Introduction to Abstract Meaning Representation AMR 1 - Basis

Bernard ESPINASSE
Aix-Marseille Université (AMU)
LIS UMR CNRS 7020, Marseille, France.

April, 2019

- Introduction to AMR
- AMR tools
- AMR applications
- AMR and different natural languages

References (1)

Books, articles and reports :

- Banarescu L., Bonial C., Cai S., Georgescu M., Griffitt K., Hermjakob U., Knight K., Koehn P., Palmer M., and Schneider N. (2013). Abstract Meaning Representation for Sembanking. In Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse, 178–186, Sofia, Bulgaria: Association for Computational Linguistics. https://amr.isi.edu/a.pdf
- AMR 1.1 specification: http://www.isi.edu/ ulf/amr/help/amr-guidelines.pdf
- Palmer M. et al, The Proposition Bank: An Annotated Corpus of Semantic Roles, Comp. Linguistics, 31(1), 1-36, (2005).
- Banarescu L., Bonial C., Cai S., Georgescu M., Griffitt K., Hermjakob U., Knight K., Koehn P., Palmer M. & Schneider N. (2012). Abstract meaning representation (amr) 1.0 specification. In *Parsing on Freebase from Question-Answer Pairs."*
- Migueles-Abraira N., A Study Towards Spanish Abstract Meaning Representation, Master Thesis, 2017, University of the Basque Country.
- …

Courses/tutorials:

- N. Schneider, J. Flanigan, T. O'Gorman, "The Logic of AMR: Pratical, Unified Graph-Based Sentence Semantics for NLP", Tutorial at the 2015 Conference of the North American Chapter of the Association for Computational Linguistics.
- …

Outline

1. Introduction to AMR

- **Computational semantic, syntactic and semantic analysis**
- **Introduction to AMR**
- **AMR formats**
- **AMR concepts**
- **What AMR represents and do not represents**

Computational semantic & syntactic analysis

- **Computational semantics** (Blackburn and Bos, 2013) :

  *discipline that combines insights from **formal semantics**, **computational linguistics**, and **automated reasoning** whose goal is to **construct semantic representations** for expressions of human language in an **automatic way**.*

- The **meaning of a sentence** depends so closely on its **syntactic structure**
- **Syntactic analysis** and **syntactic parsers** play an important role in representing such meaning:
  - Syntaxtical phrase structure helps to identify the **semantic relationship** that a *predicate* has with its given *arguments* in the description of a situation – also known as **semantic roles**

But syntactic analysis is NOT able to represent MEANING

Limitation of syntactic analysis

- Given 4 sentences (from Matthews as cited in Chomsky, 1996):

  (1) The window broke

  (2) A hammer broke the window

  (3) The workman broke the window with a hammer

  (4) The window broke with a hammer

- Syntactically speaking, ***window*** is represented:
  - as the *verb's subject* in (1) and (4)
  - as the *verb's direct object* in the (2) and (3)
- All these sentences indicate that there is a broken thing: the window.
- However, **a syntactic analysis is not able to depict** this:

*Who did **what** to **whom**, **how**, **when**, **where**, **why**, and with **what consequences**?*

From syntactic analysis to semantic analysis

Syntactic analysis:

- **Completely assumed** by many statistical parsers trained on manually annotated syntactic database of sentences often in the form of a tree (Treebank)
- The **accuracy** of state-of-the-art syntactic parsers is **around 90%.**
- One of the most well-known English language treebanks is the Penn Treebank (PTB) (Marcus et al., 1993)

Semantic analysis:

- **Not currently assumed**
- Main reason: semantic annotation is « balkanized » (Banarescu et al., 2013), divided into separate annotations
- **Lack of a unified model** to integrate various kinds of annotation data
- Initiative towards a **graph-based parsing** for a more direct semantic analysis of whole natural language sentences: **AMR**

