Who decides what a text means?

Graeme Hirst
University of Toronto
Acknowledgements

Stephen Regoczei  Jean-Pierre Corriveau
Jane Morris  Susan McRoy
Peter Heeman  David Traum
Philip Edmonds
Philosophically naïve
Philosophically naïve

– like CL itself
Text-meaning
Any complete utterance

– Spoken or written
– Interactive or not
– Long or short
Text-meaning
Text-meaning

Meaning of whole message, including subtext
Text-meaning

Meaning of whole message, including subtext

– Not just word- or sentence-meaning
Text-meaning

Meaning of whole message, including subtext

– Not just word- or sentence-meaning

– Could be more than, or less than, sum of sentence-meanings
What is the locus of text-meaning?
What is the locus of text-meaning?

1. Meaning is in the text itself
What is the locus of text-meaning?

1. Meaning is in the text itself
2. Meaning is in the writer / speaker
What is the locus of text-meaning?

1. Meaning is in the text itself
2. Meaning is in the writer / speaker
3. Meaning is in the reader / hearer
What is the locus of text-meaning?

1. Meaning is in the text itself
2. Meaning is in the writer / speaker
3. Meaning is in the reader / hearer

Or two of these, or all of these
We can ask the same question of lower-level linguistic elements

Words
Sentences
Semantic roles
Lexical relations
...

...
The same three answers are possible
The same three answers are possible

But they don’t have to be the same answer at each level
Effects of individual writer or reader are apparent only at text-meaning level, not below
Maybe …

Effects of individual writer or reader are apparent only at text-meaning level, not below

Or maybe conversely …

Individual writer or reader’s idiosyncrasies are dampened at text-meaning level
Maybe …

Effects of individual writer or reader are apparent only at text-meaning level, not below

Or maybe conversely …

Individual writer or reader’s idiosyncrasies are dampened at text-meaning level

Or maybe …
Three views of text-meaning
Three views of text-meaning

- My view here: Text is always a locus of meaning
Three views of text-meaning

- My view here: Text is always a locus of meaning
- The issue: Reader and/or writer as additional loci?
Three views of text-meaning
Three views of text-meaning

- Dominance of each view in CL varies with era
Three views of text-meaning

- Dominance of each view in CL varies with era
- CL has become less sophisticated in its view
The history of the philosophy of text-meaning in computational linguistics

1970–2009
1970–1985
1970–1985
1970–1985

- Simple utterances
1970–1985

- Simple utterances
- All texts are massively ambiguous; all texts are enthymemematic
1970–1985

• Simple utterances
• All texts are massively ambiguous; all texts are enthymematic
• Use knowledge of world and beliefs to interpret
1970–1985
1970–1985

Conceptual Structures
INFORMATION PROCESSING IN MIND AND MACHINE

J. F. Sowa
1970–1985
1970–1985

• Find the interpretation most consistent with what’s already known
1970–1985

• Find the interpretation most consistent with what’s already known

• Construe input as best match to own prior knowledge
1970–1985

• Find the interpretation most consistent with what’s already known

• Construe input as best match to own prior knowledge

3 Meaning is in the reader / hearer
Example:

The city councillors denied the demonstrators a permit because they were communists.

Who are the communists?
1985–1995
1985–1995
1985–1995

- Interactive dialogues
1985–1995

- Interactive dialogues
- Gricean and pragmatic theories of “real” language use
1985–1995

- Interactive dialogues
- Gricean and pragmatic theories of “real” language use
- Model the user to determine their goals and plans …
1985–1995

• Interactive dialogues
• Gricean and pragmatic theories of “real” language use
• Model the user to determine their goals and plans …
• … and hence real intent of their utterances
1985–1995
1985–1995
1985–1995
1985–1995

- A text means whatever the speaker thinks it means or intends it to mean
1985–1995

- A text means whatever the speaker thinks it means or intends it to mean
  
  Meaning is in the writer / speaker
1985–1995

- A text means whatever the speaker thinks it means or intends it to mean
  
  2 Meaning is in the writer / speaker

- The computer’s job is to read the user’s mind
1985–1995

Example:

Talking to domestic robot:
I’d like a beer
→ Bring me a beer and do it right now
1995–2009
1995–2009
1995–2009

• Large, non-interactive texts
1995–2009

• Large, non-interactive texts
• Statistical and machine-learning methods
1995–2009

