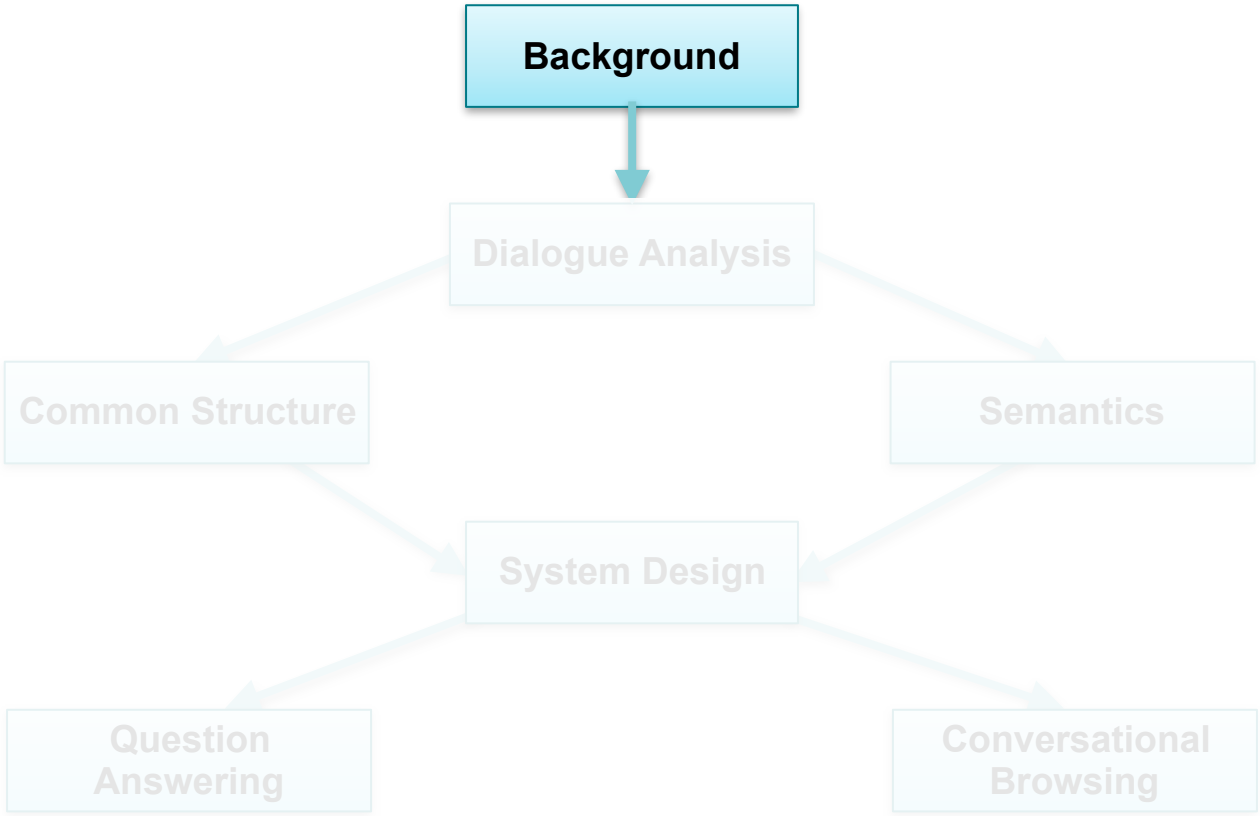
Knowledge-based Conversational Search

Svitlana Vakulenko

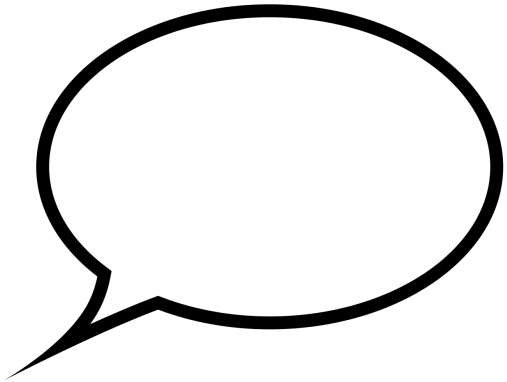
Outline


Background
Dialogue Analysis
Common Structure
Semantics
System Design
Question Answering
Conversational Browsing

Conversational Search

Conversational Search?

- Information Seeking **Task**
- Dialogue **Interface**

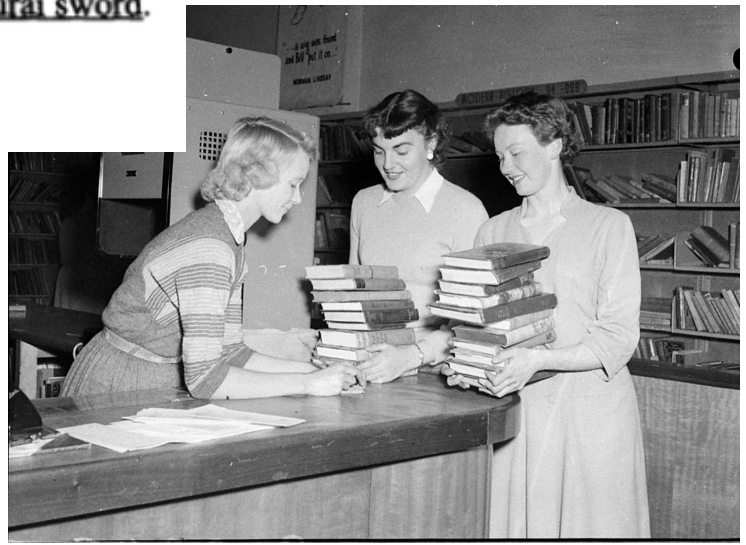


Conversational Search?

- **Information Seeking** Task
- **Dialogue** Interface

Information-seeking Dialogue

IS: Do you have the book that *they showed* in the filmstrip?
MS: Yes. Uhm. (SILENCE) You're thinking of the one that was on . . . ?
IS: Samurai.
MS: Hhh. The one. It was the picture of the guy swinging the samurai sword.
[
IS: Yeah, samurai sword.
MS: Was that Sign of the Chrysanthemum?
[
I: Yaa.

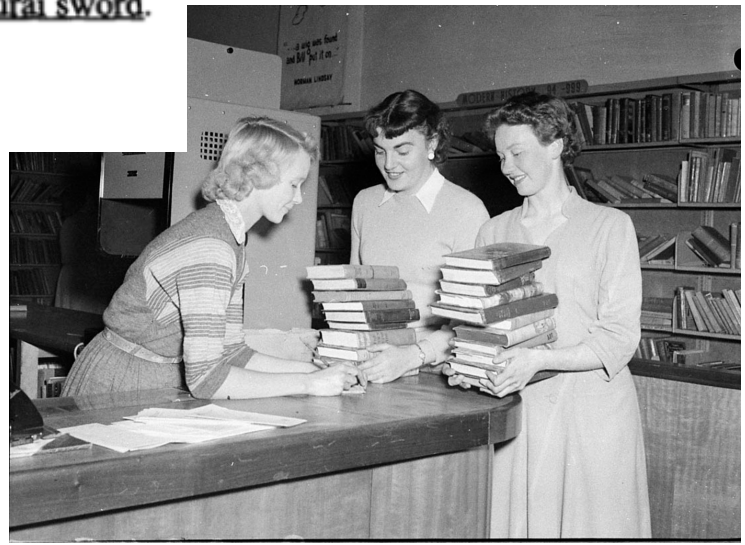


Solomon, Paul. "Conversation in information-seeking contexts: A test of an analytical framework." *Library & Information Science Research.* 1997

Information-seeking Dialogue

IS: Do you have the book that *they showed* in the filmstrip?
MS: Yes. Uhm. (SILENCE) You're thinking of the one that was on . . . ?
IS: Samurai.
MS: Hhh. The one. It was the picture of the guy swinging the samurai sword.
[
IS: Yeah, samurai sword.
MS: Was that Sign of the Chrysanthemum?
[
I: Yaa.

1. **Ambiguous (underspecified) questions**

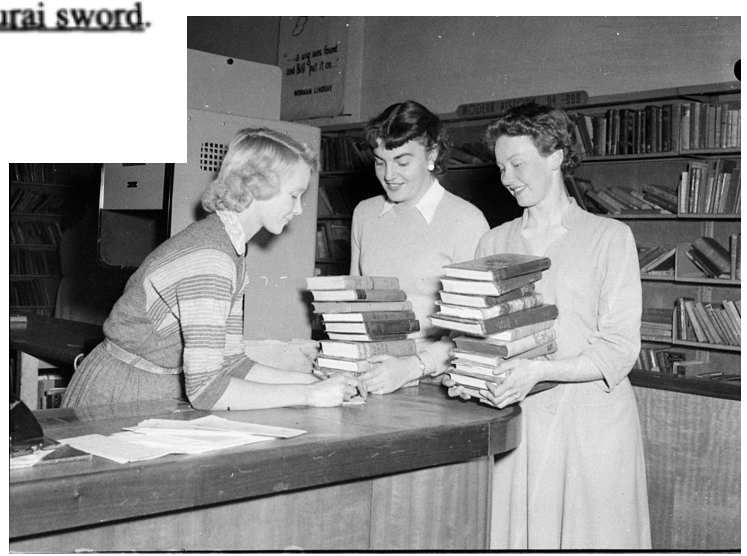


Solomon, Paul. "Conversation in information-seeking contexts: A test of an analytical framework." *Library & Information Science Research.* 1997

Information-seeking Dialogue

IS: Do you have the book that *they showed* in the filmstrip?
MS: Yes. Uhm. (SILENCE) You're thinking of the one that was on . . . ?
IS: Samurai.
MS: Hhh. The one. It was the picture of the guy swinging the samurai sword.
[
IS: Yeah, samurai sword.
MS: Was that Sign of the Chrysanthemum?
[
I: Yaa.

