

Question Generation with Minimal Recursion Semantics

Xuchen Yao¹ and Yi Zhang²

¹European Masters in Language and Communication Technologies
University of Groningen & Saarland University

²Saarland University

German Research Center for Artificial Intelligence



rijksuniversiteit
 groningen



UNIVERSITÄT
DES
SAARLANDES



German
Research Center
for Artificial
Intelligence

18 June, QG/QGSTEC/2010



Outline

Introduction

Template/Syntax/Semantics-based Approaches

Why Semantics-based?

Background

MRS/ERG/PET/LKB

System Architecture

Overview

MRS Transformation for Simple Sentences

MRS Decomposition for Complex Sentences

Language Independence and Domain Adaptability

Evaluation



Approaches

- Template-based (Mostow and Chen (2009))
 - *What did <character> <verb>?*
- Syntax-based (Wyse and Piwek (2009), Heilman and Smith (2009))
 - John plays football. (S NP (VP (V NP)))
 - John plays what? (S NP (VP (V WHNP)))
 - John does play what? (S NP (VP (Aux-V V WHNP)))
 - Does John play what? (S Aux-V NP (VP (V WHNP)))
 - What does John play? (S WHNP Aux-V NP (VP (V)))
- Semantics-based
 - play(John, football)
 - play(John, *what*)
 - play(*who*, football)

Approaches

- Template-based (Mostow and Chen (2009))
 - *What did <character> <verb>?*
- Syntax-based (Wyse and Piwek (2009), Heilman and Smith (2009))
 - John plays football. (S NP (VP (V NP)))
 - John plays what? (S NP (VP (V WHNP)))
 - John does play what? (S NP (VP (Aux-V V WHNP)))
 - Does John play what? (S Aux-V NP (VP (V WHNP)))
 - What does John play? (S WHNP Aux-V NP (VP (V)))
- Semantics-based
 - play(John, football)
 - play(John, **what**)
 - play(**who**, football)

Outline

Introduction

Template/Syntax/Semantics-based Approaches

Why Semantics-based?

Background

MRS/ERG/PET/LKB

System Architecture

Overview

MRS Transformation for Simple Sentences

MRS Decomposition for Complex Sentences

Language Independence and Domain Adaptability

Evaluation

Why Semantics-based?

- Something different than template/syntax-based.
- More intuitive?
- More language independent (universal)?
- Make use of the generation function of the English Resource Grammar
 - Sag, I. A. & Flickinger, D. Generating Questions with Deep Reversible Grammars. In *Proceedings of the First Workshop on the Question Generation Shared Task and Evaluation Challenge*. 2008.
 - Deeper is better?

Why Semantics-based?

- Something different than template/syntax-based.
- More intuitive?
- More language independent (universal)?
- Make use of the generation function of the English Resource Grammar
 - Sag, I. A. & Flickinger, D. Generating Questions with Deep Reversible Grammars. In *Proceedings of the First Workshop on the Question Generation Shared Task and Evaluation Challenge*. 2008.
 - Deeper is better?

Outline

Introduction

Template/Syntax/Semantics-based Approaches
Why Semantics-based?

Background

MRS/ERG/PET/LKB

System Architecture

Overview

MRS Transformation for Simple Sentences

MRS Decomposition for Complex Sentences

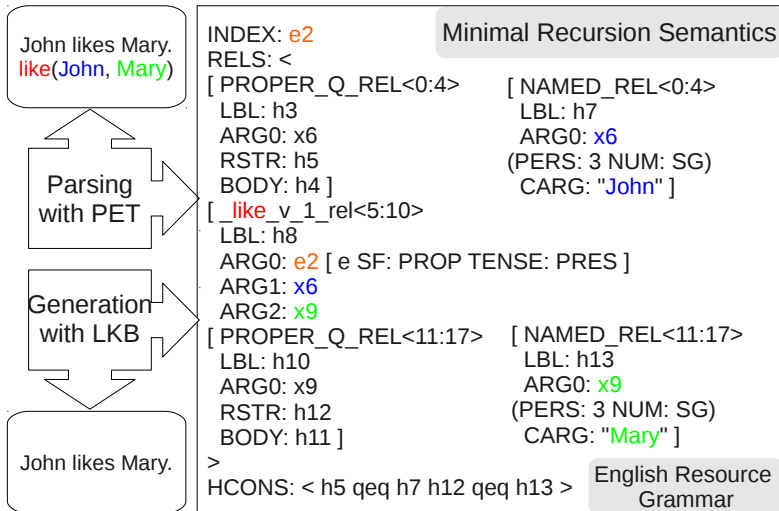
Language Independence and Domain Adaptability