Introduction to AMR

- **AMR** stands for **Abstract Meaning Representation** is a **contribution to Semantic Analysis**
- **AMR** concept firstly introduced in 1998 (Langkilde & Knight, 1998)
- **AMR** is a ***semantic representation language*** based on the assumption that we lack a **simple readable semantic bank** – or **sembank** – of natural language sentences "paired with their whole-sentence, logical meanings" (Banarescu et al., 2013)
- **AMR** is a **graph-based annotation language, encoded as** Rooted Directed Acyclic Graph permitting **rapid human annotation of corpora** with broad coverage
- **AMR** have to permit **ultimate advances in NLP tasks** : Statistical Natural Language Understanding, Statistical Machine Translation, …
- **AMR deals with**: discourse connectives, semantic roles, intra-sentential coreference, named entities (wikification), questions, negation, & modality…

AMR formats (1)

- Very roughly AMR it is supposed to represent '**who** is doing **what** to **whom**, **where**, **when** and **how**' in a sentence S.
- AMR permits to abstract away meaning from syntactical representations, in the sense that sentences which are similar in meaning should be assigned the **same AMR**, even if they are not identically worded
- AMR have 3 equivalent formats:
  - **Logic format** : a formal representation
  - **AMR format**: a textual linearization based on Penman notation (Matthiessen et al., 1991), easy for human reading and writing
  - **Graph format**: for visualisation and used by programs

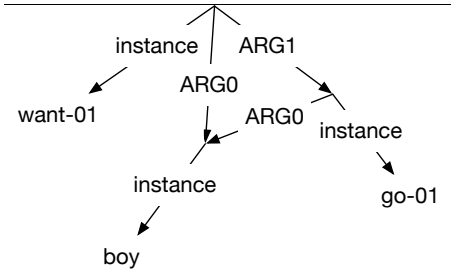
3 AMR formats (2)

Sentence : « *The boy wants to go* »

Logic formulation :

∃ w, b, g: instance(w, want-01) ∧ instance(g, go-01) ∧ instance(b, boy) ∧ arg0(w,b) ∧ arg1(w, g) ∧ arg0(g, b)

Graph formulation (DAG)



AMR formulation (based on PENMAN)

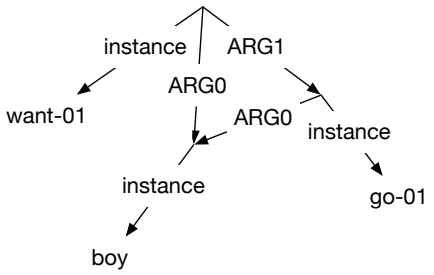
```
(w / want-01
   :arg0 (b / boy)
   :arg1 (g / go-01)
         arg0 b)
)
```


3 AMR Graph: alternative formulations

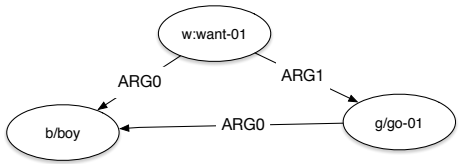
Sentence : « *The boy wants to go* »

∃ w, b, g: instance(w, want-01) ∧ instance(g, go-01) ∧ instance(b, boy) ∧ arg0(w,b) ∧ arg1(w, g) ∧ arg0(g, b)

Graph formulation 1



Graph formulation 1



AMR concepts (1)

Every AMR has a single **root** node at the top of the graph, which is considered to be the **focus**

Each **node** in the graph

- has a *variable* and represents a *semantic concept* (variable = instance of concept) a slash (/)
- variables are reused if something is referenced multiple times: *re-entrancy*
- *Semantic concepts* include PB (PropBank) framesets and English words

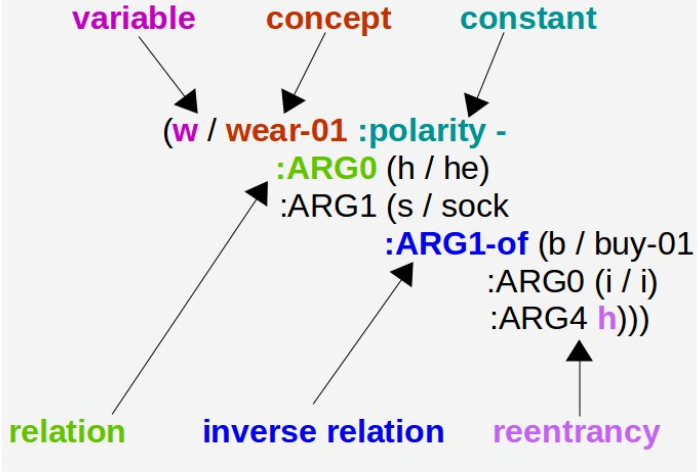
AMR concepts (2)