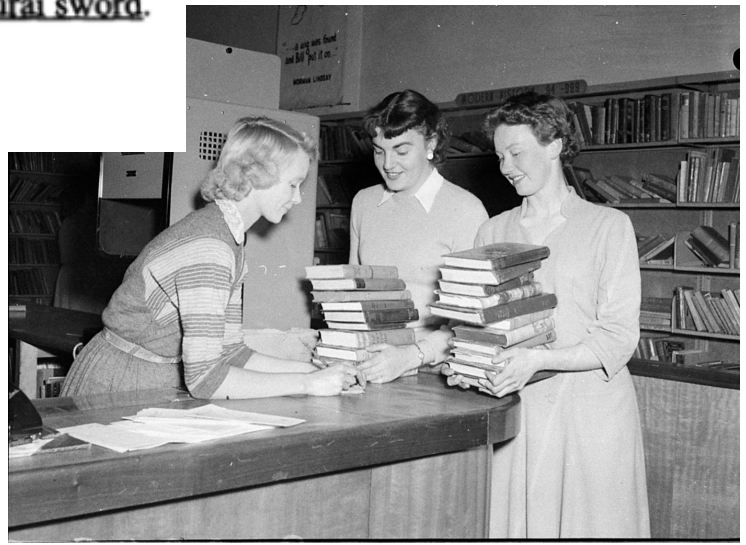
1. Ambiguous (underspecified) questions
2. **Guessing**



Solomon, Paul. "Conversation in information-seeking contexts: A test of an analytical framework." *Library & Information Science Research.* 1997

Information-seeking Dialogue

IS: Do you have the book that *they showed* in the filmstrip?
MS: Yes. Uhm. (SILENCE) You're thinking of the one that was on . . . ?
IS: Samurai.
MS: Hhh. The one. It was the picture of the guy swinging the samurai sword.
[
IS: Yeah, samurai sword.
MS: Was that Sign of the Chrysanthemum?
[
I: Yaa.



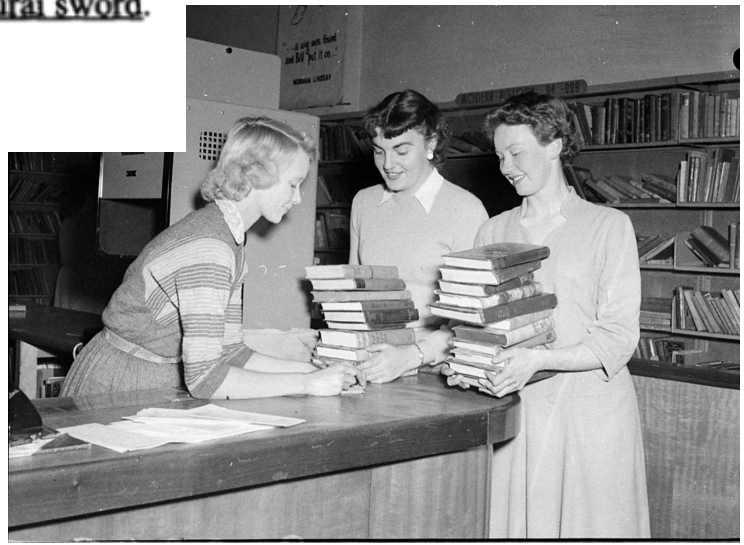
1. Ambiguous (underspecified) questions
2. Guessing
3. **Additional evidence**

Solomon, Paul. "Conversation in information-seeking contexts: A test of an analytical framework." *Library & Information Science Research.* 1997

Information-seeking Dialogue

IS: Do you have the book that *they showed* in the filmstrip?
MS: Yes. Uhm. (SILENCE) You're thinking of the one that was on . . . ?
IS: Samurai.
MS: Hhh. The one. It was the picture of the guy swinging the samurai sword.
IS: [Yeah, samurai sword.
MS: Was that Sign of the Chrysanthemum?
I: [Yaa.

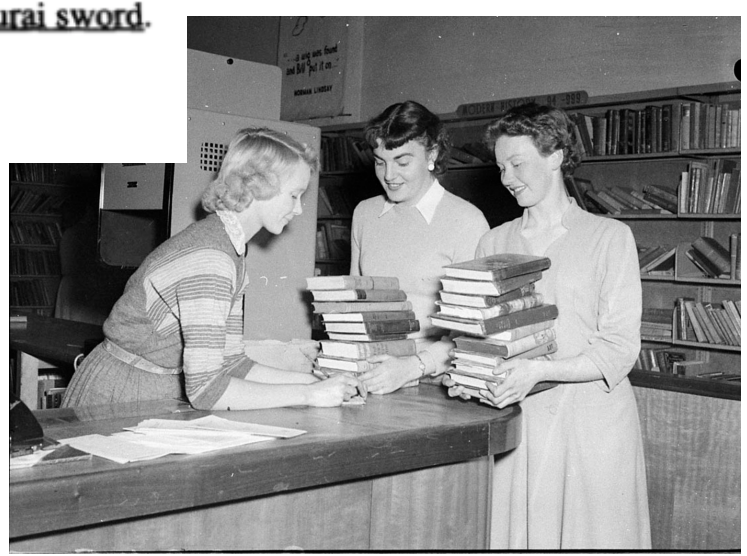
1. Ambiguous (underspecified) questions
2. Guessing
3. Additional evidence
4. **Common knowledge**



Solomon, Paul. "Conversation in information-seeking contexts: A test of an analytical framework." *Library & Information Science Research.* 1997

Information-seeking Dialogue

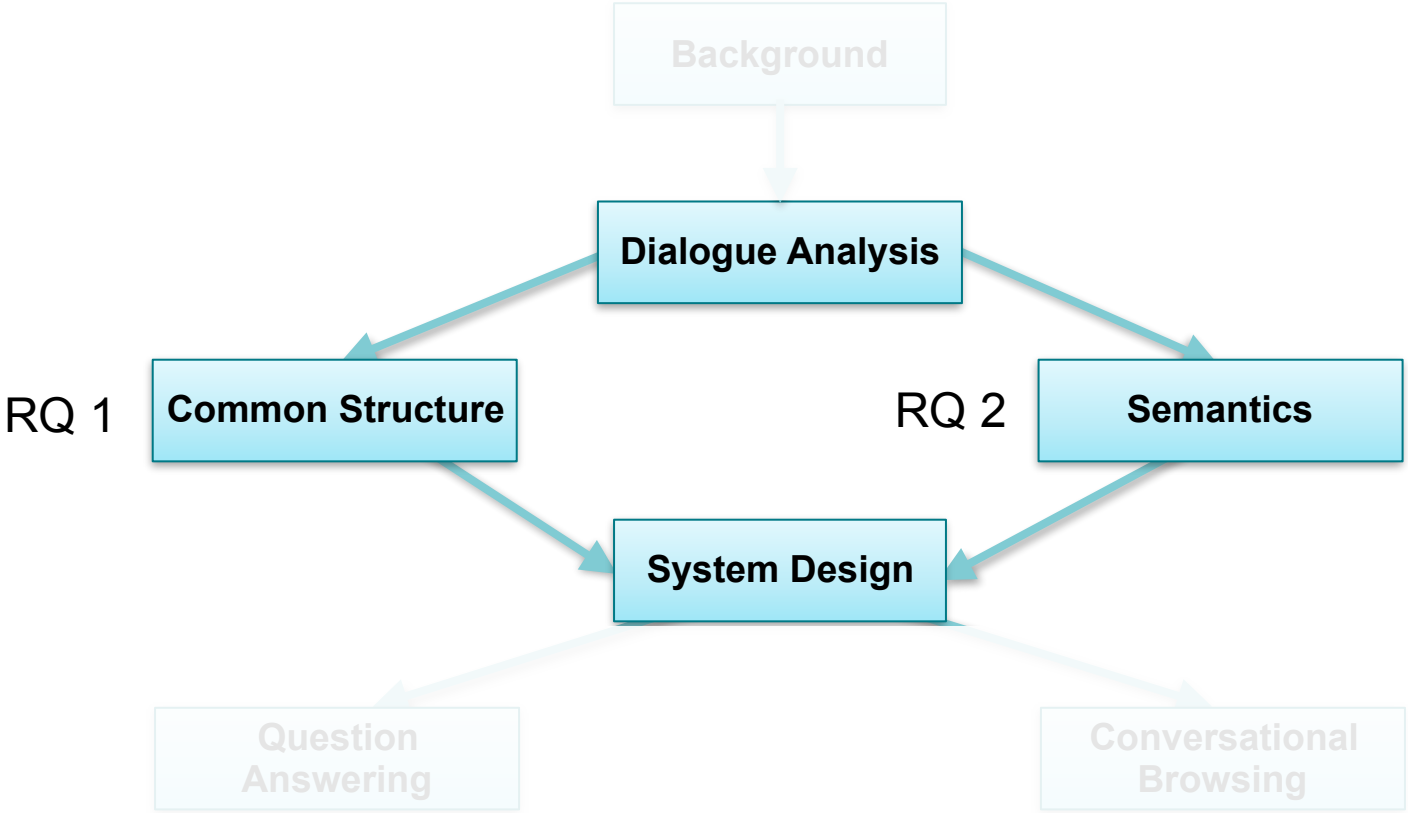
IS: Do you have the book that *they showed* in the filmstrip?
MS: Yes. Uhm. (SILENCE) You're thinking of the one that was on . . . ?
IS: Samurai.
MS: Hhh. The one. It was the picture of the guy swinging the samurai sword.
IS: [Yeah, samurai sword.
MS: Was that Sign of the Chrysanthemum?
I: [Yaa.