Evaluation



DELPH-IN (MRS/ERG/PET/LKB)

Deep Linguistic Processing with HPSG: <http://www.delph-in.net/>



Details

(THEORY)MRS: Minimal Recursion Semantics

a meta-level language for describing semantic structures in some underlying object language.

(GRAMMAR)ERG: English Resource Grammar

a general-purpose broad-coverage grammar implementation under the HPSG framework.

(TOOL)LKB: Linguistic Knowledge Builder

a grammar development environment for grammars in typed feature structures and unification-based formalisms.

(TOOL)PET: a platform for experimentation with efficient HPSG processing techniques

a two-stage parsing model with HPSG rules and PCFG models, balancing between precise linguistic interpretation and robust probabilistic coverage.

Details

(THEORY)MRS: Minimal Recursion Semantics

a meta-level language for describing semantic structures in some underlying object language.

(GRAMMAR)ERG: English Resource Grammar

a general-purpose broad-coverage grammar implementation under the HPSG framework.

(TOOL)LKB: Linguistic Knowledge Builder

a grammar development environment for grammars in typed feature structures and unification-based formalisms.

(TOOL)PET: a platform for experimentation with efficient HPSG processing techniques

a two-stage parsing model with HPSG rules and PCFG models, balancing between precise linguistic interpretation and robust probabilistic coverage.

Outline

Introduction

Template/Syntax/Semantics-based Approaches
Why Semantics-based?

Background

MRS/ERG/PET/LKB

System Architecture

Overview

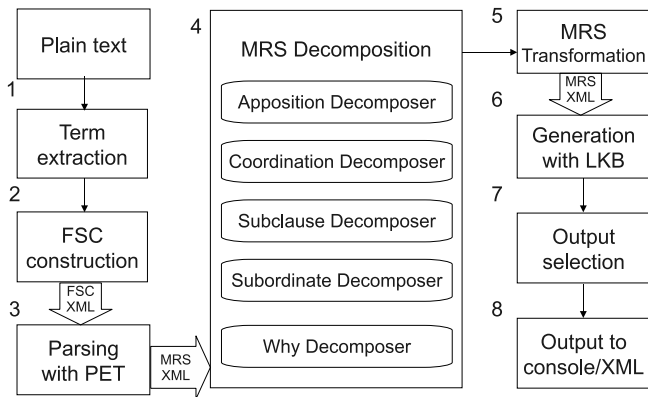
MRS Transformation for Simple Sentences
MRS Decomposition for Complex Sentences
Language Independence and Domain Adaptability

Evaluation

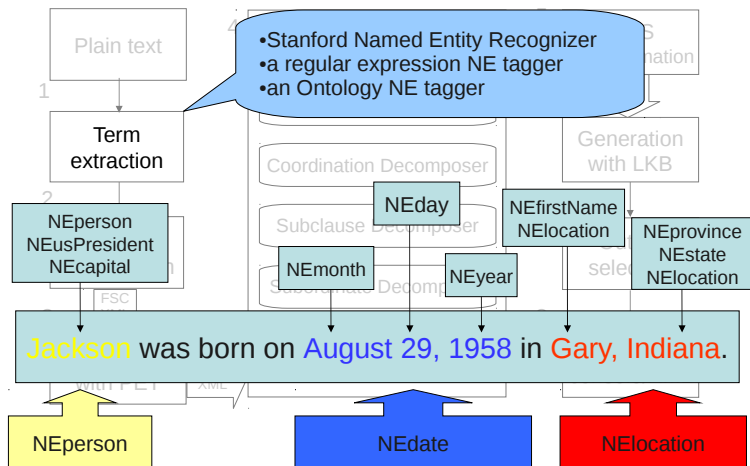


MrsQG (Task B)

