*)⟩	0.023
⟨(*agent*, *patient*), (*agent*, *in*, *in*, *patient*),(*agent*, *for*, *patient*)⟩	0.003

Sentence Compression



"In 1964 Peter Higgs wrote his second paper in Physical Review Letters explaining Higgs mechanism "



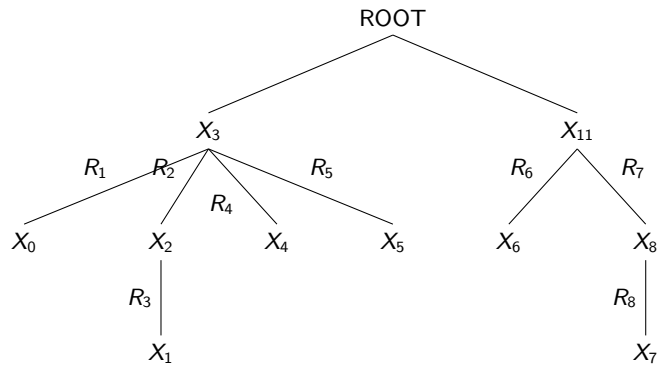
"Higgs mechanism predicted a new massive elementary particle for the first time"

⇓ **COMPRESSION**

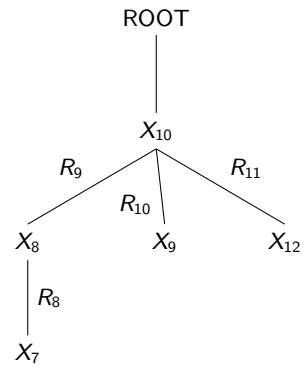
node	pos. in S	predicate/type
X_0	3, 4	higgs/per, peter/per
X_1	6	male/a
X_2	6, 7, 8	second/a, paper/a
X_3	5	write/v, **event**
X_4	10, 11, 12	physical/org review/org, letters/org
X_5	2	thing/n, 1964
X_6	6, 7, 8	−−
X_7	14	higgs/org
X_8	14, 15	mechanism/n
X_9	18, 19, 20 21, 22	new/a, elementary/a massive/a, particle/n
X_{10}	17	predict/v, **event**
X_{11}	13	explain/v, **event**
X_{12}	24, 25, 26	first/a, time/n
O_1	16	which/WDT

rel	pos. in S	predicate
R_1	5	$agent, X_3 \rightarrow X_0$
R_2	5	$patient, X_3 \rightarrow X_2$
R_3	6	$of, X_2 \rightarrow X_1$
R_4	9	$in, X_3 \rightarrow X_4$
R_5	1	$in, X_3 \rightarrow X_5$
R_6	13	$agent, X_{11} \rightarrow X_6$
R_7	13	$patient, X_{11} \rightarrow X_8$
R_8	−−	$nn, X_8 \rightarrow X_7$
R_9	17	$agent, X_{10} \rightarrow X_8$
R_{10}	17	$patient, X_{10} \rightarrow X_9$
R_{11}	23	$for, X_{10} \rightarrow X_{12}$

Sentence Compression



"In 1964 Peter Higgs wrote his second paper in Physical Review Letters explaining Higgs mechanism "



"Higgs mechanism predicted a new massive elementary particle for the first time"

COMPRESSION

node	pos. in S	predicate/type
X_0	3, 4	higgs/per, peter/per
X_1	6	male/a
X_2	6, 7, 8	second/a, paper/a
X_3	5	write/v, **event**
X_4	10, 11, 12	physical/org review/org, letters/org
X_5	2	thing/n, 1964
X_6	6, 7, 8	−−
X_7	14	higgs/org
X_8	14, 15	mechanism/n
X_9	18, 19, 20 21, 22	new/a, elementary/a massive/a, particle/n
X_{10}	17	predict/v, **event**
X_{11}	13	explain/v, **event**
X_{12}	24, 25, 26	first/a, time/n
O_1	16	which/WDT

$$\sum_x x^r_{h,w} \times P(r|h) \times P(w) \quad r \notin \{agent, patient, theme, eq\}$$

$$x^r_{h,w} = \begin{cases} 1, & \text{if } r \text{ is preserved} \\ 0, & \text{otherwise} \end{cases}$$