Graph **edges** denote *relations* between concepts

- *Semantic relations* include different types of **roles**, marked by a colon prefix (:)
- Some relations known as *constants* – get no variable, just a **value**
- Relation (role) can be *inverted* (useful for maintaining a single rooted structure)
- It is also possible to **convert** a *role* into a *concept* by reification (usefull to make a relation the focus of an AMR fragment)

AMR concepts (3)

Sentence: "*He is not wearing the socks that I bought him*"



What AMR represent …

To capture many aspects of meaning in a single simple data structure AMR:

- abstracts away from *morpho-syntactic idiosyncrasies*
- focus on *logic* rather than in *syntactic representation*

AMR uses ***PropBank*** *(PB) framesets* (Palmer et al., 2005) :

- each frame presents *annotators* with a list of *senses*
- each *sense* has its *own definition* and its *own numbered arguments* (ARG)

AMR uses of approximately a ***100 semantic relations*** (*semantic roles*) organised in role categories

AMR *does not dictate imperative modelling rules*, but promotes *personal interpretation* about how strings are related to meanings.

What AMR do not represent … *currently* …

In order to :

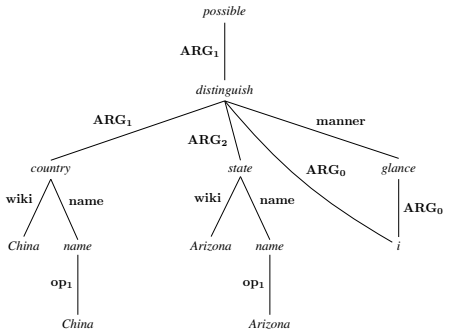
- obtain a *simple representation*
- to assign the *same AMR* to sentences that have the *same basic meaning*

AMR :

- does not represents *tense* and *number* of *verbs*
- do not represents *word category* and *order of words*
- drops in a sentence :
  - articles
  - most prepositions except for time and location prepositions
- AMR is NOT an interlingua, it is biased towards English.

AMRs examples … *(Le Petit Prince, Saint-Exupery)*

- ***At a glance I can distinguish China from Arizona***
  « D' un coup d'œil je peux distinguer la Chine de l'Arizona »
- ***At a glance I am able to distinguish China from Arizona***
  « D' un coup d'œil je suis capable de distinguer la Chine de l'Arizona »
- ***At a glance it is possible for me to distinguish China from Arizona***
  « D' un coup d'œil c'est possible pour moi de distinguer la Chine de l'Arizona »
  All have the Same AMR :



2. AMR Tools

- **AMR alignments**
- **AMR parsers**
- **AMR visualization & comparisons**
- **AMR text generators**

AMR tools (1)

AMR Alignments:

- AMR parser and generators are often trained using an AMR corpus, a large set of AMR graphs and corresponding reference sentences
- For such training procedures, it is useful to somehow link vertices of each AMR graph **G** to corresponding words of its reference sentence **s**.
- These links are commonly referred to as an « alignment » and several methods have been proposed for automatically generating such alignments.
- The aligner by Pourdamghani et al. (2014) is available at :
  - http://isi.edu/~damghani/papers/Aligner.zip
- The aligner by Flanigan et al. (2016) is available at :
  - github.com/jflanigan/jamr

AMR tools (2)

AMR parsers (from text to AMR format):

- **JAMR**: Flanigan J. et al. « A Discriminative Graph-Based Parser for the Abstract Meaning Representation » (ACL 2014)
  - https://github.com/jflanigan/jamr
- **CAMR**: Wang C. et al., « A Transition-based Algorithm for AMR Parsing » (NAACL 2015)
  - https://github.com/c-amr/camr
- **AMR-Eager**: Damonte M. et al. « An incremental parser for abstract meaning representation » (EACL 2017).
  - https://github.com/mdtux89/amr-eager