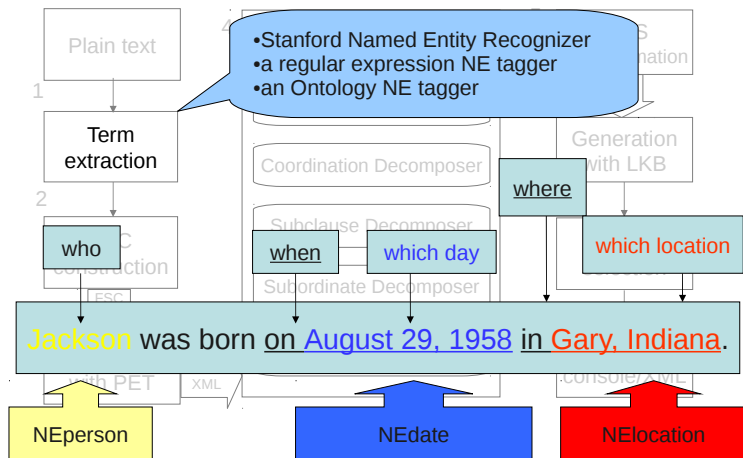
<http://code.google.com/p/mrsqg/>



Term Extraction



Term Extraction



Outline

Introduction

Template/Syntax/Semantics-based Approaches
Why Semantics-based?

Background

MRS/ERG/PET/LKB

System Architecture

Overview

MRS Transformation for Simple Sentences

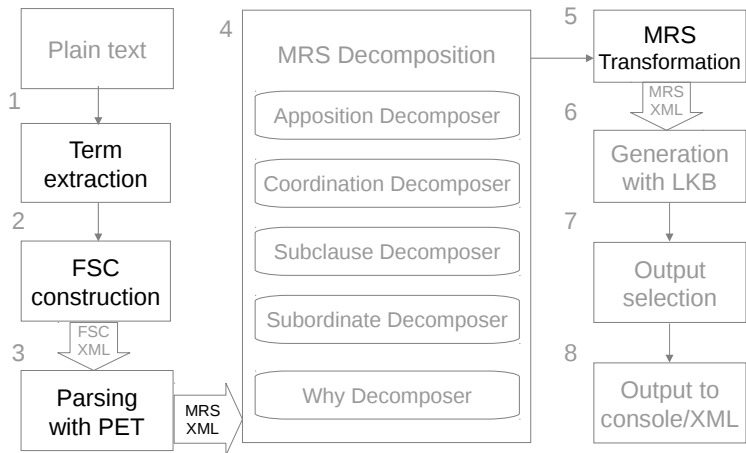
MRS Decomposition for Complex Sentences

Language Independence and Domain Adaptability

Evaluation



MRS Transformation





WHO

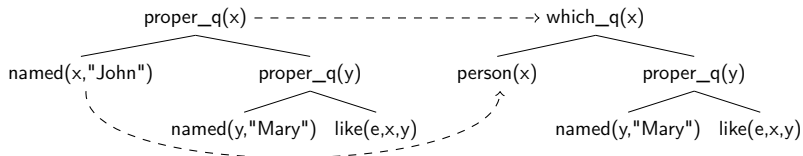


Figure: "John likes Mary" \rightarrow "Who likes Mary?"

WHERE

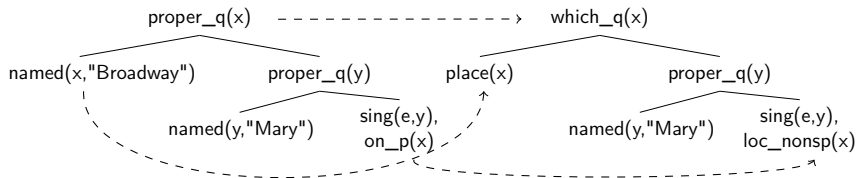


Figure: “Mary sings on Broadway.” → “Where does Mary sing?”

WHEN

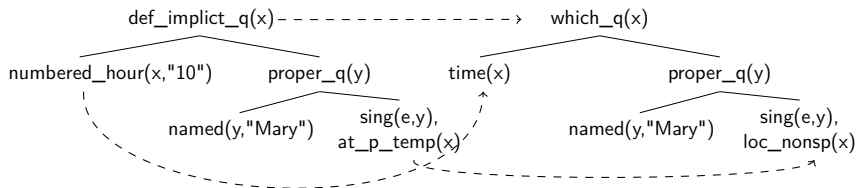


Figure: “Mary sings at 10.” → “When does Mary sing?”

