AMR-to-Text Generation **via GF**

Normunds Grūzītis

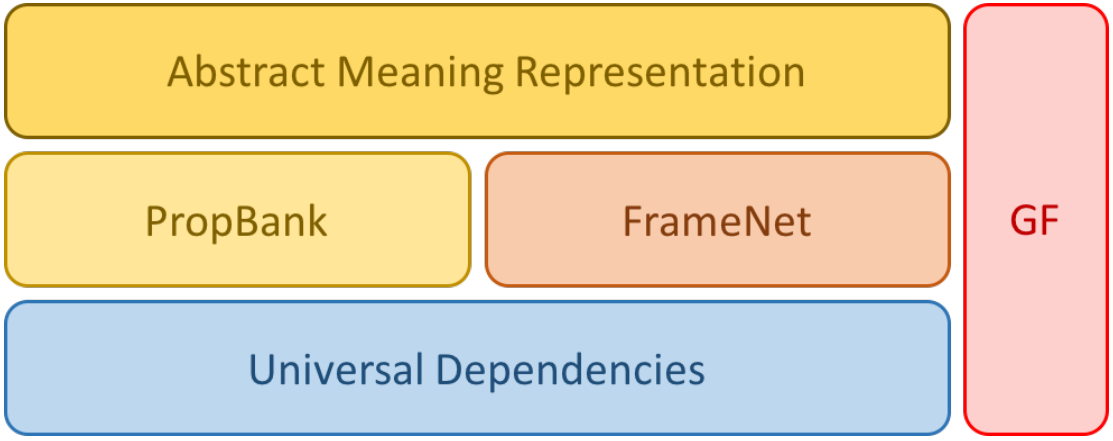
University of Latvia, Institute of Mathematics and Computer Science
National information agency LETA

GF Summer School 2017, Rīga, Latvia

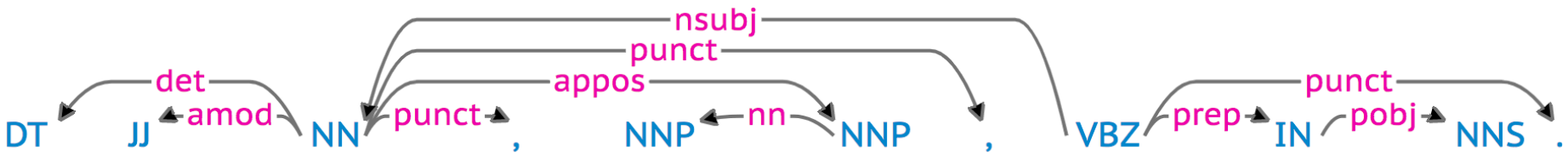
This work has received funding in part from the Latvian State research programs SOPHIS and NexIT, the EU Horizon 2020 project SUMMA (grant No. 688139), and the European Regional Development Fund (grant No. 1.1.1.1/16/A/219).

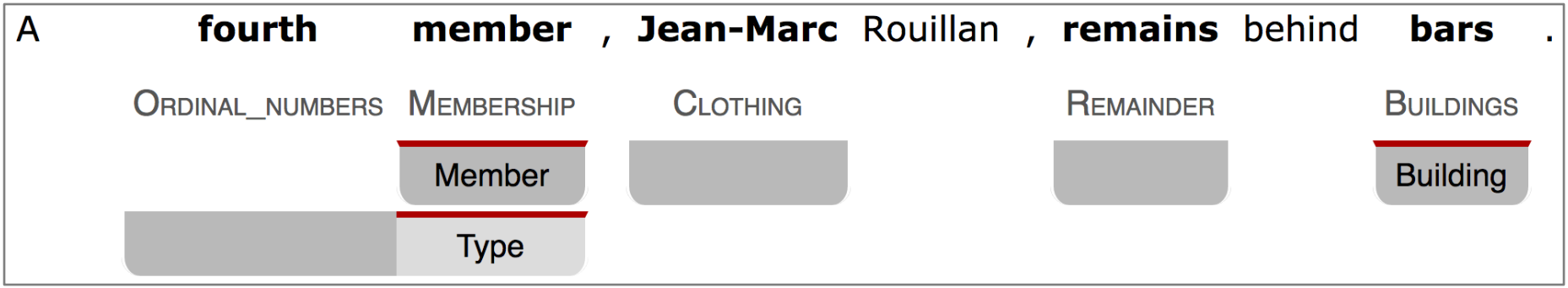
Agenda

- **Frame semantics**
  - FrameNet
  - PropBank
- **AMR**
- Text-to-AMR parsing, **AMR-to-text generation**
  - SemEval 2016
  - SemEval 2017



Semantic Role Labelling (SRL)





FrameNet

TurboParser + SEMAFOR: http://demo.ark.cs.cmu.edu/parse

	A	fourth	member	,	Jean-Marc	Rouillan	,	remains	behind	bars	.
member.01											
remain.01	A1								A3		

PropBank

LTH parser: http://barbar.cs.lth.se:8081/

FrameNet (https://framenet.icsi.berkeley.edu)

Remainder

Definition:

A Resource is depleted by some process, resulting in there being only a Remainder in existence some time into or after the process. The original Resource may be indicated metonymically via reference to an Original_owner. The Location where the Resource is found may be expressed.