1. Ambiguous (underspecified) questions
2. Guessing
3. Additional evidence
4. Common knowledge
5. **Feedback**

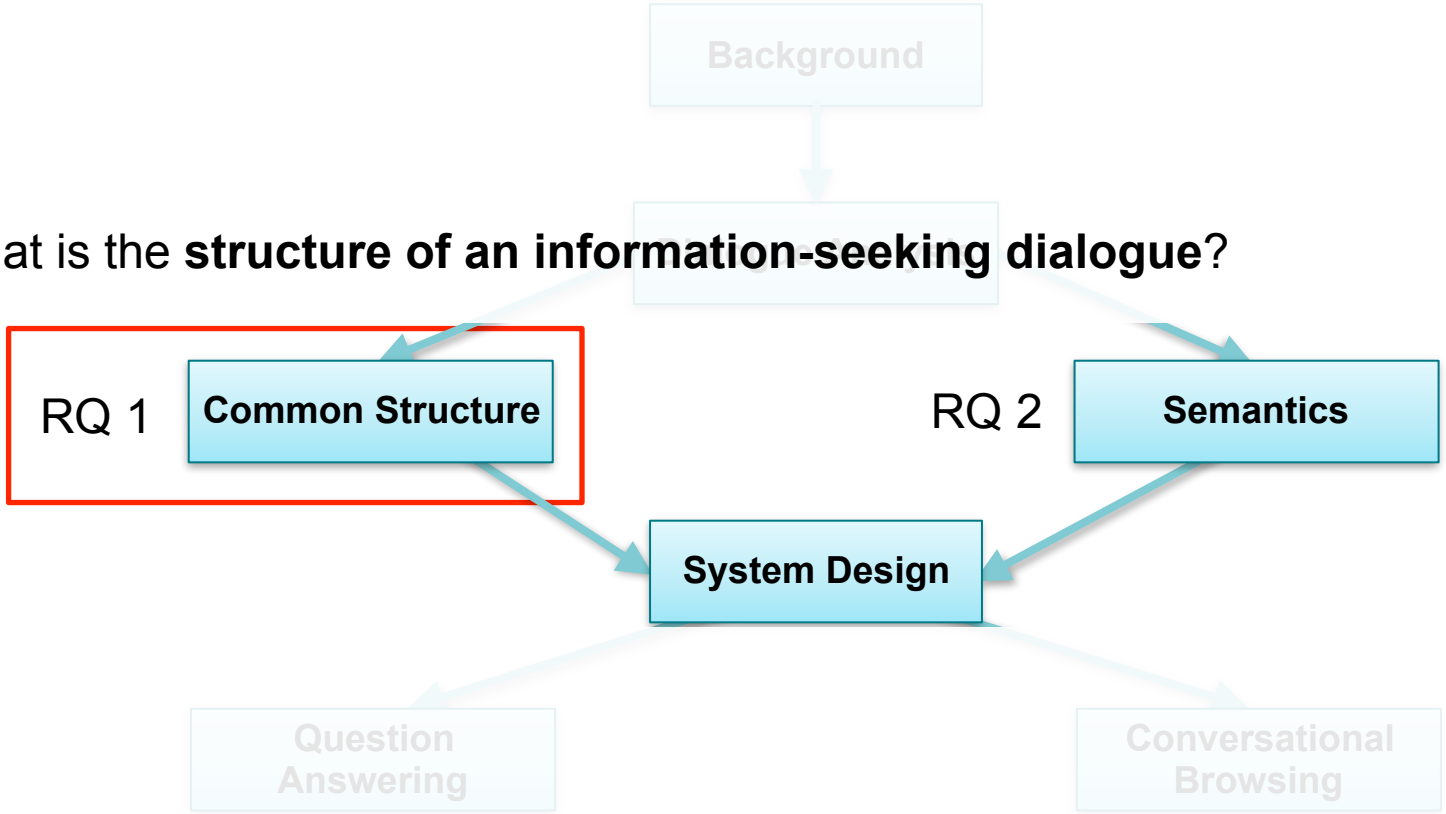
Solomon, Paul. "Conversation in information-seeking contexts: A test of an analytical framework." *Library & Information Science Research.* 1997

Outline

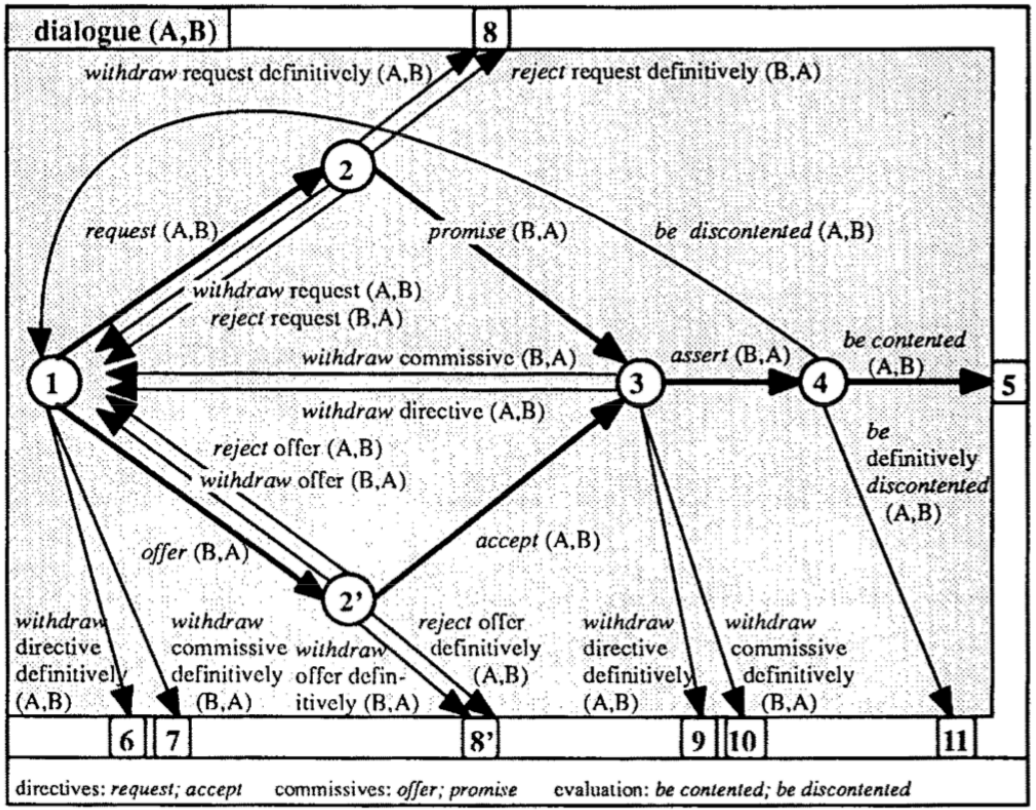

Background
Dialogue Analysis
RQ 1
Common Structure
RQ 2
Semantics
System Design
Question Answering
Conversational Browsing

Outline

What is the **structure of an information-seeking dialogue**?

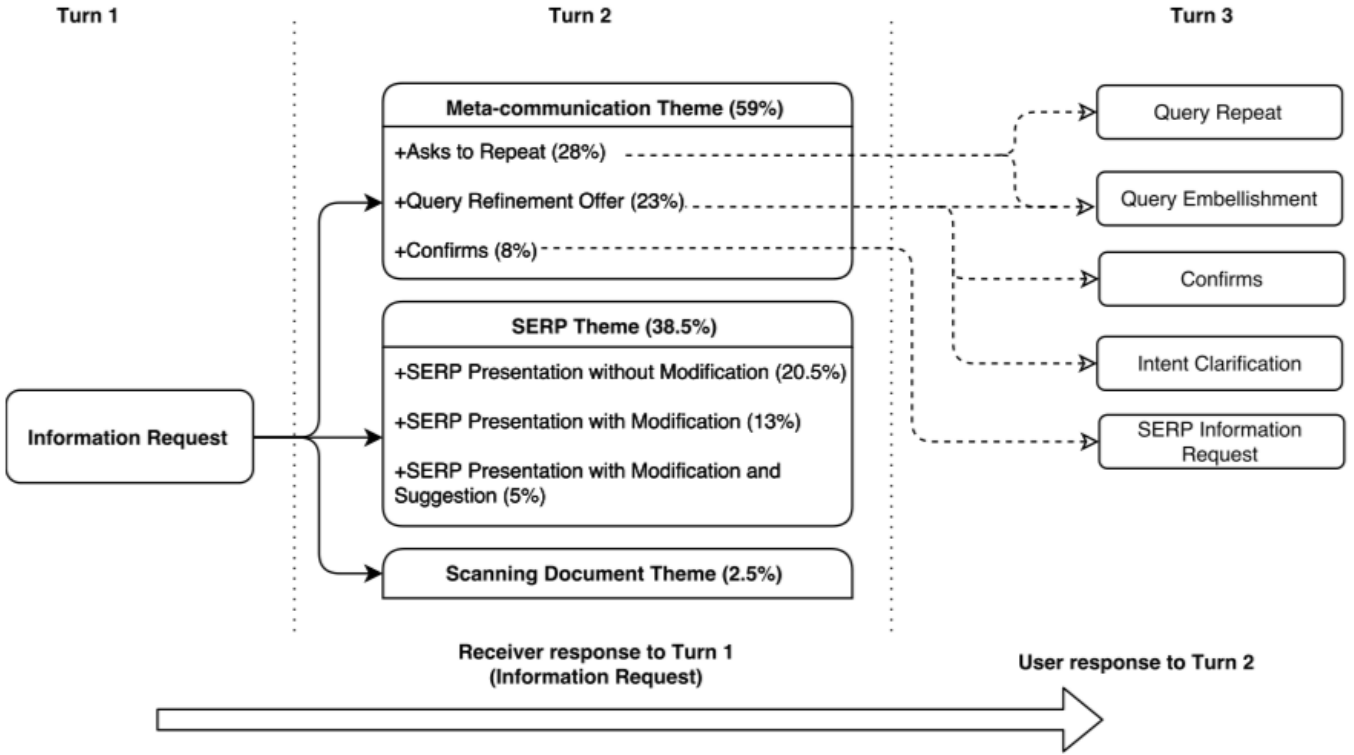


COnversational Roles model



Sitter, Stefan and Adelheit Stein. “Modeling the illocutionary aspects of information-seeking dialogues.” *Information Processing & Management.* 1992

Spoken Conversational Search model



Trippas, Johanne et al. “How do people interact in conversational speech-only search tasks.” CHIIR. 2017

Conversation Mining

Process Mining

- system log (workflow) analysis
  - P - process
  - T - trace, where $T \in P$
  - $T = < e_1, e_2, e_3 \ldots >$, where e - event and $e \in E$

Wil M. P. van der Aalst: Process Mining - Data Science in Action, Second Edition. Springer 2016

Process Mining

- system log (workflow) analysis
  - P - process
  - T - trace, where $T \in P$
  - $T = < e_1, e_2, e_3 \ldots >$, where e - event and $e \in E$

P_1

$T_1 : e_1\ e_2\ e_3$

$T_2 : e_2\ e_1\ e_3$

$T_3 : e_1\ e_3$

Wil M. P. van der Aalst: Process Mining - Data Science in Action, Second Edition. Springer 2016