WHY

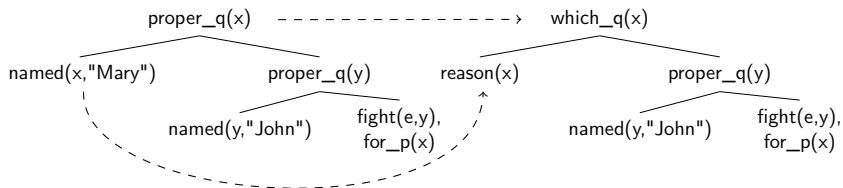


Figure: "John fights for Mary." → "Why does John fight?"

Outline

Introduction

Template/Syntax/Semantics-based Approaches
Why Semantics-based?

Background

MRS/ERG/PET/LKB

System Architecture

Overview

MRS Transformation for Simple Sentences

MRS Decomposition for Complex Sentences

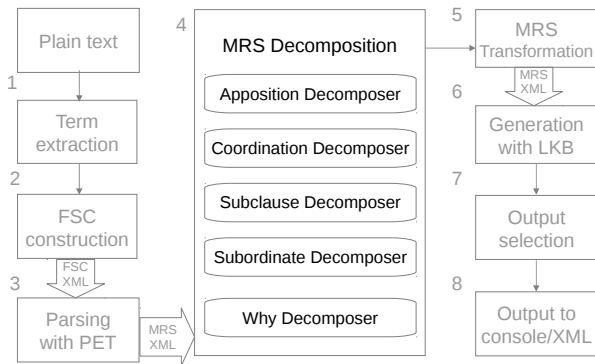
Language Independence and Domain Adaptability

Evaluation



MRS Decomposition

Complex Sentences -> Simple Sentences



Subclause Decomposer

identifies the verb, extracts its arguments and reconstructs MRS

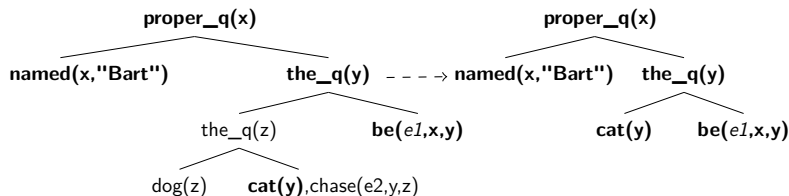


Figure: "Bart is the cat that chases the dog." → "Bart is the cat."

Subclause Decomposer

identifies the verb, extracts its arguments and reconstructs MRS

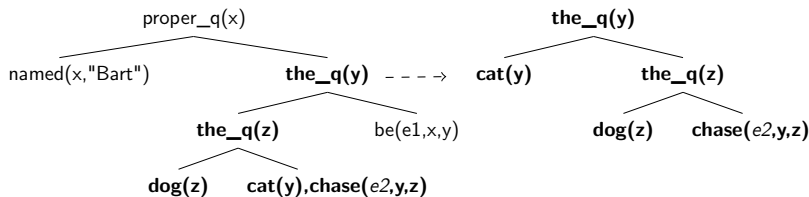
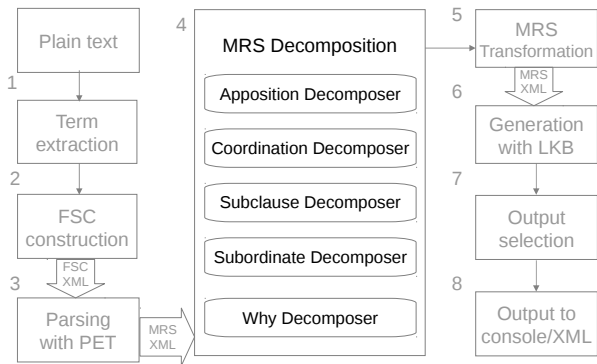


Figure: "Bart is **the cat that chases the dog.**" → "The cat chases the dog."

MRS Decomposition

Complex Sentences -> Simple Sentences



Outline

Introduction

Template/Syntax/Semantics-based Approaches
Why Semantics-based?