AMR tools (3)

AMR parsers (from text to AMR format) - cont.:

- **Neural-AMR**: Konstak et al., « Neural AMR: Sequence-to-Sequence Models for Parsing and Generation », ArXiv 2017.
  - https://github.com/sinantie/NeuralAmr
- **GREW**: Perrier G., Guillaume B., Bonfante G. (LORIA), « Application of Graph Rewriting to Natural Language Processing » (Wiley 2018) & « Application de la réécriture de graphes au traitement automatique des langues » (EMI-2018)
  - http://parse.grew.fr/
  - http://match.grew.fr/
- **…**

AMR tools (4)

AMR visualisation & comparisons:

- **AMRICA** : visualize an AMR or the difference between 2 AMRs to help users diagnose inter-annotator disagreement or errors from an AMR parser.
  - Saphra N., Lopez A., « AMRICA: an AMR Inspector for Cross-language Alignments » (ACL 2015).
  - http://github.com/nsaphra/AMRICA
- **Smatch** : an algorithm to compute a metric that calculates the degree of overlap between two semantic feature AMR structures.
  - Cai S., Knight K., « Smatch: an evaluation metric for semantic feature structures » (ACL 2013).
  - https://amr.isi.edu/eval/smatch/compare.html
- **…**

AMR tools (5)

AMR text generators (from AMR to text):

- **Flanigan generator:** Flanigan, J., Dyer, C., Smith, N. A., and Carbonell, J. « ***Generation*** from abstract meaning representation using tree transducers » (NAC-ACL 2016)
  - http://github.com/jflanigan/jamr/tree/Generator.
- **Song generator:** Song, L., Zhang, Y., Peng, X., Wang, Z., Gildea, D. (2016). AMR-to-text ***generation*** as a traveling salesman problem » (EMNLP 2016).
  - http://github.com/xiaochang13/AMR-generation
- **Neural-AMR**: Konstak et al., « Neural AMR: Sequence-to-Sequence Models for Parsing and ***Generation*** », ArXiv 2017.
  - https://github.com/sinantie/NeuralAmr
- **…**

AMR and neural Networks and Deep Learning …

- Neural-AMR: Konstak et al. (2017). **Neural AMR: Sequence-to-Sequence Models for Parsing and *Generation*** ».
  - https://github.com/sinantie/NeuralAmr
- Rik van Noord and Johan Bos (2018). **Dealing with Co-reference in Neural Semantic Parsing**
- Rik van Noord and Johan Bos (2017). **Neural Semantic Parsing by Character-based Translation: Experiments with Abstract Meaning Representations**
- Xiaochang Peng, Chuan Wang, Daniel Gildea and Nianwen Xue (2017). **Addressing the Data Sparsity Issue in Neural AMR Parsing**
- Hardy and Andreas Vlachos (2018). **Guided Neural Language Generation for Abstractive Summarization using Abstract Meaning Representation**
- William R. Foland Jr. & James H. Martin (2017). **Abstract Meaning Representation Parsing using LSTM Recurrent Neural Networks**
- Jan Buys & Phil Blunsom (2017). **Robust Incremental Neural Semantic Graph Parsing**
- …

3. AMR applications

- **SemEval 2016 & SemEval 2017**
- **Automatic Abstractive Summarization**
- **Events extraction**
- **Various applications**

SemEVAL 2016

Objectives (http://alt.qcri.org/semeval2016/task8/):

- Participants where be provided with parallel English-AMR training data.
- They will had to parse new English data and return the obtained AMRs.
- Participants could use any resources at their disposal (but may not hand-annotate the blind data or hire other human beings to hand-annotate the blind data).
- The SemEval trophy goes to the system with the highest Smatch score.

SemEVAL 2017

Objectives (http://alt.qcri.org/semeval2017/task9/):

- **Subtask 1- Parsing Biomedical Data**: participants were asked to produce Abstract Meaning Representation (AMR) (Banarescu et al., 2013) graphs for a set of English sentences in the biomedical domain.
- **Subtask 2 - AMR-to-English Generation**: participants were asked to generate English sentences given AMR graphs in the news/forum domain.

