Towards Tackling Winograd Schemas
Applying Linguistic Theories and World Knowledge

Peter Schüller
Measuring Artificial Intelligence

- Turing Test:
  - Jury is chatting with Computer and with Human.
  - Jury has to find out who is Computer and who is Human.
  - Limited time.

Problems:
- based on deception
- no objective measurement, time intensive

- Winograd Schemas as a new Measure of Artificial Intelligence
  - Natural Language Processing task
  - Anaphora Resolution (Winograd 1972)

Advantages:
- clearly defined natural task
- requires reasoning, understanding, and world knowledge
- automatic measurement possible
The city councilmen refused the demonstrators a permit because they feared violence. Who feared violence?

A1 The demonstrators.

A2 The city councilmen.
Winograd Schema Challenge (Levesque, Davis, and Morgenstern 2012)

Q  The city councilmen refused the demonstrators a permit because they feared violence. Who feared violence?
A1  The demonstrators.
A2  The city councilmen.

Q  The city councilmen refused the demonstrators a permit because they advocated violence. Who advocated violence?
Winograd Schema Challenge (Levesque, Davis, and Morgenstern 2012)

Q  The [city councilmen]$_c$ refused the [demonstrators]$_d$ a permit because [they]$_{c,d}$ feared violence. (Who feared violence?)

A1  The demonstrators.

A2  The city councilmen.

Q  The city councilmen refused the demonstrators a permit because they advocated violence. Who advocated violence?

- Two NPs of the same gender, animacy, and number.
- One ambiguous pronoun reference to these NPs.
- Text comprehension with one binary question, coreference determines outcome of the question.
- Switching one special word changes the answer.

⇒ Easily solved by humans, not easily solvable by computers
Natural Language Processing: Statistical versus Symbolical

- Statistical:
  - Corpora with Annotations (manual effort)
  - Learning Algorithms (strengths and weaknesses come in here)
    
    Machine translation on the web, search, ...
  - Robust (+) but correct only to certain extent (-)
  - Big Success
Natural Language Processing: Statistical versus Symbolical

► Statistical:
  ▶ Corpora with Annotations (manual effort)
  ▶ Learning Algorithms (strengths and weaknesses come in here)

  Machine translation on the web, search, ...

  ▶ Robust (+) but correct only to certain extent (-)
  ▶ Big Success

► Symbolical:
  ▶ Rules (if . . . then)
  ▶ Knowledge bases (collections of rules, manual effort)
  ▶ Reasoning Algorithms (e.g., deduction, abduction)

  “A wall that becomes higher with each step towards it.”
  (Kemal Oflazer, 2013, pers. comm.)

  ▶ Precise (+) but Fragile (-)
  ▶ Out of Fashion

Claim of Winograd Schema Challenge: must be solved symbolically
Applying Linguistic Theories and World Knowledge

- Cultural Conventions
- Text Structure (in this talk: Rhetorical Structure Theory)
- Pragmatics (in this talk: Relevance Theory)
Applying Linguistic Theories and World Knowledge

- Cultural Conventions
- Text Structure (in this talk: Rhetorical Structure Theory)
- Pragmatics (in this talk: Relevance Theory)
- **Foundations** for the above: Semantics and Reasoning
World Knowledge and Cultural Conventions (1)

(1) [Joan]j made sure to thank [Susan]s for all the help [she]s had given

(2) [Joan]j made sure to thank [Susan]s for all the help [she]j had received.

• (Linguistic) Knowledge
  ‘A makes sure to do B’ ⇒ ‘A does B’
  ‘A gives help to B’ ⇔ ‘B receives help from A’

• Conventions (per default reasoning, exceptions are possible):
  ‘A gives help to B’ ⇒ \textit{usually} ‘B thanks A’
  (‘A gives help to B’ ⇒ ‘B has a plausible reason to thank A’)

World Knowledge and Cultural Conventions (2)

(3) The [drain]$_d$ is clogged with [hair]$_h$. [It]$_h$ has to be removed.

(4) [The drain]$_d$ is clogged with [hair]$_h$. [It]$_d$ has to be cleaned.

• (Linguistic) Knowledge

‘A clogs B’ and ‘A is removed from B’ ⇒ ‘A no longer clogs B’
‘B is clogged’ and ‘someone cleans B’ ⇒ ‘B no longer is clogged’
‘A has to be done’ ⇒ ‘A is a goal to be achieved’
‘A has to be done’ ⇒ ‘result of A is desirable for the speaker’

• (Conventions) Usually...

‘we do not want to have clogged drains’
‘we want to maintain the house’ (not remove the drain)
‘we want to have clean hair’ (if the hair is connected to the body)
(3) The [drain]$_d$ is clogged with [hair]$_h$. [lt]$_h$ has to be removed.