Process Mining

- system log (workflow) analysis
  - P - process
  - T - trace, where $T \in P$
  - $T = < e_1, e_2, e_3 \ldots >$, where e - event and $e \in E$

P_1

$T_1 : e_1\ e_2\ e_3$ e_1 - place order

$T_2 : e_2\ e_1\ e_3$ e_2 - send request

$T_3 : e_1\ e_3$ e_3 - deliver order

Wil M. P. van der Aalst: Process Mining - Data Science in Action, Second Edition. Springer 2016

Process Mining

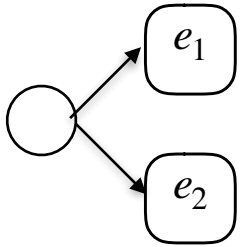
- system log (workflow) analysis
  - P - process
  - T - trace, where $T \in P$
  - $T = < e_1, e_2, e_3 \ldots >$, where e - event and $e \in E$

P_1

$T_1 : e_1\ e_2\ e_3$ e_1 - place order

$T_2 : e_2\ e_1\ e_3$ e_2 - send request

$T_3 : e_1\ e_3$ e_3 - deliver order



Process Mining

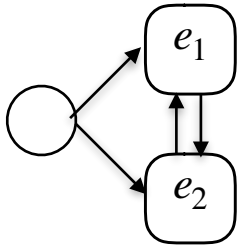
- system log (workflow) analysis
  - P - process
  - T - trace, where $T \in P$
  - $T = < e_1, e_2, e_3 \ldots >$, where e - event and $e \in E$

P_1

$T_1 : e_1\ e_2\ e_3$ e_1 - place order

$T_2 : e_2\ e_1\ e_3$ e_2 - send request

$T_3 : e_1\ e_3$ e_3 - deliver order



Wil M. P. van der Aalst: Process Mining - Data Science in Action, Second Edition. Springer 2016

Process Mining

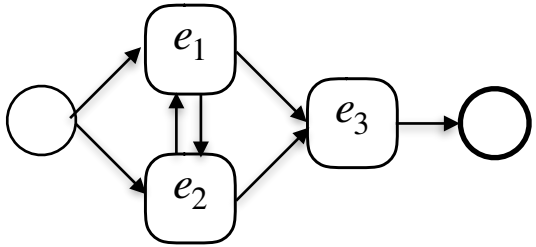
- system log (workflow) analysis
  - P - process
  - T - trace, where $T \in P$
  - $T = < e_1, e_2, e_3 \ldots >$, where e - event and $e \in E$

P_1

$T_1 : e_1\ e_2\ e_3$ e_1 - place order

$T_2 : e_2\ e_1\ e_3$ e_2 - send request

$T_3 : e_1\ e_3$ e_3 - deliver order



Wil M. P. van der Aalst: Process Mining - Data Science in Action, Second Edition. Springer 2016

Conversation Mining

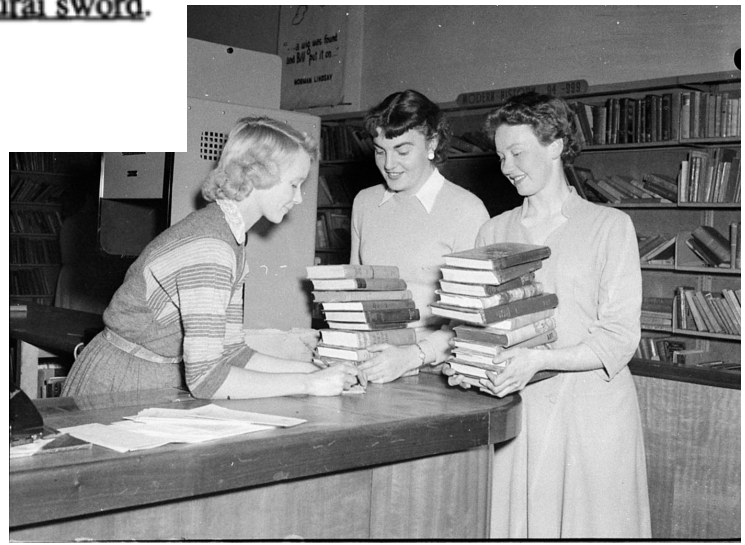
- conversation analysis
  - P - communication process
  - T - dialogue transcript, where $T \in P$
  - $T = < u_1, u_2, u_3 \ldots >$, where u - utterance and $u \in U$

Conversation Mining

- conversation analysis
  - P - communication process
  - T - dialogue transcript, where $T \in P$
  - $T = < u_1, u_2, u_3 \ldots >$, where u - utterance and $u \in U$
  - utterance labeling: $U \rightarrow L$, where $|U| > |L|$
  - $T = < l_1, l_2, l_3 \ldots >$, where l - label and $l \in L$

Example

IS: Do you have the book that *they showed* in the filmstrip?
MS: Yes. Uhm. (SILENCE) You're thinking of the one that was on . . . ?
IS: Samurai.
MS: Hhh. The one. It was the picture of the guy swinging the samurai sword.
[
IS: Yeah, samurai sword.
MS: Was that Sign of the Chrysanthemum?
[
I: Yaa.

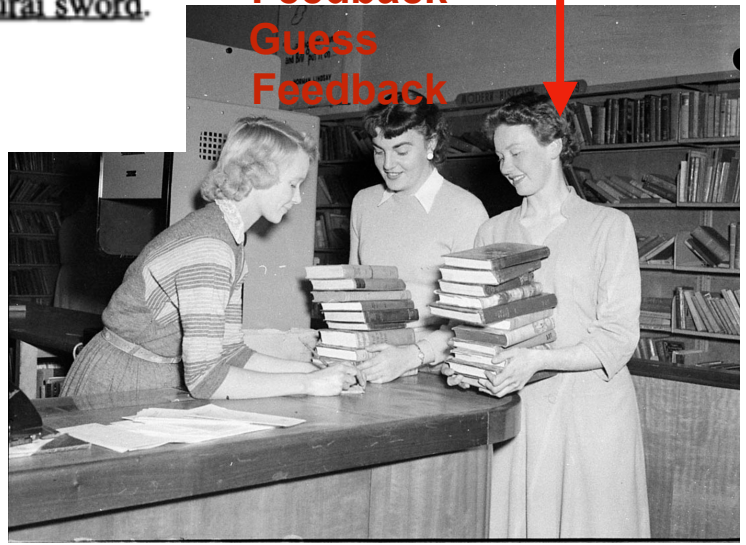


Solomon, Paul. "Conversation in information-seeking contexts: A test of an analytical framework." *Library & Information Science Research.* 1997

Example

IS: Do you have the book that *they showed* in the filmstrip?
MS: Yes. Uhm. (SILENCE) You're thinking of the one that was on . . . ?
IS: Samurai.
MS: Hhh. The one. It was the picture of the guy swinging the samurai sword.
[
IS: Yeah, samurai sword.
MS: Was that Sign of the Chrysanthemum?
[
I: Yaa.

Question
Guess
Hint
Guess
Feedback
Guess
Feedback



Solomon, Paul. "Conversation in information-seeking contexts: A test of an analytical framework." *Library & Information Science Research.* 1997

Datasets

Seeker (User) - Intermediary (Agent) - [Information Source]

- **Spoken Conversation Search** 39 dialogues 13 labels
- **Open Data Exploration** 26 dialogues 20 labels
- **DSTC1** bus schedules 15,866 dialogues 37 labels
- **DSTC2** restaurant reservation 2,118 dialogues 21 labels

Datasets

Seeker (User) - Intermediary (Agent) - [Information Source]

- **Spoken Conversation Search** 39 dialogues 13 labels
- **Open Data Exploration** 26 dialogues 20 labels
- **DSTC1** bus schedules 15,866 dialogues 37 labels
- **DSTC2** restaurant reservation 2,118 dialogues 21 labels

human-human

human-machine

QRFA Annotation Schema

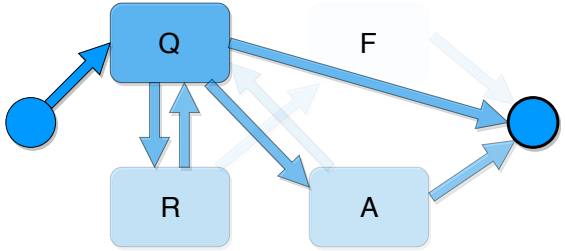
		Proactive
User	**Q**uery	Information Prompt
Agent	**R**equest	Offer Understand

QRFA Annotation Schema

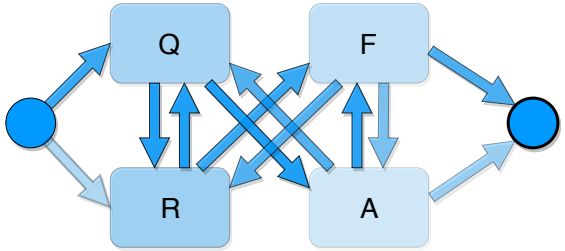
	Proactive		Reactive	
User	**Q**uery	Information	**F**eedback	Positive
		Prompt		Negative
Agent	**R**equest	Offer	**A**nswer	Results
		Understand		Backchannel
				Empty

