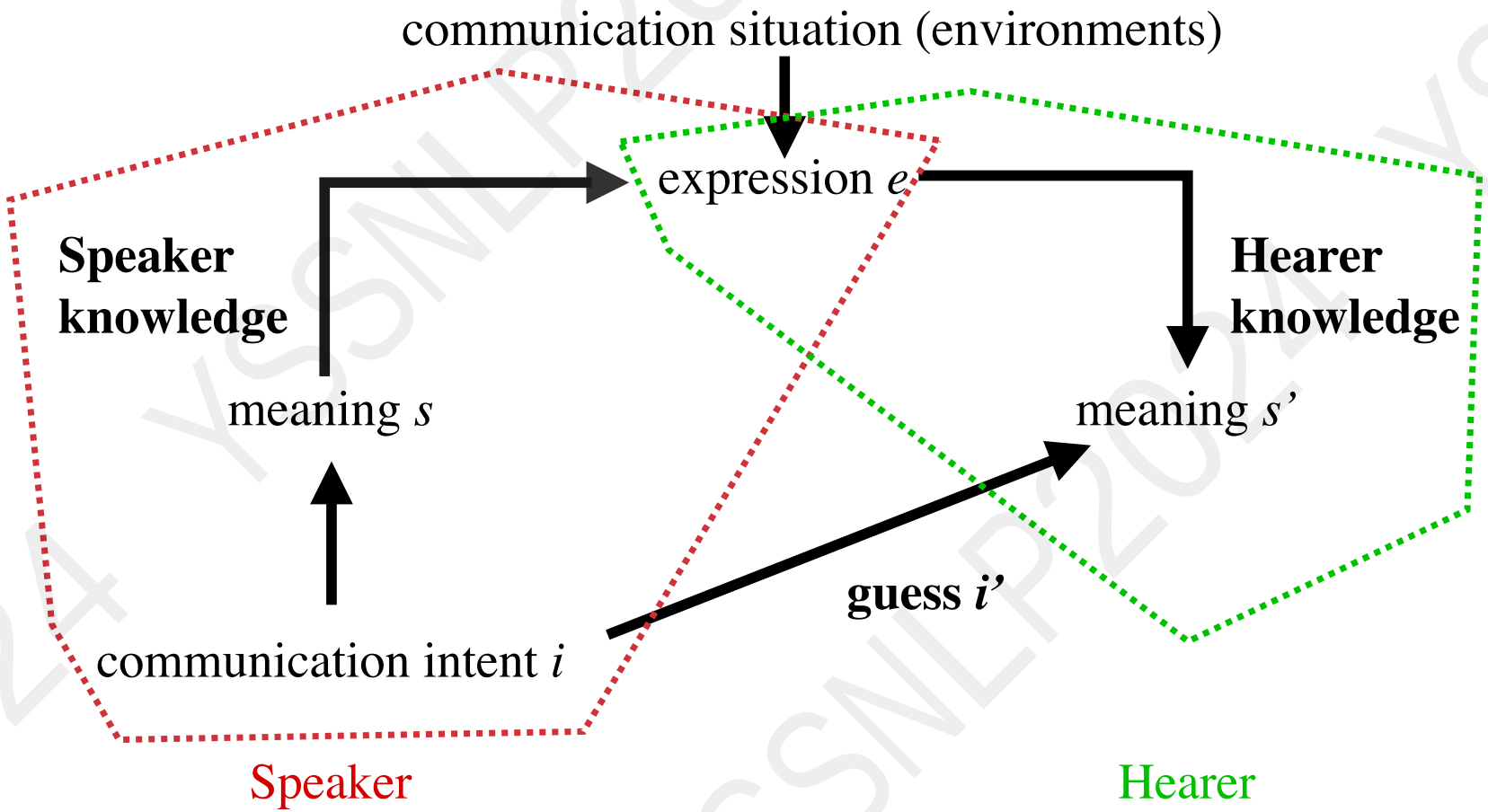
基于话语表示理论的语义分析

Semantic Parsing in Discourse Representation Theory

刘江鸣　　云南大学

2024年06月16日

Octopuses and Language (Bender and Koller, 2020)



- (Bender and Koller, 2020) Emily M. Bender, and Alexander Koller. Climbing towards NLU: On meaning, form, and understanding in the age of data. *ACL*. 2020.