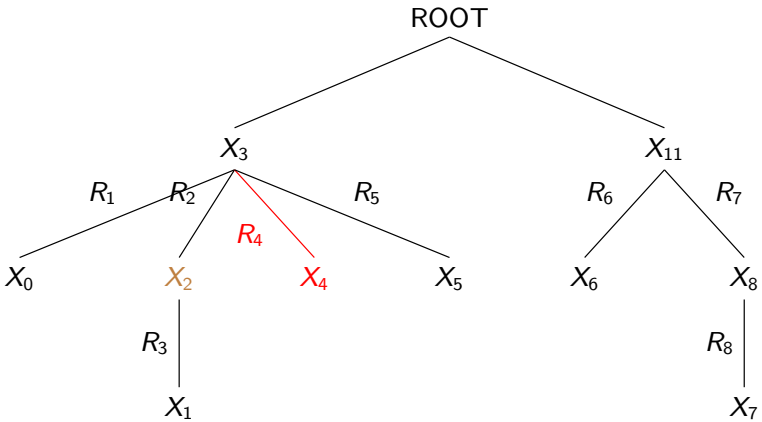
$P(r|h)$: Conditional probability of relation r given head label h

$P(w)$: Relative frequency of w in Simple Wikipedia

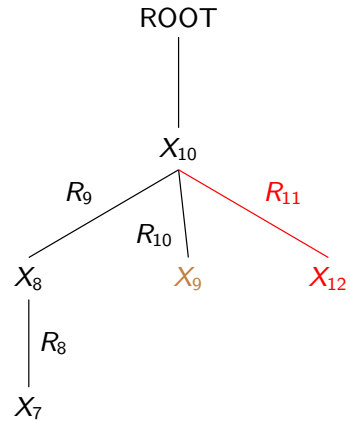
rel	pos. in S	predicate
R_1	5	$agent, X_3 \rightarrow X_0$
R_2	5	$patient, X_3 \rightarrow X_2$
R_3	6	$of, X_2 \rightarrow X_1$
R_4	9	$in, X_3 \rightarrow X_4$
R_5	1	$in, X_3 \rightarrow X_5$
R_6	13	$agent, X_{11} \rightarrow X_6$
R_7	13	$patient, X_{11} \rightarrow X_8$
R_8	−−	$nn, X_8 \rightarrow X_7$
R_9	17	$agent, X_{10} \rightarrow X_8$
R_{10}	17	$patient, X_{10} \rightarrow X_9$
R_{11}	23	$for, X_{10} \rightarrow X_{12}$

Optimized using ILP, constrained to well-formedness

Sentence Compression



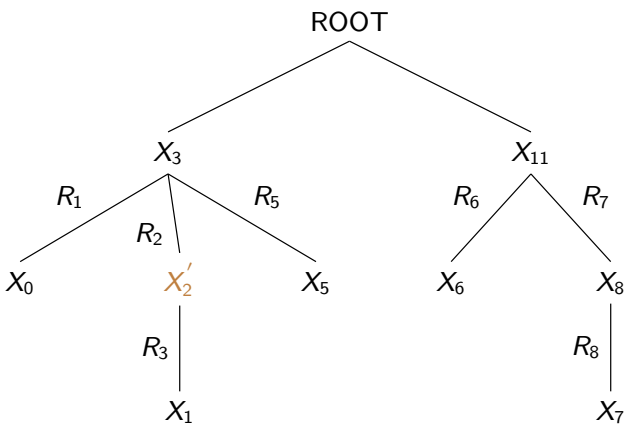
"*In 1964 Peter Higgs wrote his second paper in Physical Review Letters explaining Higgs mechanism* "



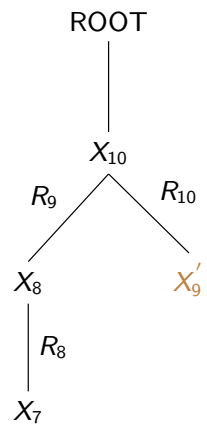
"*Higgs mechanism predicted a new massive elementary particle for the first time*"

node	pos. in S	predicate/type
X_0	3, 4	higgs/per, peter/per
X_1	6	male/a
X_2	6, 7, 8	second/a, paper/a
X_3	5	write/v, **event**
X_4	10, 11, 12	physical/org review/org, letters/org
X_5	2	thing/n, 1964
X_6	6, 7, 8	−−
X_7	14	higgs/org
X_8	14, 15	mechanism/n
X_9	18, 19, 20 21, 22	new/a, elementary/a massive/a, particle/n
X_{10}	17	predict/v, **event**
X_{11}	13	explain/v, **event**
X_{12}	24, 25, 26	first/a, time/n
O_1	16	which/WDT