FOUNDATIONS OF STATISTICAL NATURAL LANGUAGE PROCESSING

Christopher D. Manning and Hinrich Schütze

25
1995–2009
1995–2009

- Text regarded as *objet trouvé* (‘found object’)
1995–2009

- Text regarded as *objet trouvé* (‘found object’)
- Meaning is “extracted” by “processing” the words and their context
1995–2009

- Text regarded as *objet trouvé* (‘found object’)
- Meaning is “extracted” by “processing” the words and their context

1 Meaning is in the text
1995–2009

- Text regarded as *objet trouvé* (‘found object’)
- Meaning is “extracted” by “processing” the words and their context
  1 Meaning is in the text
- “The text is all we have.”
Examples:

Find articles on raptor migration in Colorado.

Find follow-ups to this news story.

Summarize this report.

Monitor this chat room.
Roles of the linguistic computer
Roles of the linguistic computer

Roles of the linguistic computer


1985–1995: Servant of the user
Roles of the linguistic computer

1985–1995: Servant of the user
1995–2009: Reader and transformer of text
Computational linguistics vacillates between the three views of locus of text-meaning
Computational linguistics vacillates between the three views of locus of text-meaning

But computational linguists don’t notice and don’t care
Computational linguistics vacillates between the three views of locus of text-meaning.

But computational linguists don’t notice and don’t care.

Philosophically naïve.
Two types of system
Two types of system

- **Observer**: Reads external text on behalf of a user
Two types of system

- **Observer**: Reads external text on behalf of a user
- **Conversant**: Actively participates in a dialogue with a user
CL’s naïve assumptions about meaning
CL’s naïve assumptions about meaning

- User or writer is perfect language user
CL’s naïve assumptions about meaning

- User or writer is perfect language user
- If observer: User’s knowledge and agenda are same as the writer’s
CL’s naïve assumptions about meaning

- User or writer is perfect language user
- If observer: User’s knowledge and agenda are same as the writer’s
- If conversant: System’s knowledge and agenda are same as user’s
CL’s naïve assumptions about meaning

• Meaning is conveyed solely by positives
CL’s naïve assumptions about meaning

- Meaning is conveyed solely by positives
- No distinction between meaning and interpretation
2009–2016
2009–2016

- Elimination of assumption of identical agendas
2009–2016

- Elimination of assumption of identical agendas
- Interpretation distinguished from meaning
2009–2016

- Elimination of assumption of identical agendas
- Interpretation distinguished from meaning
- Return of in-reader and in-writer views
Google has turned us all into researchers
Google has turned us all into researchers but with only an *impoverished* view of meaning.
In 2009

My meaning

Strings

Their meaning

Text

Strings

Match?
By 2016

My meaning in text

Their meaning in text

Match?
By 2016

My meaning in text → Match? ← Their meaning in text

1. What does this mean for me?
2. What are they trying to say?
What does this mean for me?
What does this mean for me?

- **Goal:** Research intermediaries that can interpret from the user’s perspective
What does this mean for me?

**Goal:** Research intermediaries that can interpret from the user’s perspective

To get at reader’s meaning, system first needs to understand their purpose and their viewpoint
What does this mean for me?

- A document may answer a user’s question without any intent by the author to do so.
What does this mean for me?

- A document may answer a user’s question without any intent by the author to do so.
- Especially abstract, wide-ranging, or unusual questions and query-oriented multi-document summarization.
2009–2016

Examples:

Find evidence that …

… Norway is capable of developing WMD

… society is too tolerant of drunk drivers

… the President is doing a great job
2009–2016

- Learning by reading
  - Integrating content of new document into existing knowledge base
“The text is all we have.”
“The text is all we have.”

We know our own beliefs and goals
What are they trying to say?
2009–2016

What are they trying to say?

- **Goal:** Research intermediaries that can interpret text from the writer’s perspective
2009–2016

What are they trying to say?

- **Goal:** Research intermediaries that can interpret text from the writer’s perspective
- Hermeneutic (interpretive) task
2009–2016

What are they trying to say?

- **Goal:** Research intermediaries that can interpret text from the writer’s perspective
- Hermeneutic (interpretive) task
- Intelligence gathering
2009–2016

- Examples:
2009–2016

• Examples:
  – Sentiment analysis and classification
2009–2016

- Examples:
  - Sentiment analysis and classification
  - Opinion extraction and ideological analysis
Examples:

- Sentiment analysis and classification
- Opinion extraction and ideological analysis
- Learning by reading: answering test questions
2009–2016

Examples:

- Sentiment analysis and classification
- Opinion extraction and ideological analysis
- Learning by reading: answering test questions
- Interlingual machine translation
“All we have is the text.”
“All we have is the text.”