Two Types of Text Understanding

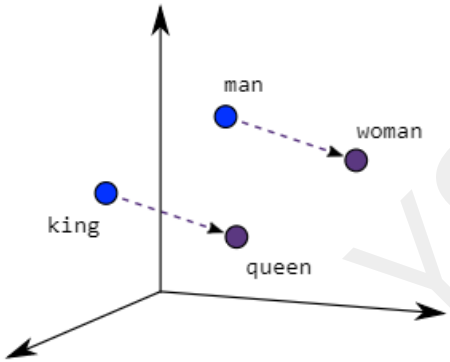
from Shay Cohen in RISE

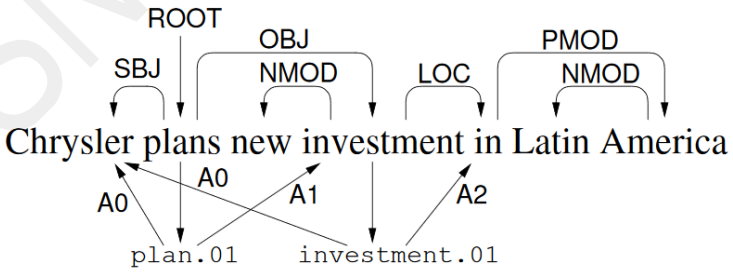
Understanding, loosely, can refer to two different things:

- Attaching meaning (representations?) which is divorced of direct denotation (large language model)
- Attaching meaning which is either executable and has **denotation** or is grounded in the environment (semantic parsing)

Continuous : gradient Symbolic: interpretable

Three Types of Semantics





b1

b1 ← x1

b3

b1 ← ¬ b2 ← x2 b2 ← s1 b3 ← e1 b1 ← t1

b1 ← time.n.08(t1)
b1 ← t1 = now
b3 ← touch.v.01(e1)
b3 ← Time(e1, t1)
b3 ← Theme(e1, x2)
b3 ← Agent(e1, x1)
b2 ← blue.a.01(s1)
b2 ← button.n.02(x2)
b2 ← Colour(x2, s1)

b1 ← person.n.01(x1)
b1 ← x1 = hearer

Distributed semantics
(word2vec, LDA, LLM, …)

Frame semantics
(SRL, AMR, …)

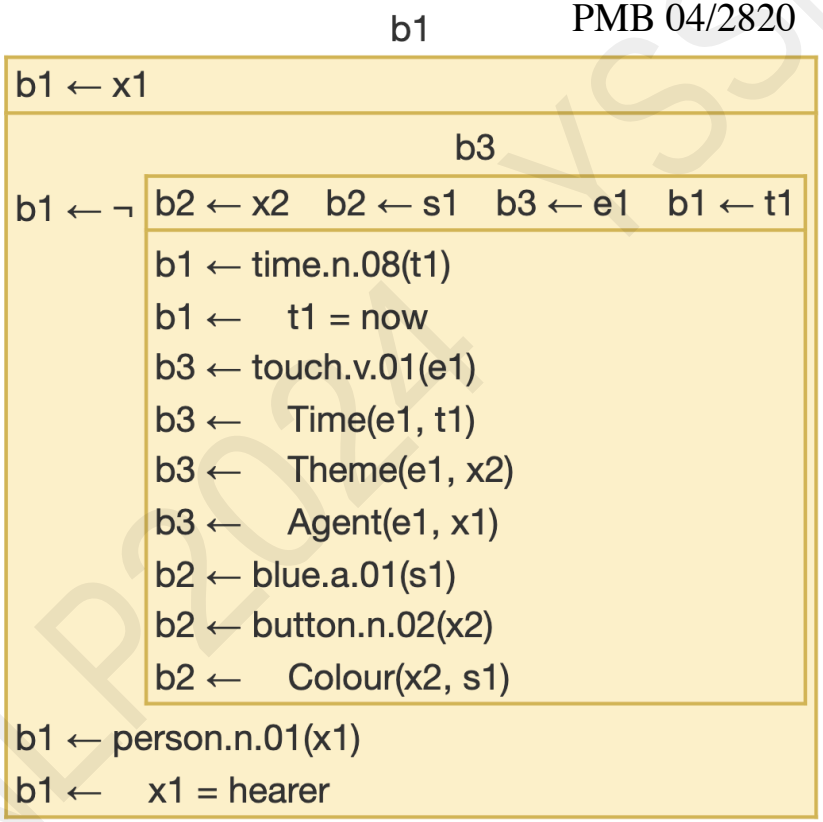
Model-theoretic semantics
(MRS, DRT, …)

Discourse Representation Theory (Kamp, 1981; Kamp and Reyle, 1993)

- Model-theoretic semantics
  - A special form of truth-conditional semantics
- Dynamic semantics
  - Compositional
- First-order logic

$\exists x_1(\text{person.n.01}(x_1) \wedge x_1 = \text{hearer} \wedge \neg \exists x_2 \exists s_1 \exists e_1 \exists t_1(\text{touch.v.01}(e_1) \ldots))$

PMB 04/2820



Don't touch that blue button.