(4) [The drain]$_d$ is clogged with [hair]$_h$. [lt]$_d$ has to be cleaned.

• (Linguistic) Knowledge
  ‘A clogs B’ and ‘A is removed from B’ ⇒ ‘A no longer clogs B’
  ‘B is clogged’ and ‘someone cleans B’ ⇒ ‘B no longer is clogged’
  ‘A has to be done’ ⇒ ‘A is a goal to be achieved’
  ‘A has to be done’ ⇒ ‘result of A is desirable for the speaker’

• Conventions (usually . . . )
  ‘we do not want to have clogged drains’
  ‘we want to maintain the house’ (not remove the drain)
  ‘we want to have clean hair’ (if the hair is connected to the body)
Text Structure

(5) The [city councilmen]\textsubscript{c} refused the [demonstrators]\textsubscript{d} a permit because [they]\textsubscript{c} feared violence.

(6) The [city councilmen]\textsubscript{c} refused the [demonstrators]\textsubscript{d} a permit because [they]\textsubscript{d} advocated violence.

- Utterances depend on each other (overt or covert)
- Conjunctions can indicate overt dependency
  Examples: so, because, unless, although, but, for
- Dependency evokes expectations on the meaning
  Example: ‘A because B’ \(\Rightarrow\) B should be a plausible cause for A
  \(\Rightarrow\) A should be a plausible result of B
- **Cohesion** (binding principles are obeyed) restricts less than
  **Coherence** (semantical dependencies are fulfilled)
Some Theories of Text Structure

- **Rhetorical Structure Theory (RST, Mann and Thompson 1988)**
  Systematic and detailed theory, constraints in natural language.
  Used for *generating* discourses with computers (Hovy 1993).
  Used for *parsing* overt discourse structure (Marcu 2000).

- **Coherence and Structure of Discourse (Hobbs 1985)**
  Similar to RST, but simpler, and based on formal logic.

- **G&S (Grosz and Sidner 1986)**
  Attention (focus), intention, and sequence of discourses.

- **Discourse Representation Theory (Kamp and Reyle 1993)**
  Formal model-theoretic semantics with multiple sentences.

- **Centering Theory (Grosz, Joshi, and Weinstein 1995)**
  Relate coherence of discourse with minimal processing load.

- **Veins Theory (Cristea, Marcu, Ide, and Tablan 2000)**
  Statistical anaphora disambiguation (application of RST).
Rhetorical Structure Theory (RST)

- Hierarchical Text Structure with Nucleus (head) and Satellite
- Schemas with constraints and effect

(7) \[ I \text{ took the } [\text{water bottle}]_w \text{ out of the } [\text{backpack}]_b \] \_1
\[ \text{so that } [\text{it}]_{w,b?} \text{ would be lighter} \] \_2.

1-2

Volitional Cause

1 (Nucleus) 2 (Satellite)

C  N (1): volitional action performed by an agent A
C  N+S (1-2): S is a situation that could have caused A to do N
C  N+S (1-2): the reader can consider N (on its own) unmotivated
C  N is more central to the writer's purpose than S
E  reader recognizes situation in S as cause for action in N
Rhetorical Structure Theory (RST)

- Hierarchical Text Structure with Nucleus (head) and Satellite
- Schemas with constraints and effect

(7) \[ I \text{ took the } [ \text{water bottle}_w \text{ out of the } [ \text{backpack}_b ]_1 \text{ so that } [ \text{it}_w,b? \text{ would be lighter} ]_2. \]

1-2

Volitional Cause

1 (Nucleus) 2 (Satellite)

C N (1): volitional action performed by an agent A
C N+S (1-2): S is a situation that could have caused A to do N
C N+S (1-2): the reader can consider N (on its own) unmotivated
C N is more central to the writer's purpose than S
E reader recognizes situation in S as cause for action in N
Rhetorical Structure Theory (RST)

- Hierarchical Text Structure with Nucleus (head) and Satellite
- Schemas with constraints and effect

(7) \[ \text{I took the [water bottle]}_w \text{ out of the [backpack]}_b \]_1
\[ \text{so that [it]}_w,b? \text{ would be handy} \]_2.

1-2
Volitional Cause
1 (Nucleus) \quad 2 (Satellite)

C N (1): volitional action performed by an agent A
C N+S (1-2): S is a situation that could have caused A to do N
C N+S (1-2): the reader can consider N (on its own) unmotivated
C N is more central to the writer's purpose than S

E reader recognizes situation in S as cause for action in N
Issues coming up when using RST

- How to detect relations?
- How many and which relations are there?
- How plausible is plausible enough?
- Reasoning about belief, knowledge, intentions ...
Relevance Theory (Wilson and Sperber 2006)

- Inferential Pragmatics: ‘explain how the hearer infers the speaker’s meaning on the basis of the evidence provided’.
- Utterance creates expectation of its own optimal relevance.
- Relevance increases with cognitive effect (implications, revision of assumptions)
- Relevance decreases with processing effort (perception, memory, inference)

[Sam’s drawing]$_s$ was hung just above [Tina’s $∅$]$_t$ and it did look much better with another one below [it]?