QRFA Model

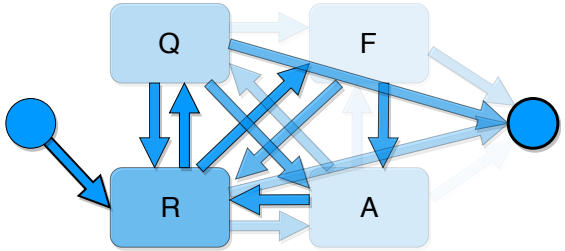
SCS



ODE

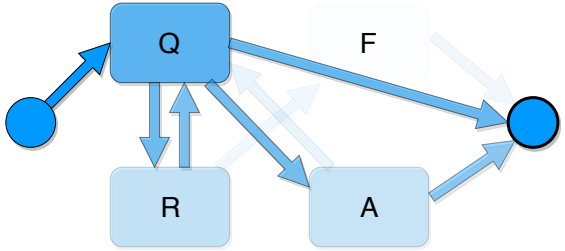


DSTC1

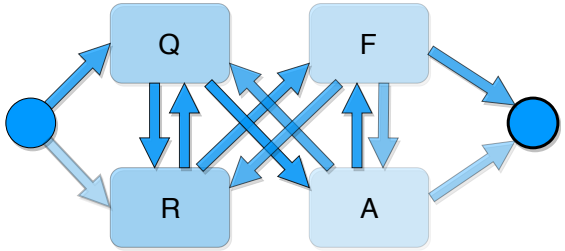


QRFA Model

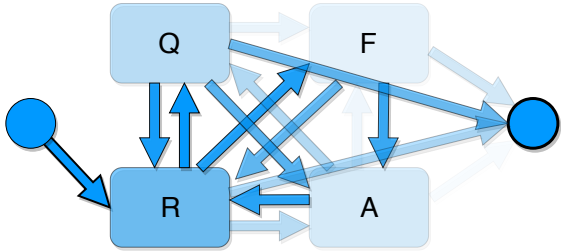
SCS

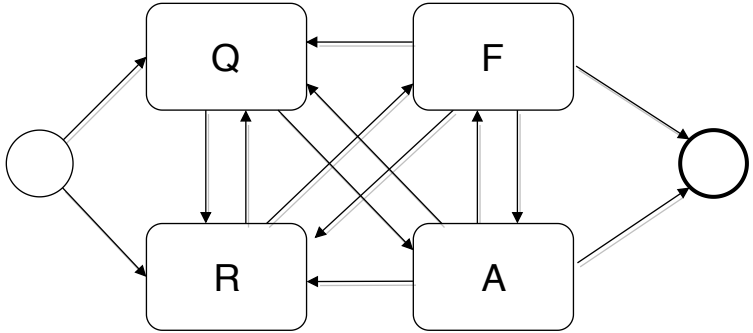


ODE

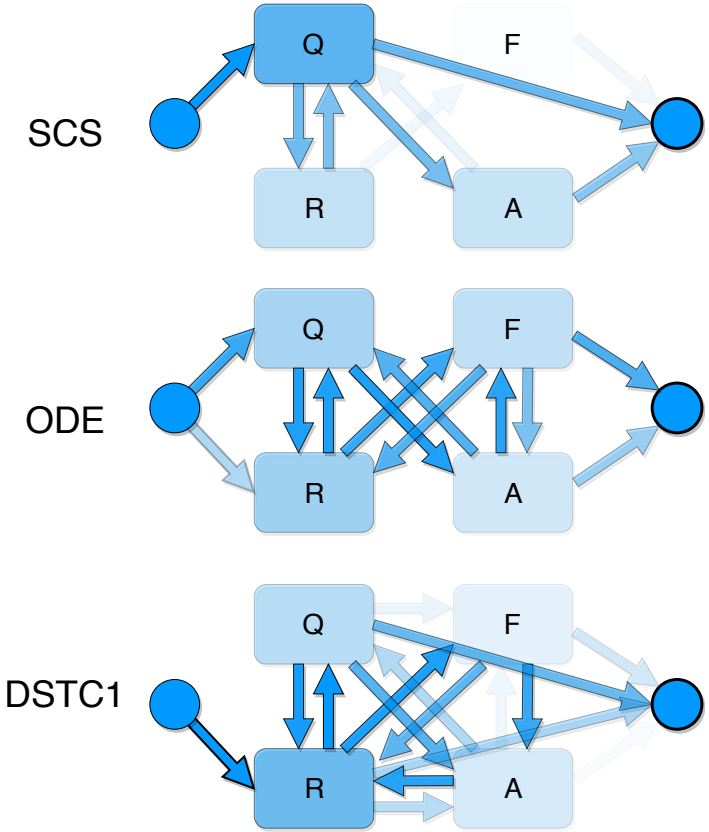


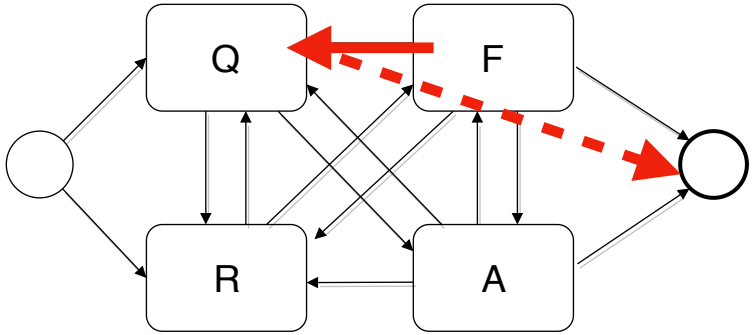
DSTC1





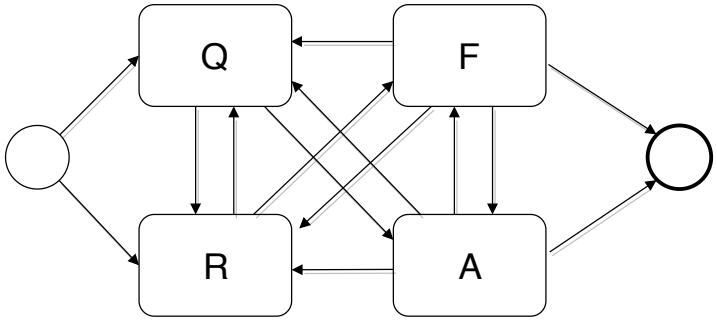
QRFA Model





Model Evaluation

- **Simplicity**, fitness, generalisation

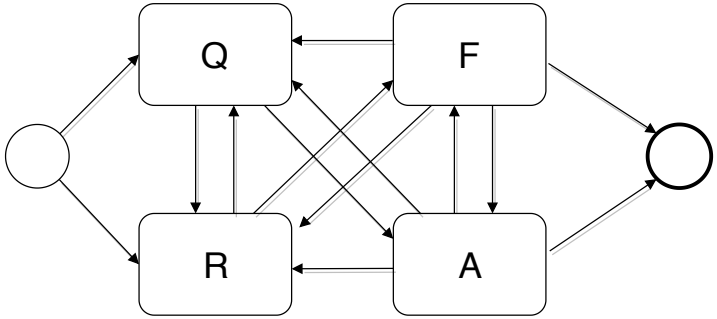


Model Fitness

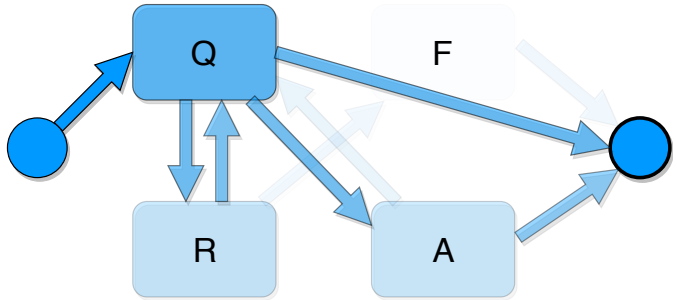
- **Fitness**/dialogue: QRFA Min 0.65 Max 1 Average 0.96 SD 0.04

  COR Min 0.38 Max 0.93 Average 0.67 SD 0.1

- Dialogues with **fitness=1:** QRFA 0.61 COR 0.01



SCS

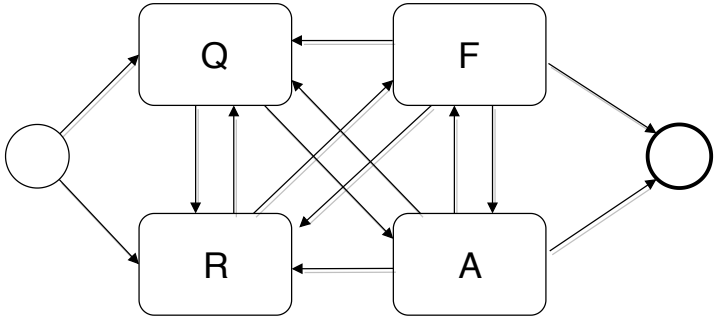


Model Generalisation

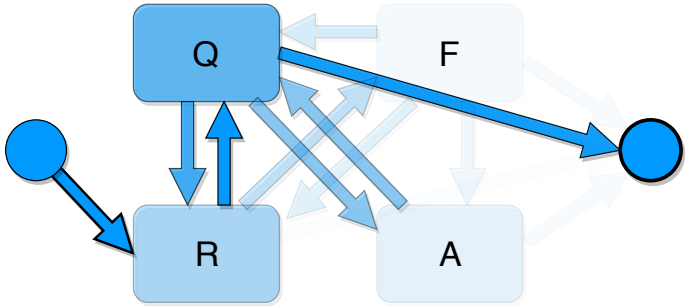
- **Fitness**/dialogue: QRFA Min 0.8 Max 1 Average 0.99 SD 0.02

  COR Min 0.53 Max 0.91 Average 0.7 SD 0.05

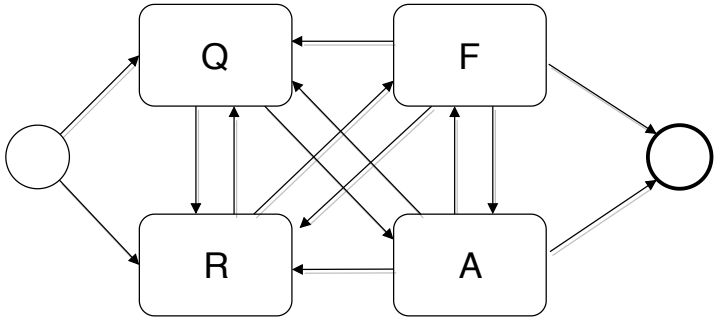
- **Break-down detection**: P 0.92 R 0.55