COMPRESSION



"*In 1964 Peter Higgs wrote his paper explaining Higgs mechanism*"



"*Higgs mechanism predicted a new particle*"

rel	pos. in S	predicate
R_1	5	$agent, X_3 \rightarrow X_0$
R_2	5	$patient, X_3 \rightarrow X_2$
R_3	6	$of, X_2 \rightarrow X_1$
R_4	9	$in, X_3 \rightarrow X_4$
R_5	1	$in, X_3 \rightarrow X_5$
R_6	13	$agent, X_{11} \rightarrow X_6$
R_7	13	$patient, X_{11} \rightarrow X_8$
R_8	−−	$nn, X_8 \rightarrow X_7$
R_9	17	$agent, X_{10} \rightarrow X_8$
R_{10}	17	$patient, X_{10} \rightarrow X_9$
R_{11}	23	$for, X_{10} \rightarrow X_{12}$

Sentence Simplification Framework: Summary

In 1964 Peter Higgs published his second paper in Physical Review Letters describing Higgs mechanism which predicted a new massive spin-zero boson for the first time.

⇓ **Context-aware Lexical Simplification**

⇓ **Resource:** Comparable English and Simple Wikipedia

In 1964 Peter Higgs wrote his second paper in Physical Review Letters explaining Higgs mechanism which predicted a new massive elementary particle for the first time.

Sentence Simplification Framework: Summary

In 1964 Peter Higgs published his second paper in Physical Review Letters describing Higgs mechanism which predicted a new massive spin-zero boson for the first time.

⇓ **Context-aware Lexical Simplification**

⇓ **Resource:** Comparable English and Simple Wikipedia

In 1964 Peter Higgs wrote his second paper in Physical Review Letters explaining Higgs mechanism which predicted a new massive elementary particle for the first time.

⇓ **Syntactic Simplification using Deep Semantic Representation**
Sentence Splitting and Compression

⇓ **Resource:** Simple Wikipedia

In 1964 Peter Higgs wrote his paper explaining Higgs mechanism.
Higgs mechanism predicted a new elementary particle.

Outline of the talk

Outline of the talk

Evaluations

- Evaluated on the Zhu testset consisting of 100/131 complex/simple sentences [Zhu et al., 2010]

Evaluations

- Evaluated on the Zhu testset consisting of 100/131 complex/simple sentences [Zhu et al., 2010]

- compared against 4 supervised methods
  - Zhu's *Monolingual tree-based translation model* [Zhu et al., 2010]
  - Woodsend's *Quasi-synchronous grammar and integer linear programming* [Woodsend and Lapata, 2011]
  - Wubben's *Simplification as monolingual translation* [Wubben et al., 2012]
  - Narayan's *Hybrid Simplification with Deep Semantics and SMT* [Narayan and Gardent, 2014]

Evaluations

- Evaluated on the Zhu testset consisting of 100/131 complex/simple sentences [Zhu et al., 2010]

- compared against 4 supervised methods
  - Zhu's *Monolingual tree-based translation model* [Zhu et al., 2010]
  - Woodsend's *Quasi-synchronous grammar and integer linear programming* [Woodsend and Lapata, 2011]
  - Wubben's *Simplification as monolingual translation* [Wubben et al., 2012]
  - Narayan's *Hybrid Simplification with Deep Semantics and SMT* [Narayan and Gardent, 2014]

- Evaluation metrics
  - Automatic evaluation metrics: BLEU score, Edit distance, and Split
  - Human evaluation metrics: Simplicity, Fluency, and Adequacy

Automatic evaluation

System	BLEU	Edit distance to complex sentences		Edit distance to simple sentences		Sentences with splits
		LD	No edit	LD	No edit	
GOLD	100	12.24	3	0	100	28
Zhu	37.4	7.87	2	14.64	0	80
Woodsend	42	8.63	24	16.03	2	63
Wubben	41.4	3.33	6	13.57	2	1
Narayan	**53.6**	6.32	4	11.53	3	10
Unsup	38.5	6.75	3	14.29	0	49

Table: BLEU score, Average Levenshtein distance (LD) to complex and simple reference sentences per system ; number of input sentences for which no simplification occur (No edit).