- (Kamp, 1981) Hans Kamp. 1981. A Theory of Truth and Semantic Representation. In Formal Methods in the Study of Language, 1, 277-322.
- (Kamp and Reyle, 1993) Hans Kamp and Uwe Reyle. 1993. From Discourse to Logic; An Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory. Kluwer, Dordrecht

Discourse Representation Theory

- Logics
  - Propositional Logic (e.g., implication)

    Premise 1: $P \rightarrow Q$

    Premise 2: P

    Conclusion: Q

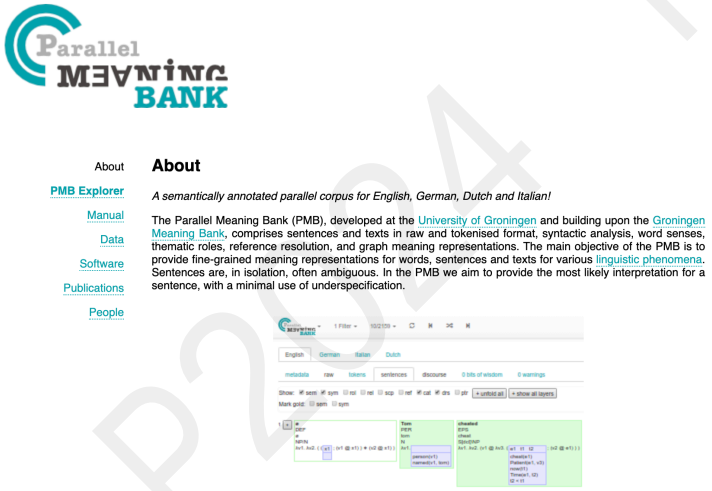
  - Modal Logic (e.g., possibility and necessity)

    "It is not necessary that X" is logically equivalent to "It is possible that not X".
    "It is not possible that X" is logically equivalent to "It is necessary that not X"