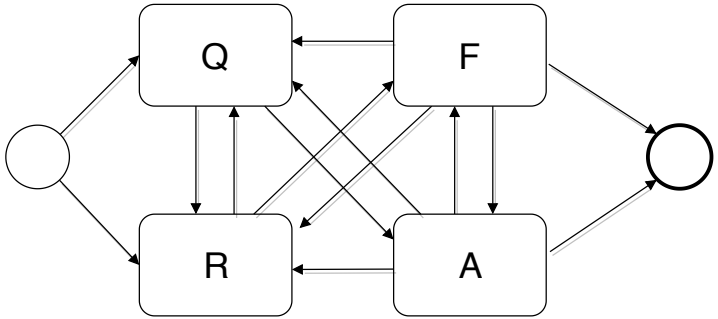
DSTC2



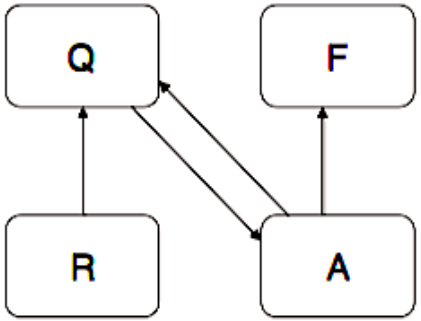
Interaction Modes



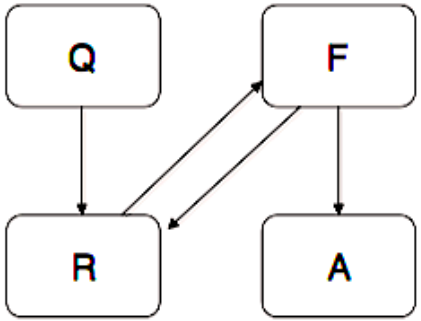
Interaction Modes



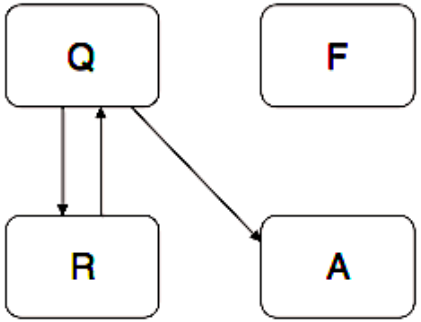
(a) Question Answering



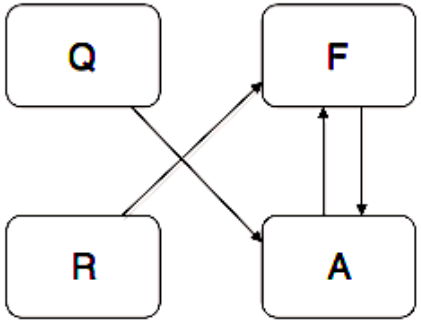
(c) Offer Refinement



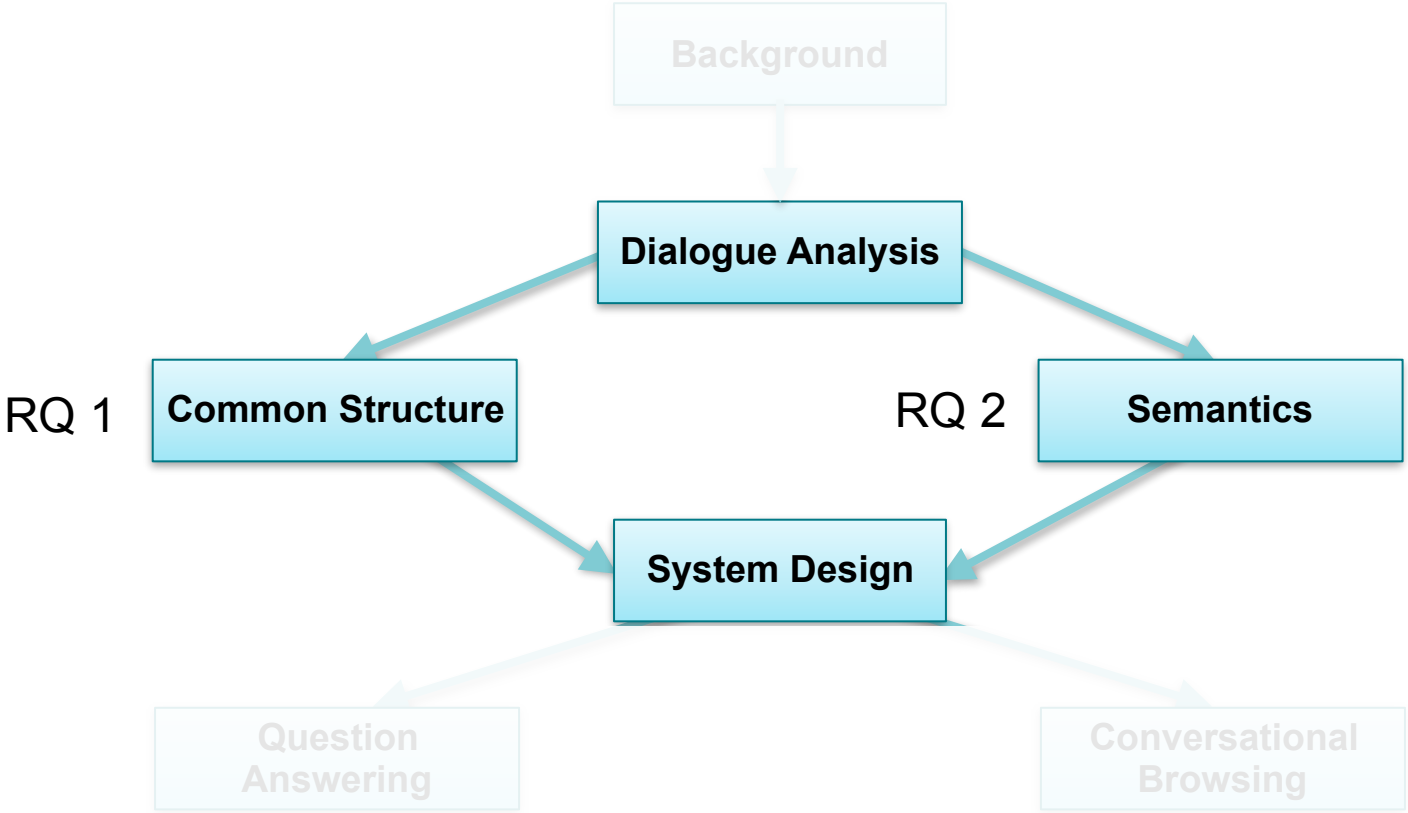
(b) Query Refinement



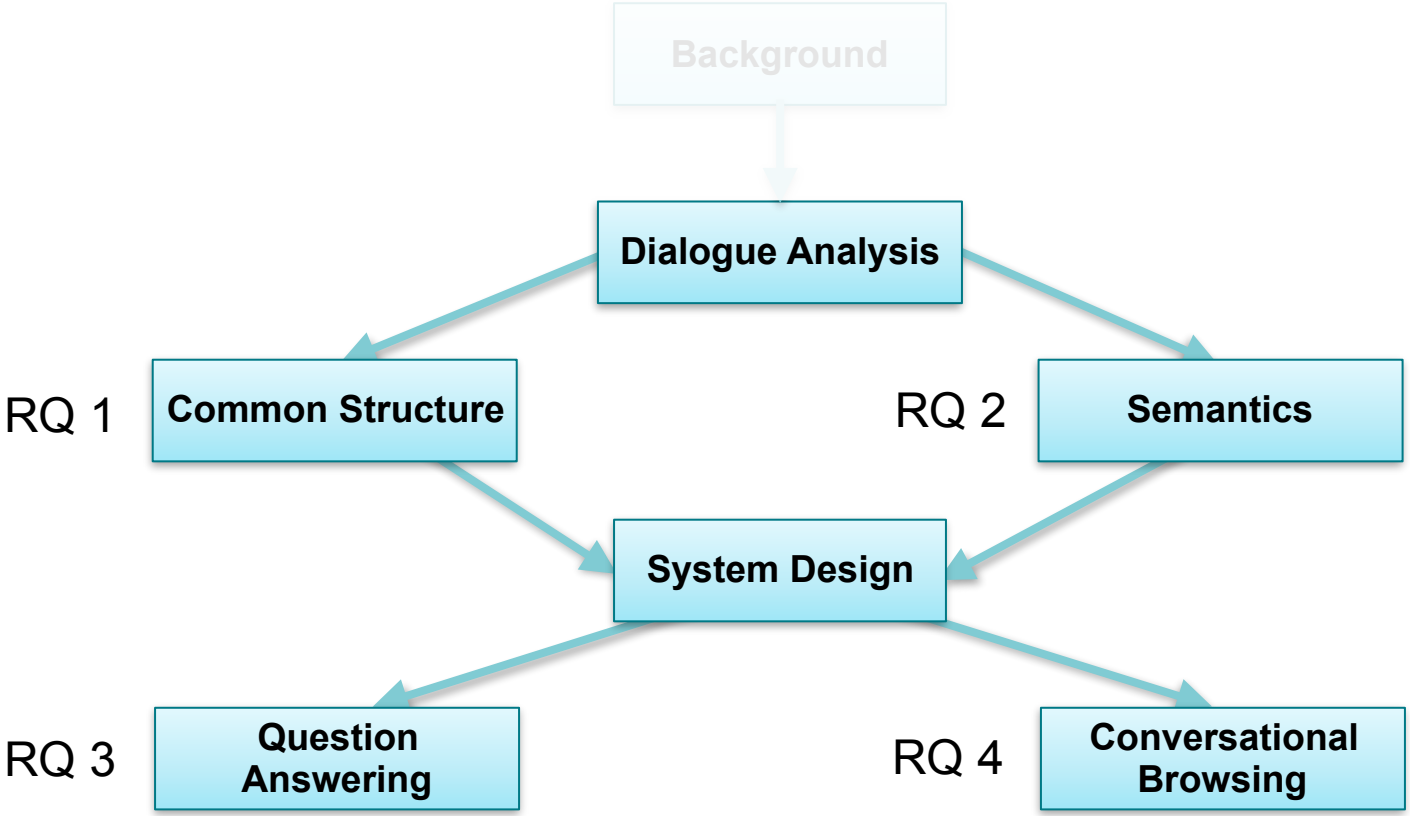
(d) Answer Refinement



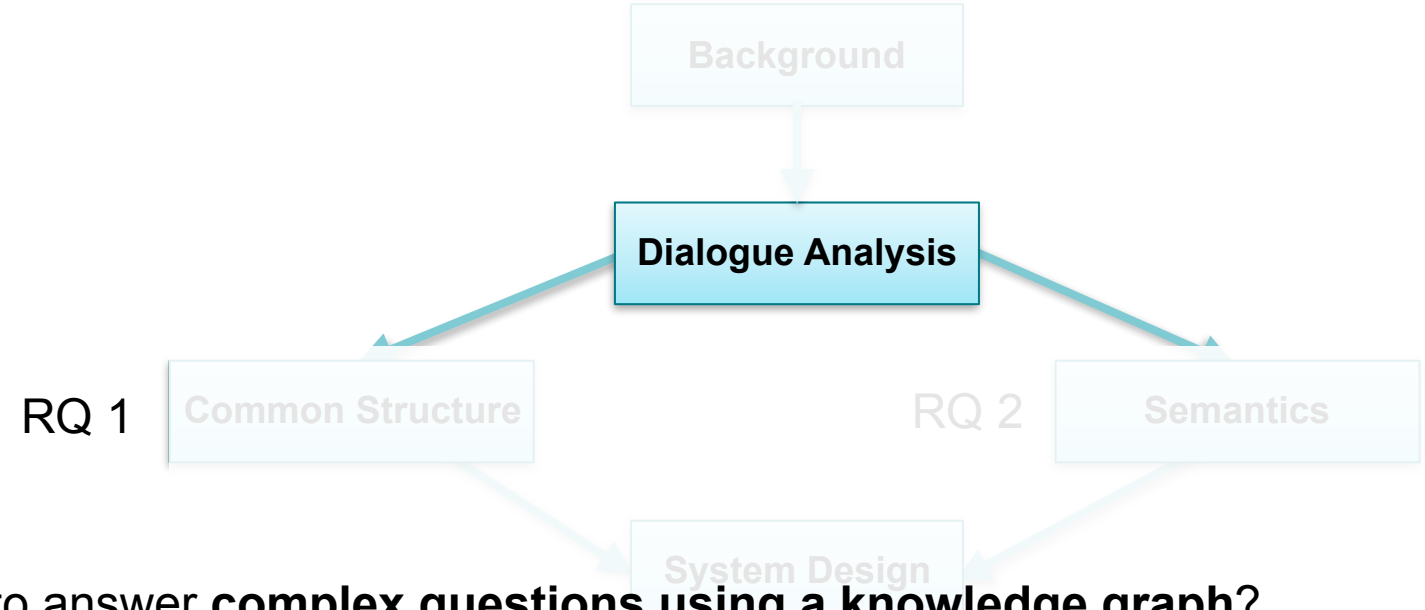
Outline


Background
Dialogue Analysis
RQ 1
Common Structure
RQ 2
Semantics
System Design
Question Answering
Conversational Browsing

Outline


Background
Dialogue Analysis
RQ 1
Common Structure
RQ 2
Semantics
System Design
RQ 3
Question Answering
RQ 4
Conversational Browsing

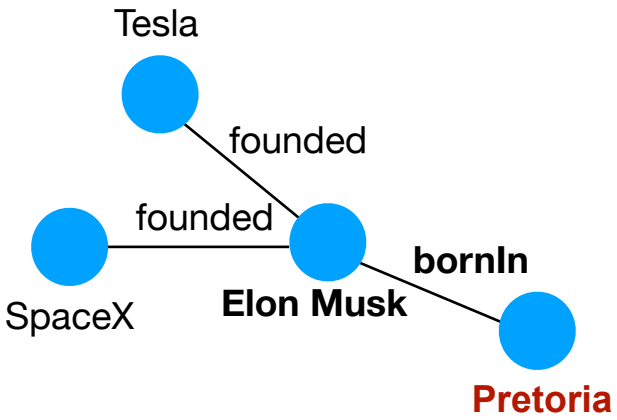
Outline



How to answer **complex questions using a knowledge graph**?

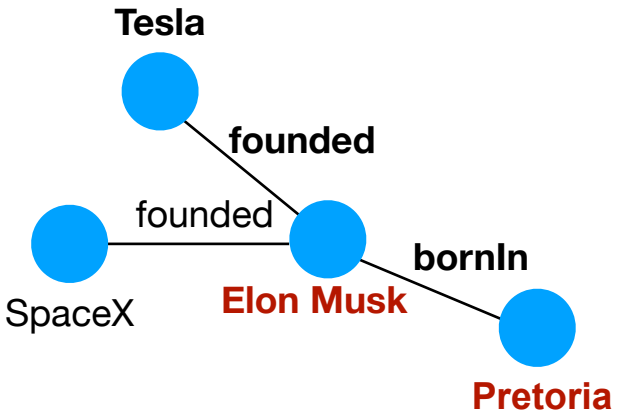
Simple Question Answering

Where is the **Elon Musk** **born**?