Human evaluation

- 18 randomly selected sentences
- ratings collected from 18 participants
- judgments on a slider scale from 0 to 5 for *simplicity*, *fluency*, and *adequacy*

Systems	Simplicity	Fluency	Adequacy
GOLD	3.62	4.69	3.80
Zhu	2.62	2.56	2.47
Woodsend	1.69	3.15	3.15
Wubben	1.52	3.05	3.38
Narayan	2.30	3.03	3.35
UNSUP	2.83	3.56	2.83

Human evaluation

- Pairwise comparisons and statistical significance
- One way ANOVA with post-hoc Tukey HSD tests

Systems	GOLD	Zhu	Woodsend	Wubben	Narayan
Zhu	◊□△				
Woodsend	◊□▲	◊■▲			
Wubben	◊□▲	◊■△	♦■▲		
Narayan	◊□▲	♦■▲	♦■▲	♦■▲	
Unsup	♦□△	♦□▲	◊■▲	◊■▲	♦■▲

Table: ◊/♦ is/not significantly different (sig. diff.) wrt simplicity. □/■ is/not sig. diff. wrt fluency. △/▲ is/not sig. diff. wrt adequacy. (significance level: $p < 0.05$)

Qualitative Analysis

Complex. This array distributes data across multiple disks, but the array is seen by the computer user and operating system as one single disk.

Zhu. This array sells data across multiple disks but the array is seen. The computer user and operating as one disk. [0]

Woodsend. This array distributes data across several disks. The array is seen by the computer user and operating system as one single disk. [4.85]

Unsup. This array moves data across disks. The array is seen by the computer user and operating system as one disk. [4.83]

Qualitative Analysis

Complex. Here, simulations of evolution using evolutionary algorithms and artificial life started with the work of Nils Aall Barricelli in the 1960s, and was extended by Alex Fraser, who published a series of papers on simulation of artificial selection.

Zhu. Here, simulations of evolution using evolutionary algorithms and artificial life started with the work of nils aall barricelli in the 1960s, and was extended by alex fraser. Alex fraser published a series of papers on feeling of artificial selection. [3.6]

Woodsend. It has Genetic algorithms. Evolution is using artificial life. This is Nils Aall Barricelli in the 1960s. Here, simulations started with the work. He was extended by Alex Fraser. He published a group of papers on simulation of artificial selection. [0.33]

Unsup. Here, models of evolution using evolutionary algorithms and artificial life started with the work in the 1960s, and was led by Alex Fraser. Alex Fraser wrote a series of papers on model of selection. [5]

Conclusion

We set a benchmark for an **unsupervised approach** to sentence simplification:

- A *deep semantic* based **syntactic simplification**: Better split capabilities
- Competitive with other state-of-the-art supervised methods

[Biran et al., 2011] Biran, O., Brody, S., and Elhadad, N. (2011).
Putting it simply: a context-aware approach to lexical simplification.
In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2*, pages 496–501. Association for Computational Linguistics.

[Bott et al., 2012] Bott, S., Saggion, H., and Mille, S. (2012).
Text simplification tools for spanish.
In *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC)*, pages 1665–1671.

[Canning, 2002] Canning, Y. M. (2002).
Syntactic simplification of Text.
PhD thesis, University of Sunderland.

[Carroll et al., 1999] Carroll, J., Minnen, G., Pearce, D., Canning, Y., Devlin, S., and Tait, J. (1999).
Simplifying text for language-impaired readers.

In *Proceedings of 9th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, volume 99, pages 269–270. Citeseer.

[Chandrasekar et al., 1996] Chandrasekar, R., Doran, C., and Srinivas, B. (1996).
Motivations and methods for text simplification.
In *Proceedings of the 16th International conference on Computational linguistics (COLING)*, pages 1041–1044. Association for Computational Linguistics.

[Chandrasekar and Srinivas, 1997] Chandrasekar, R. and Srinivas, B. (1997).
Automatic induction of rules for text simplification.
Knowledge-Based Systems, 10(3):183–190.

[Coster and Kauchak, 2011] Coster, W. and Kauchak, D. (2011).
Learning to simplify sentences using wikipedia.
In *Proceedings of the Workshop on Monolingual Text-To-Text Generation*, pages 1–9. Association for Computational Linguistics.

[Filippova and Strube, 2008] Filippova, K. and Strube, M. (2008). Dependency tree based sentence compression. In *Proceedings of the Fifth International Natural Language Generation Conference (INLG)*, pages 25–32. Association for Computational Linguistics.

[Knight and Marcu, 2000] Knight, K. and Marcu, D. (2000). Statistics-based summarization-step one: Sentence compression. In *Proceedings of the Seventeenth National Conference on Artificial Intelligence (AAAI) and Twelfth Conference on Innovative Applications of Artificial Intelligence (IAAI)*, pages 703–710. AAAI Press.