We know the writer and the context
Future roles of the linguistic computer
Future roles of the linguistic computer

– Servant of the user
Future roles of the linguistic computer

– Servant of the user
– “Neutral” reader and transformer of text
Future roles of the linguistic computer

- Servant of the user
- “Neutral” reader and transformer of text
- Proxy for the world
Future roles of the linguistic computer

- Servant of the user
- “Neutral” reader and transformer of text
- Proxy for the world
- Proxy for the user in the world
Future roles of the linguistic computer

- Mediates between the user and the world
Future roles of the linguistic computer

- Mediates between the user and the world
- Interprets the world to me
Future roles of the linguistic computer

– **Mediates** between the user and the world

– Interprets the world to me

– Interprets me to the world
Recovering from misunderstanding
Interpretive freedom is not unlimited
Interpretive freedom is not unlimited

Even in the reader-based view, readers can be wrong
Interpretive freedom is not unlimited

The text is a given
Interpretive freedom is not unlimited

The text is a given

– Mishearing, misreading are errors
Interpretive freedom is not unlimited

The rules of language and linguistic processing are given
Interpretive freedom is not unlimited

The rules of language and linguistic processing are given

– Anaphora resolution, homonym disambiguation, phrase attachment, …
Interpretive freedom is not unlimited

But the text might be mis-generated with respect to intent
Interpretive freedom is not unlimited

But the text might be mis-generated with respect to intent

– Typos, malapropisms, slips of the tongue, …
Interpretive freedom is not unlimited

But the text might be mis-generated with respect to intent

– Typos, malapropisms, slips of the tongue, …

– Unintended ambiguities, misleading cues
If present text is unexpected or uninterpretable
If present text is unexpected or uninterpretable
then hypothesize a present or earlier misunderstanding
If present text is unexpected or uninterpretable

then hypothesize a present or earlier misunderstanding

by self or other
If present text is unexpected or uninterpretable
then hypothesize a present or earlier misunderstanding
by self or other

Re-interpret or clarify
Example

Data from Terasaki 1976
Example

**MOTHER:** Do you know who’s going to that meeting?

*Data from Terasaki 1976*
Example

MOTHER: Do you know who’s going to that meeting?

RUSS: Who?
Example

**MOTHER:** Do you know who’s going to that meeting?

**RUSS:** Who?

**MOTHER:** I don’t know.

Data from Terasaki 1976
MOTHER: Do you know who’s going to that meeting?

RUSS: Who?

MOTHER: I don’t know.

RUSS: Oh. Probably Mrs McOwen and some of the teachers.

Data from Terasaki 1976
Collaborative repair of misunderstanding

- Repair of text-level misunderstanding
Collaborative repair of misunderstanding

- Repair of text-level misunderstanding
- Speaker and listener negotiate and refine meaning of prior utterance
Collaborative repair of misunderstanding

- Repair of text-level misunderstanding
- Speaker and listener negotiate and refine meaning of prior utterance
- Integrates speaker-based and listener-based views of meaning
Collaborative repair of misunderstanding

• Repair of text-level misunderstanding
• Speaker and listener negotiate and refine meaning of prior utterance
• Integrates speaker-based and listener-based views of meaning
• Computational models of this process (McRoy and Hirst 1995)
Conclusion
Conclusion

- Three loci of text-meaning
  - in text, in writer, in reader
Conclusion

- Three loci of text-meaning
  - in text, in writer, in reader
- CL varies in its view
  - but has lately forgotten the writer and reader
Conclusion

- Three loci of text-meaning
  - in text, in writer, in reader
- CL varies in its view
  - but has lately forgotten the writer and reader
- New applications will bring them back
Conclusion

- Further sophistication in text-meaning
Conclusion

• Further sophistication in text-meaning
  — Collaborative construction of meaning in interaction and elicitation of knowledge
Conclusion

- Further sophistication in text-meaning
  - Collaborative construction of meaning in interaction and elicitation of knowledge
  - Searching for and reconciling different interpretations of text
Future role of the linguistic computer
Future role of the linguistic computer

– Mediation and reconciliation
Future role of the linguistic computer

- Mediation and reconciliation
- Peace in the Middle East