A total of 5 sites participated in the parsing subtask, and 4 participated in the generation subtask.

AMR Applications (1)

Automatic Abstractive Summarization:

- Liu, F., Flanigan, J., Thomson, S., Sadeh, N., and Smith, N. A. (2015). **Toward abstractive summarization using semantic representations**. ACL 2015.
- Liao K., Lebanoff L., Liu F. (2018). **Abstract Meaning Representation for Multi-Document Summarization**
- Dohare S. , Gupta V., Karnick H. (2018). **Unsupervised Semantic Abstractive Summarization**, ACL 2018 .
- Dohare S. , Gupta V., Karnick H. (2017). **Text Summarization using Abstract Meaning Representation**.
- Moreda P., Suarez A., Lloret E., Saquete E., Moreno I. (2018) **From Sentences to Documents: Extending Abstract Meaning Representation for Understanding Documents**
- …

AMR Applications (2)

Event extraction:

- Rao S., Marcu D., Knight K., Daume H., (2017). **Biomedical Event Extraction using Abstract Meaning Representation**
- Cao K., **Improving Event Extraction-Casting a Winder Net** (2017) PhD thesis
- Lifu Huang L., Taylor Cassidy T., Xiaocheng Feng X., Heng Ji H., Clare R. Voss C., Jiawei Han J., Avirup Sil A. (2016). **Liberal Event Extraction and Event Schema Induction**.
- Peng H., Song Y. and Roth D. (2016). **Event Detection and Co-reference with Minimal Supervision** (EMNLP'16)
- Li X., Nguyen T., Cao K., Grishman R. (2015) **Improving Event Detection with Abstract Meaning Representation**-ACL 2015
- …

AMR Applications (3)

Various applications:

- Bonial C., Donatelli L., Ervin J., Voss C.R. (2019). **Abstract Meaning Representation for Human-Robot Dialogue**
- Issa F., Damonte M., B. Cohen S.B., Yan X., Chang Y. (2018). **Abstract Meaning Representation for Paraphrase Detection** (NAACL 2018)
- Shen H., (2018) **Semantic Parsing in Spoken Language Understanding using Abstract Meaning Representation**, Master thesis.
- Wang Y. et al. (2017). **Dependency and AMR Embeddings for Drug-Drug Interaction Extraction from Biomedical Literature** ACM- ACM-BCB'17
- Lai Dac Viet L.D., Vu Trong Sinh V.T., Nguyen Le Minh N., Ken Satoh K. (2017). **ConvAMR: Abstract meaning representation parsing for legal document**
- …

4. AMR and different natural languages

- ..

AMR and different natural languages

AMR is originally developed for English language, but it exists various tentatives to use it with different natural languages :

- Noelia Migueles-Abraira, Rodrigo Agerri, Arantza Diaz de Ilarraza (2018). **Annotating Abstract Meaning Representations for Spanish**
- Rafael Torres Anchieta, Thiago Alexandre Salgueiro Pardo (2018). **Towards AMR-BR: A SemBank for Brazilian Portuguese Language**
- Marco Damonte Shay B. Cohen (2018). **Cross-lingual Abstract Meaning Representation Parsing** (NAACL 2018)
- Noelia Migueles-Abraira (2018) **A Study Towards Spanish Abstract Meaning Representation**, Master Thesis.
- Elior Sulem, Omri Abend, Ari Rappoport (2015). **Conceptual Annotations Preserve Structure Across Translations: A French-English Case Study**
- Lucy Vanderwende, Arul Menezes, Chris Quirk (2015). **An AMR parser for English, French, German, Spanish and Japanese and a new AMR-annotated corpus** (Microsoft Research)
- Boliang Zhang, Ying Lin, Xiaoman Pan, Di Lu, Jonathan May, Kevin Knight, Heng Ji (2018). **ELISA-EDL: A Cross-lingual Entity Extraction, Linking and Localization System**