Background

MRS/ERG/PET/LKB

System Architecture

Overview

MRS Transformation for Simple Sentences

MRS Decomposition for Complex Sentences

Language Independence and Domain Adaptability

Evaluation

Language Independence

MrsQG aims to stay language-neutral based on a semantics transformation of sentences.

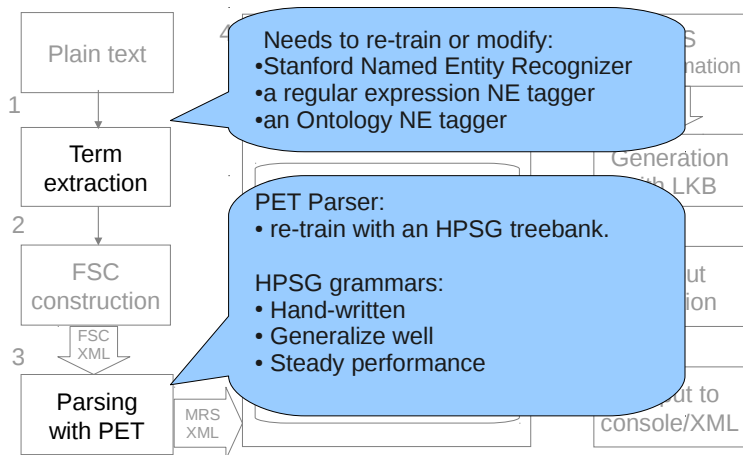
In Principle

It needs little modification to adapt to other languages.

In Practice

It is difficult to guarantee absolute language independence.

Domain Adaptability



QGSTEC2010

The Question Generation Shared Task and Evaluation Challenge (QGSTEC) 2010

Task B: QG from Sentences.

Participants are given one complete sentence from which their system must generate questions.

1. **RELEVANCE.** Questions should be relevant to the input sentence.
2. **QUESTION TYPE.** Questions should be of the specified target question type.
3. **SYNTACTIC CORRECTNESS AND FLUENCY.** The syntactic correctness is rated to ensure systems can generate sensible output.
4. **AMBIGUITY.** The question should make sense when asked more or less out of the blue.
5. **VARIETY.** Pairs of questions in answer to a single input are evaluated on how different they are from each other.

Examples

TEXT: Alexander Graham Bell, who had risen to prominence through his invention of the telephone, took a great interest in recording sounds, even suggesting to Edison that they might collaborate.

WHO: Who took a great interest in recording sounds?

WHO: Who is Alexander Graham Bell?

WHAT: A great interest in what did Alexander Graham Bell take?

WHAT: What did Alexander Graham Bell take a great interest in?

WHY: Why Alexander Graham Bell took a great interest in recording sounds?

WHY: Why do they collaborate?

Outline

Introduction

Template/Syntax/Semantics-based Approaches
Why Semantics-based?

Background

MRS/ERG/PET/LKB

System Architecture

Overview

MRS Transformation for Simple Sentences

MRS Decomposition for Complex Sentences

Language Independence and Domain Adaptability

Evaluation

Conclusion

- semantics-based (easy in theory, difficult in practice)
 - multi-linguality
 - cross-domain
- deep grammar (worry less, wait more)
 - generation \leftrightarrow grammaticality
 - heavy machinery

Conclusion

- semantics-based (easy in theory, difficult in practice)
 - multi-linguality
 - cross-domain
- deep grammar (worry less, wait more)
 - generation \leftrightarrow grammaticality
 - heavy machinery

References

- M. Heilman and N. A. Smith. Question Generation via Overgenerating Transformations and Ranking. Technical report, Language Technologies Institute, Carnegie Mellon University Technical Report CMU-LTI-09-013, 2009.
- Jack Mostow and Wei Chen. Generating instruction automatically for the reading strategy of self-questioning. In *Proceeding of the 2009 conference on Artificial Intelligence in Education*, Amsterdam, The Netherlands, 2009. ISBN 978-1-60750-028-5.
- Brendan Wyse and Paul Piwek. Generating Questions from OpenLearn study units. In *Proceedings of the 2nd Workshop on Question Generation In Craig, S.D. & Dicheva, S. (Eds.) (2009) AIED 2009: 14th International Conference on Artificial Intelligence in Education: Workshops Proceedings*, 2009.