[Narayan and Gardent, 2014] Narayan, S. and Gardent, C. (2014). Hybrid simplification using deep semantics and machine translation. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics.

[Siddharthan, 2002] Siddharthan, A. (2002).
An architecture for a text simplification system.
In *Proceedings of the Language Engineering Conference (LEC)*, pages 64–71. IEEE Computer Society.

[Siddharthan, 2010] Siddharthan, A. (2010).
Complex lexico-syntactic reformulation of sentences using typed dependency representations.
In *Proceedings of the 6th International Natural Language Generation Conference (INLG)*, pages 125–133. Association for Computational Linguistics.

[Siddharthan, 2011] Siddharthan, A. (2011).
Text simplification using typed dependencies: a comparison of the robustness of different generation strategies.
In *Proceedings of the 13th European Workshop on Natural Language Generation (ENLG)*, pages 2–11. Association for Computational Linguistics.

[Siddharthan and Mandya, 2014] Siddharthan, A. and Mandya, A. (2014).
Hybrid text simplification using synchronous dependency grammars with hand-written and automatically harvested rules.
In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, pages 722–731, Gothenburg, Sweden. Association for Computational Linguistics.

[Vickrey and Koller, 2008] Vickrey, D. and Koller, D. (2008).
Sentence simplification for semantic role labeling.
In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics (ACL) and the Human Language Technology Conference (HLT)*, pages 344–352.

[Watanabe et al., 2009] Watanabe, W. M., Junior, A. C., Uzêda, V. R., Fortes, R. P. d. M., Pardo, T. A. S., and Aluísio, S. M. (2009).
Facilita: reading assistance for low-literacy readers.
In *Proceedings of the 27th ACM international conference on Design of communication*, pages 29–36. ACM.

[Woodsend and Lapata, 2011] Woodsend, K. and Lapata, M. (2011).
Learning to simplify sentences with quasi-synchronous grammar and integer programming.
In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 409–420. Association for Computational Linguistics.

[Wubben et al., 2012] Wubben, S., van den Bosch, A., and Krahmer, E. (2012).
Sentence simplification by monolingual machine translation.
In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (ACL): Long Papers-Volume 1*, pages 1015–1024. Association for Computational Linguistics.

[Xu et al., 2016] Xu, W., Napoles, C., Pavlick, E., Chen, Q., and Callison-Burch, C. (2016).
Optimizing statistical machine translation for text simplification.
Transactions of the Association for Computational Linguistics.

[Yamada and Knight, 2001] Yamada, K. and Knight, K. (2001).
A syntax-based statistical translation model.
In *Proceedings of the 39th Annual Meeting on Association for Computational Linguistics (ACL)*, pages 523–530. Association for Computational Linguistics.

[Zhu et al., 2010] Zhu, Z., Bernhard, D., and Gurevych, I. (2010).
A monolingual tree-based translation model for sentence simplification.
In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING)*, pages 1353–1361, Stroudsburg, PA, USA. Association for Computational Linguistics.

Modular evaluation

System	Levenshtein Edit distance				BLEU Scores with respect to	
	Complex to System		System to Simple			
	LD	No edit	LD	No edit	complex	simple
complex	0	100	12.24	3	100	49.85
LexSimpl	2.07	22	13.00	1	82.05	44.29
Split	2.27	51	13.62	1	89.70	46.15
Deletion	2.39	4	12.34	0	85.15	47.33
LexSimpl-Split	4.43	11	14.39	0	73.20	41.18
LexSimpl-Deletion	4.29	3	13.09	0	69.84	41.91
Split-Deletion	4.63	4	13.42	0	77.82	43.44
LexSimpl-Split-Deletion	6.75	3	14.29	0	63.41	38.47
GOLD (simple)	12.24	3	0	100	49.85	100

Pairwise Split evaluation

System pairs		Average Score (number of split sentences)					
A	B	ALL-A	ALL-B	ONLY-A	BOTH-AB		ONLY-B
					A	B	
	GOLD		3.85(28)	2.15(32)	2.80(17)	3.70(17)	4.05(11)
	Zhu		2.25(80)	1.53(4)	2.45(45)	2.42(45)	2.02(35)
UNSUP	Woodsend	2.37(49)	2.08(63)	2.42(11)	2.36(38)	2.29(38)	1.78(25)
	Wubben		2.73(1)	2.32(48)	4.75(1)	2.73(1)	0(0)
	Narayan		2.09(10)	2.29(41)	2.78(8)	1.79(8)	3.81(2)