Complex Question Answering

Where is the **founder** of **Tesla** **born**?

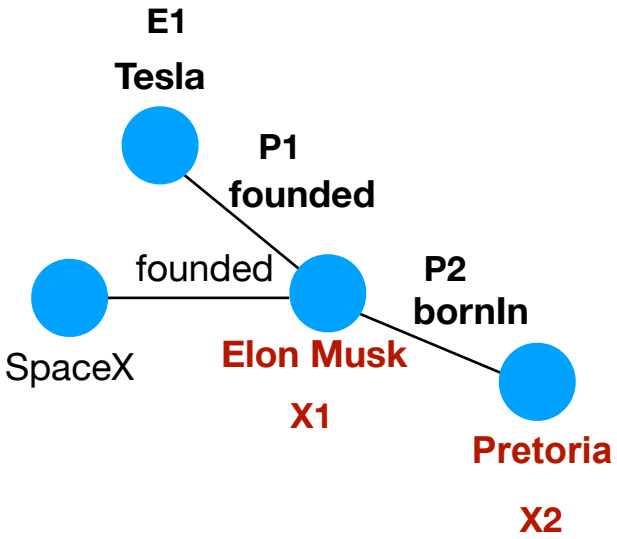


Message Passing

Parsing

Where is the **founder** of **Tesla** **born**?

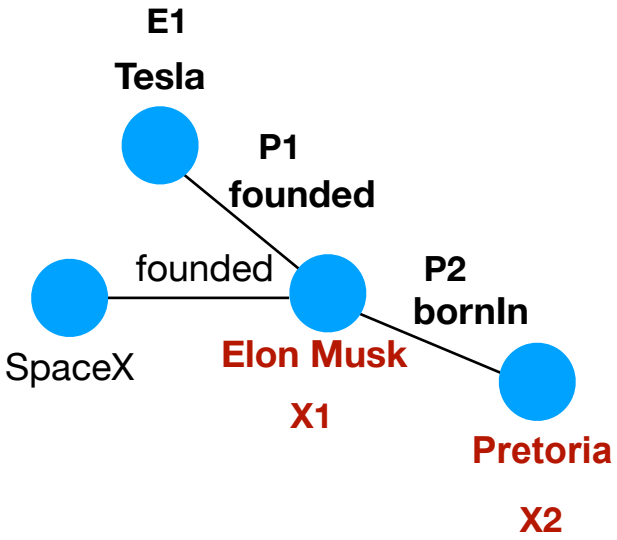
P1 **E1** **P2**



Matching

Where is the **founder** of **Tesla** **born**?

P1 E1 P2



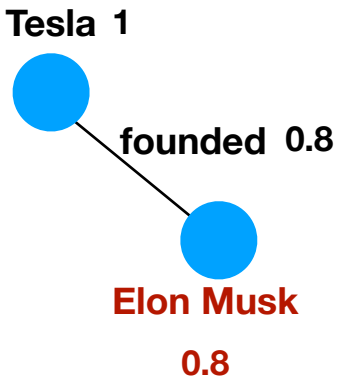
P1		E1	
founder	1	Tesla	1
founded	0.8	Nicola Tesla	0.6
		Tesla coil	0.5
P2			
bornIn	0.8		

Message Passing I

Where is the **founder** of **Tesla** **born**?

P1 E1 P2

P1		E1	
founder	1	Tesla	1
founded	0.8	Nicola Tesla	0.6
		Tesla coil	0.5



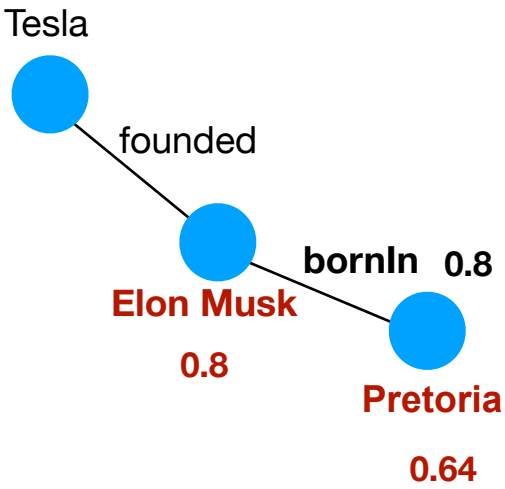
Message Passing II

Where is the **founder** of **Tesla** **born**?

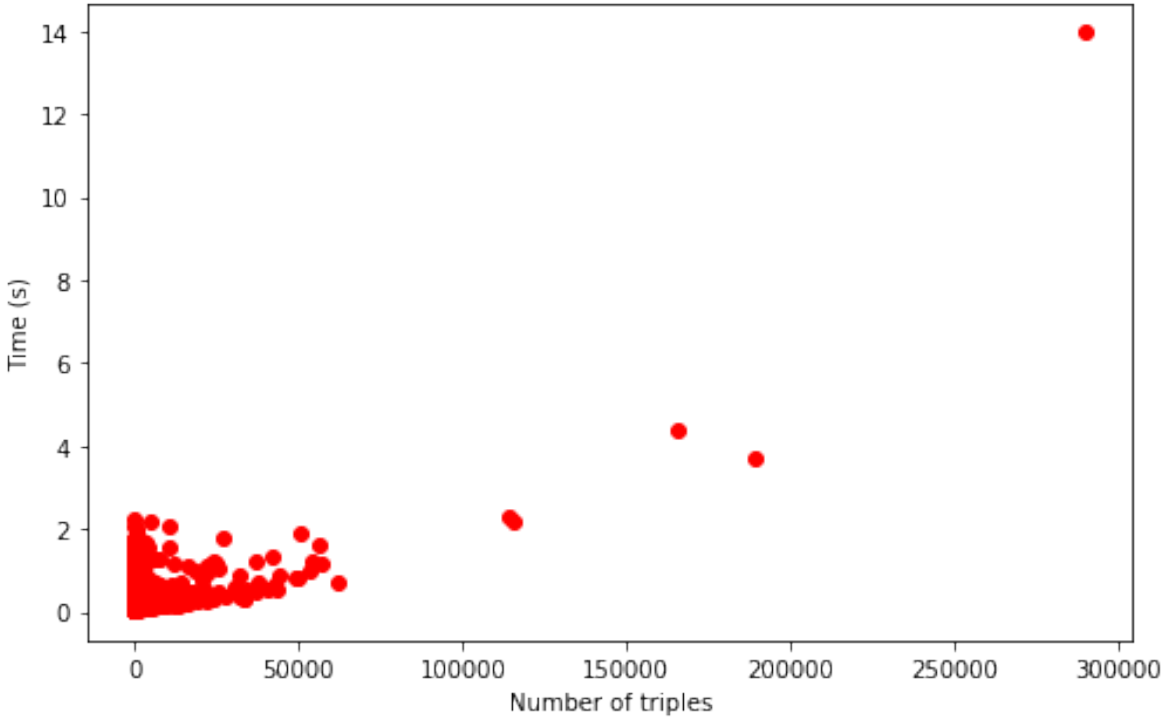
P1 **E1** **P2**

P2 **E2**

bornIn **0.8** **Elon Musk** **0.8**



Results: Scalability



Processing times per question from the LC-QuAD test split
(Min: 0.01s. Median: 0.68s. Mean: 0.72s. Max: 13.97s)