[Sam’s drawing]$_s$ was hung just above [Tina’s $∅$]$_t$ and it did look much better with another one above [it]?.
Application of Relevance Theory

There is a $[\text{gap}]_g$ in the $[\text{wall}]_w$.  
You can see the garden $\text{through } it_g$.  

There is a $[\text{gap}]_g$ in the $[\text{wall}]_w$.  
You can see the garden $\text{behind } it_w$.  

The $[\text{drain}]_d$ is clogged with $[\text{hair}]_h$.  $[\text{It}]_d$ has to be removed.  

The $[\text{drain}]_d$ is clogged with $[\text{hair}]_h$.  $[\text{It}]_d$ has to be cleaned.  

$[\text{Tom}]_t$ threw his schoolbag down to $[\text{Ray}]_r$ after $[\text{he}]_t$ reached the top of the stairs.  

$[\text{Tom}]_t$ threw his schoolbag down to $[\text{Ray}]_r$ after $[\text{he}]_t$ reached the bottom of the stairs.  

There is a $[\text{pillar}]_p$ between me and $[\text{the stage}]_s$, and I can't see $[\text{it}]_p$.  
There is a $[\text{pillar}]_p$ between me and $[\text{the stage}]_s$, and I can't see around $[\text{it}]_p$.  
Application of Relevance Theory

There is a \([\text{gap}]_g\) in the \([\text{wall}]_w\).
You can see the garden through \(\text{it}_g\).
There is a \([\text{gap}]_g\) in the \([\text{wall}]_w\).
You can see the garden behind \(\text{it}_w\).

The \([\text{drain}]_d\) is clogged with \([\text{hair}]_h\). \([\text{It}]_h\) has to be removed.
\([\text{The drain}]_d\) is clogged with \([\text{hair}]_h\). \([\text{It}]_d\) has to be cleaned.
Application of Relevance Theory

There is a \([\text{gap}]_g\) in the \([\text{wall}]_w\). You can see the garden through it\(_g\).

There is a \([\text{gap}]_g\) in the \([\text{wall}]_w\). You can see the garden behind it\(_w\).

The \([\text{drain}]_d\) is clogged with \([\text{hair}]_h\). It\(_h\) has to be removed.

The drain\(_d\) is clogged with hair\(_h\). It\(_d\) has to be cleaned.

Tom\(_t\) threw his schoolbag down to Ray\(_r\) after he\(_?\) reached the top of the stairs.

Tom\(_t\) threw his schoolbag down to Ray\(_r\) after he\(_?\) reached the bottom of the stairs.
Application of Relevance Theory

There is a \([\text{gap}]_g\) in the \([\text{wall}]_w\).
You can see the garden through \(it_g\).

There is a \([\text{gap}]_g\) in the \([\text{wall}]_w\).
You can see the garden behind \(it_w\).

The \([\text{drain}]_d\) is clogged with \([\text{hair}]_h\). \([\text{It}]_h\) has to be removed.
\([\text{The drain}]_d\) is clogged with \([\text{hair}]_h\). \([\text{It}]_d\) has to be cleaned.

\([\text{Tom}]_t\) threw his schoolbag down to \([\text{Ray}]_r\)
after \([\text{he}]_2\) reached the top of the stairs.
\([\text{Tom}]_t\) threw his schoolbag down to \([\text{Ray}]_r\)
after \([\text{he}]_2\) reached the bottom of the stairs.

There is \([\text{a pillar}]_p\) between me and \([\text{the stage}]_s\),
and I can’t see \([\text{it}]_2\).

There is \([\text{a pillar}]_p\) between me and \([\text{the stage}]_s\),
and I can’t see around \([\text{it}]_2\).
Current and Future Work

▶ Everything!

for example Deferrence (Nunberg 2006):

Q  Who is the ham sandwich?
A  The ham sandwich is at table 7.

▶ Realize Relevance Theory in computational logic

▶ Corpus of schemas on limited vocabulary
→ Winograd Schema competition?
Current and Future Work

▶ Everything!

for example Deference (Nunberg 2006):

  Q  Who is the ham sandwich?
  A  The ham sandwich is at table 7.

▶ Realize Relevance Theory in computational logic

▶ Corpus of schemas on limited vocabulary
  → Winograd Schema competition?

Thank you very much for your attention!