Discourse Representation Structure Parsing

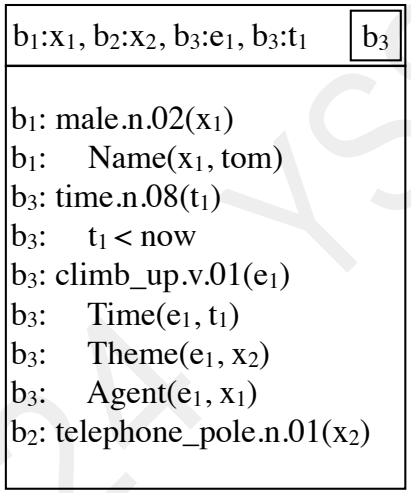
- Groningen Meaning Bank

- Parallel Meaning Bank v5.1.0

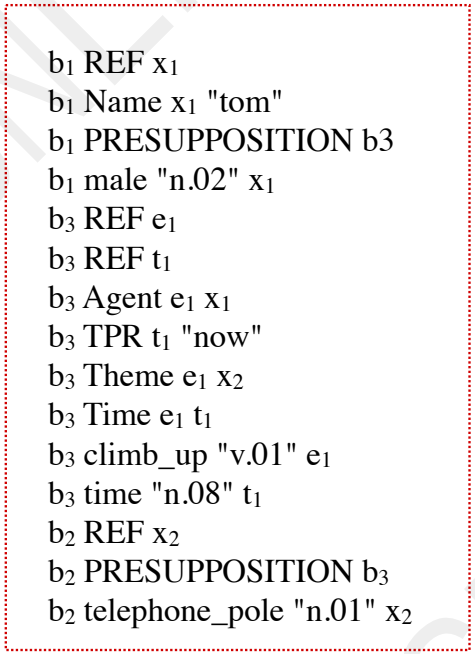


Discourse Representation Structure Parsing

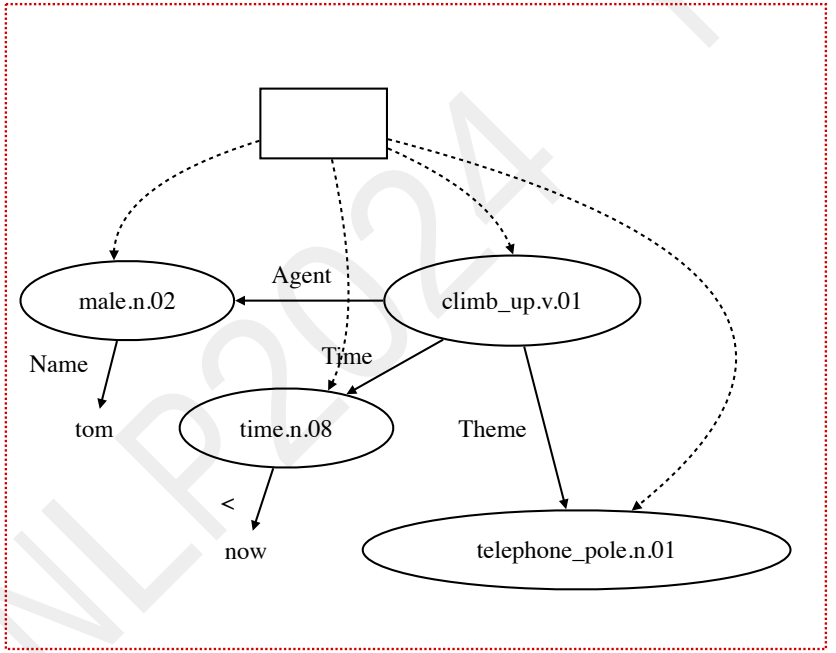
Three DRS notations



box



clause



graph

Discourse Representation Structure Parsing — Results

PMB 3.0.0[clause]	EN	DE	IT	NL
SPAR	40.8	/	/	/
SIM-SPAR	57.7	/	/	/
Boxer (Bos, 2015)	72.2	/	/	/
Neural Boxer (van Noord et al., 2018)	83.2	/	/	/
Neural Boxer (van Noord et al., 2019)	88.9	81.9	**80.5**	71.1
BILSTM-char (Wang et al., 2021)	88.1	/	/	/
MACT (Liu, 2024)	**89.1**	**82.7**	80.2	**80.1**

(Bos, 2015) Johan Bos. 2015. Open-domain semantic parsing with Boxer. In B. Megyesi (Ed.), Proceedings of NODALIDA 2015, Vilnius, Lithuania, pp. 301–304.
(van Noord et al., 2018) Rik van Noord, Lasha Abzianidze, Antonio Toral, and Johan Bos. 2018. Exploring Neural Methods for Parsing Discourse Representation Structures. TACL.
(van Noord et al., 2019) Rik van Noord, Antonio Toral, and Johan Bos. 2019. Linguistic Information in Neural Semantic Parsing with Multiple Encoders. In Proceedings of IWCS.
(Wang et al., 2021) Chunliu Wang, Rik van Noord, Arianna Bisazza, and Johan Bos. 2021. Input Representations for Parsing Discourse Representation Structures: Comparing English with Chinese. In Proceedings of ACL.
(Liu, 2024) Jiangming Liu. 2024. Model-Agnostic Cross-Lingual Training for Discourse Representation Structure Parsing. In Proceedings of LREC-COLING.

Discourse Representation Structure Parsing — Results

PMB 4.0.0[graph]	EN	DE	IT	NL
UD Boxer (Poelman et al., 2022)	81.8	77.5	79.1	75.8
Neural Boxer (Poelman et al., 2022)	92.5	74.7	75.4	71.6
MLM (Wang et al., 2023)	94.7	92.0	93.1	92.6
MACT (Liu, 2024a)	96.3	92.8	93.0	93.1
MACT (Liu, 2024b)	**97.0**	**94.7**	**94.3**	**94.4**

(Poelman et al., 2022) Wessel Poelman, Rik van Noord, and Johan Bos. 2022. Transparent semantic parsing with Universal Dependencies using graph transformations. In Proceedings of COLING.
(Wang et al., 2023) Chunliu Wang, Huiyuan Lai, Malvina Nissim, and Johan Bos. 2023. Pre-trained language-meaning models for multilingual parsing and generation. In Findings of ACL.
(Liu, 2024a) Jiangming Liu. 2024. Model-Agnostic Cross-Lingual Training for Discourse Representation Structure Parsing. In Proceedings of LREC-COLING.
(Liu, 2024b) Jiangming Liu. 2024. Soft Well-Formed Semantic Parsing with Score-Based Selection. In Proceedings of LREC-COLING.

Challenges and Future Works

- The propositional logic and modal logic calculus system.
- Incremental semantic parsing on long texts (dialogues).
- Factual evaluation, including propositional logics and modal logics.
- Universal semantics.
- Connections between symbolic semantics and continuous semantics (neural-symbolic inference).
- ……

Q & A