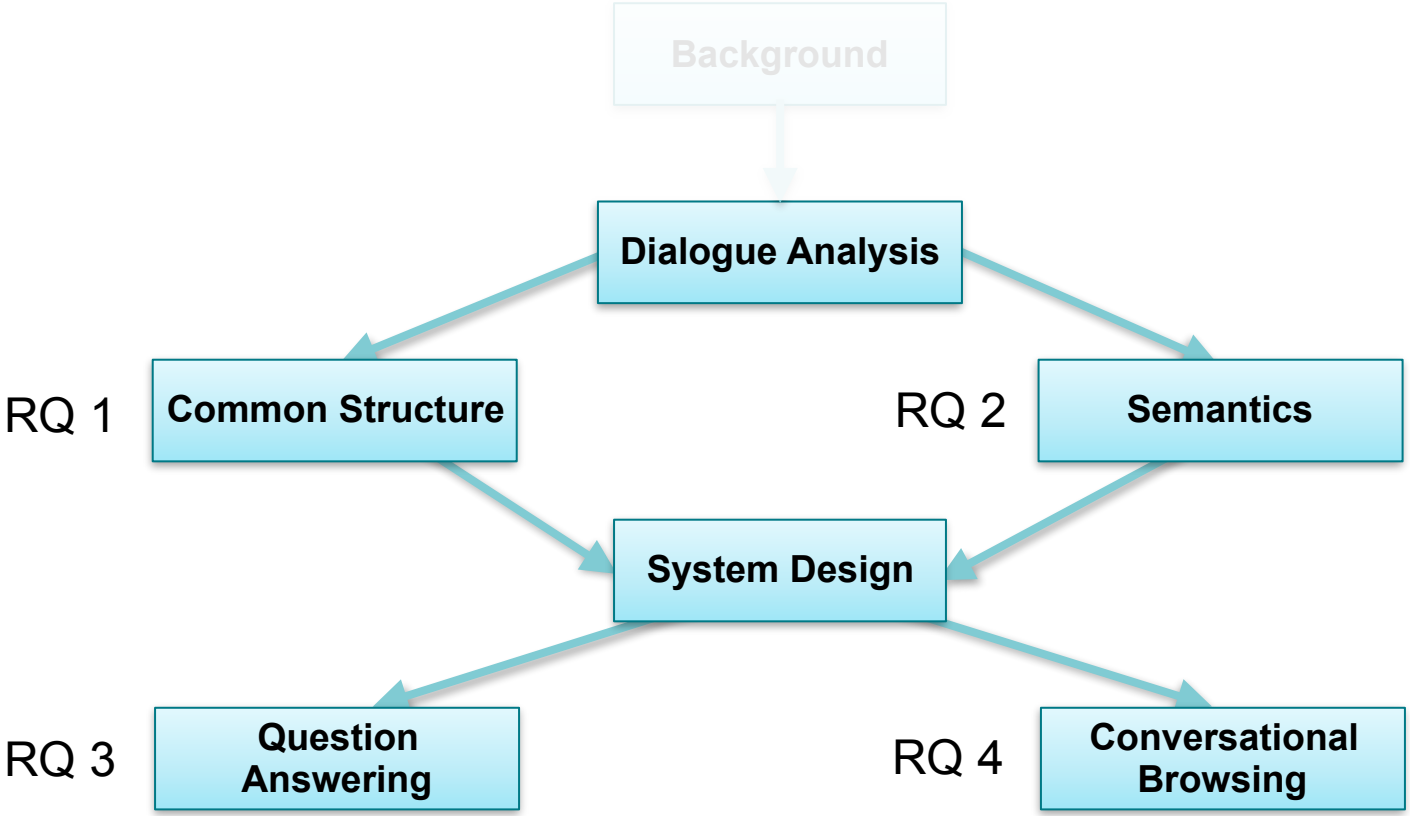
http://www.rdfhdt.org

Results: Performance

Approach	*P*	*R*	*F*	*Runtime*
Baseline	**0.22**	**0.38**	**0.28**	**1.50s/q**
Us	**0.25**	**0.50**	**0.33**	**0.72s/q**

Baseline: WDAqua, template-based SPARQL generator

Outline


Background
Dialogue Analysis
RQ 1
Common Structure
RQ 2
Semantics
System Design
RQ 3
Question Answering
RQ 4
Conversational Browsing

List of Publications

Vakulenko, Svitlana, et al. "QRFA: A Data-Driven Model of Information Seeking Dialogues. *ECIR*. 2019. Best User Paper Award

Vakulenko, Svitlana, et al. "Measuring semantic coherence of a conversation." *ISWC*. 2018. Spotlight paper

Vakulenko, Svitlana, et al. "Message Passing for Complex Question Answering over Knowledge Graphs." *CIKM.* 2019.

Vakulenko, Svitlana, et al. "Conversational browsing: Dialog-based Access to Structured Information Sources." *Under review*. 2019.

List of Publications

1. Svitlana Vakulenko, Javier D. Fernandez, Axel Polleres, Maarten de Rijke and Michael Cochez. Message Passing for Complex Question Answering over Knowledge Graphs. CIKM. 2019.
2. Svitlana Vakulenko, Vadim Savenkov and Maarten de Rijke. Conversational browsing: Dialog-based Access to Structured Information Sources. (Under review)
3. Sophia Keyner, Vadim Savenkov and Svitlana Vakulenko. Open Data Chatbot. ESWC 2019.
4. Svitlana Vakulenko, Kate Revoredo, Claudio Di Ciccio and Maarten de Rijke. QRFA: A Data-Driven Model of Information Seeking Dialogues. ECIR. 2019. *Best paper (User track).*
5. Svitlana Vakulenko, Maarten de Rijke, Michael Cochez, Vadim Savenkov and Axel Polleres. Measuring Semantic Coherence of a Conversation. ISWC. 2018. *Spotlight paper*.
6. Svitlana Vakulenko, Ilya Markov and Maarten de Rijke. Conversational Exploratory Search via Interactive Storytelling. SCAI@ICTIR. 2017.
7. Svitlana Vakulenko and Vadim Savenkov. TableQA: Question Answering on Tabular Data. SEMANTiCS. 2017. *Best poster and demo nomination*.
8. Sebastian Neumaier, Vadim Savenkov and Svitlana Vakulenko. Talking Open Data. ESWC. 2017.
9. Kathrin Figl, Samuel Kieling, Christiane Rank, Svitlana Vakulenko. The Effect of Fake News Flags on the Believability of Social Media Posts. ICIS. 2019.
10. Svitlana Vakulenko, Lyndon Nixon, and Mihai Lupu. Character-based Neural Embeddings for Tweet Clustering. SocialNLP@EACL. 2017.
11. Albert Weichselbraun, Stefan Gindl, Fabian Fischer, Svitlana Vakulenko, and Arno Scharl. Aspect-based extraction and analysis of affective knowledge from social media streams. In IEEE Intelligent Systems. 2017.
12. Svitlana Vakulenko, Albert Weichselbraun, and Arno Scharl. Detection of valid sentiment-target pairs in online product reviews and news media articles. WI. 2016.
13. Svitlana Vakulenko, Max Gobel, Arno Scharl and Lyndon Nixon. Know which way the Wind Blows: Visualising the Propagation of News on the Web. NewsIR@ECIR. 2016.
14. Gregory Katsios, Svitlana Vakulenko, Anastasia Krithara and Georgios Paliouras. Towards Open Domain Event Extraction from Twitter: REVEALing Entity Relations. DERIVE@ESWC. 2015.
15. Svitlana Vakulenko, Oliver Muller, Jan vom Brocke. Enriching iTunes App Store Categories via Topic Modeling. ICIS. 2014.
16. Nadine Reuter, Svitlana Vakulenko, Jan vom Brocke, Stefan Debortoli, Oliver Muller. Identifying the Role of Information Systems in Achieving Energy-Related Environmental Sustainability using Text Mining. ECIS. 2014.
17. Andrea Herbst, Alexander Simons, Jan vom Brocke, Oliver Muller, Stefan Debortoli, Svitlana Vakulenko. Identifying and characterizing topics in enterprise content management: a latent semantic analysis of vendor case studies. ECIS. 2014.
18. Peep Kungas, Siim Karus, Svitlana Vakulenko, Marlon Dumas, Cristhian Parra, Fabio Casati. Reverse-engineering conference rankings: what does it take to make a reputable conference? Scientometrics. 2013.

Contributions

- **Conversation mining** *methodology* + **QRFA** *data-driven model*
- **Semantic coherence** as a classification *task* + *experimental evaluation*
- **Message-passing** *algorithm* for complex question answering from KGs
- *Task, dataset, prototype* + *user studies* for **conversational browsing**

Acknowledgement

This work was supported by the following projects:

EU H2020 programme under the MSCA-RISE agreement 645751 (RISE_BPM)

Project Open Data for Local Communities funded by the Austrian Federal Ministry of Transport, Innovation and Technology (BMVIT) under the program "ICT of the Future", between November 2016 and April 2019.

More information https://iktderzukunft.at/en/

Austrian Ministry
for Transport,
Innovation and Technology