There are only two poems LEFT from Quayle

After the storm, only two houses were LEFT DNI

Not much was LEFT in the till

Nothing REMAINED for the youngest brother from the inheritance

Only this school is LEFT from the 1970s; all the other ones are newer.

Lexical Unit	LU Status	Lexical Entry Report	Annotation Report	Annotator ID	Created Date
left.a	**Created**	Lexical entry	Annotation	JKR	03/16/2004 12:23:35 PST Tue
remain.v	**Finished_Initial**	Lexical entry	Annotation	JKR	03/16/2004 12:24:22 PST Tue
remains.n	**Created**	Lexical entry	Annotation	KmG	03/14/2007 03:20:12 PDT Wed
remnant.n	**Finished_Initial**	Lexical entry	Annotation	JKR	03/16/2004 12:21:10 PST Tue

FrameNet (https://framenet.icsi.berkeley.edu)

State_continue

Definition:

Despite some implication that a State would be interrupted, the Entity remains in the specified State. Note that State includes locative relations.

Online security remains elusive.

FEs:

Core:

Entity [ent]	A concrete or abstract Entity.

Core Unexpressed:

State [sta]	The State of an Entity.

remain.v	**Finished_Initial**	Lexical entry	Annotation	JKR	06/30/2004 10:30:48 PDT Wed
rest.v	**Created**	Lexical entry	Annotation	RLG	08/21/2006 02:26:58 PDT Mon
stay.v	**Created**	Lexical entry	Annotation	MJE	08/19/2004 05:38:49 PDT Thu

FrameNet (FN)

- A lexico-semantic resource based on the theory of frame semantics (Fillmore et al. 2003)
  - A semantic **frame** represents a cognitive, prototypical situation (scenario) characterized by **frame elements** (**FE**) – semantic valence
  - Frames are "evoked" in sentences by **target words** – **lexical units** (**LU**)
  - FEs are mapped based on the syntactic valence of the LU
    - The syntactic valence patterns are derived from **FN-annotated corpora** (for an increasing number of languages, incl. Latvian)
  - FEs are split into **core** and **non-core** ones
    - Core FEs uniquely characterize the frame and syntactically **tend** to correspond to verb arguments
    - Non-core FEs are not specific to the frame and **typically** are adjuncts

Berkeley FrameNet as Interlingua

Desiring	
Definition:	An EXPERIENCER desires that an EVENT occur. In some cases, the EXPERIENCER is an active participant in the EVENT, and in such cases the EVENT itself is often not mentioned, but rather some FOCAL_PARTICIPANT which is subordinately involved.
Core FEs:	EVENT, EXPERIENCER, FOCAL_PARTICIPANT, LOCATION_OF_EVENT
Non-core FEs:	CAUSE, DEGREE, DURATION, MANNER, PLACE, PURPOSE_OF_EVENT, REASON, ROLE_OF_FOCAL_PARTICIPANT, TIME, TIME_OF_EVENT

Introduced in **BFN**, reused in SweFN

want.v..6412

Examples	Valence patterns	
40	EVENT	EXPERIENCER
(22)	VPto.Dep	NP.Ext
14	EXPERIENCER	FOCAL_PARTICIPANT
(10)	NP.Ext	NP.Obj
(1)	PP[*by*].Dep	NP.Ext

Some valence patterns found in **BFN**

e.g. "[*I*]$_{\text{Experiencer}}$ *do n't WANT* [*to deceive anyone*]$_{\text{Event}}$"

|

an embedded frame

känna_för.vb..1

Examples	Valence patterns	
1	EVENT	EXPERIENCER
(1)	VB.INF.VG	NN.SS
2	EXPERIENCER	FOCAL_PARTICIPANT
(2)	NN.SS	NN.OO

Some valence patterns found in **SweFN**

e.g. "[*Jag*]$_{\text{Experiencer}}$ *KÄNNER FÖR* [*en tur på landet*]$_{\text{Focal\_participant}}$"

FrameNet and GF

- Existing FNs are not entirely formal and computational
  - A limited but computational FN-**based** GF grammar and lexicon
- Grammatical Framework:
  - Separates between an **abstract syntax** and **concrete syntaxes**
  - Provides a general-purpose **resource grammar library** (RGL)
- The language-independent layer of FrameNet (frames and FEs) – the abstract syntax
  - The language-specific layers (surface realization of frames and FEs; LUs) – concrete syntaxes
- RGL can be used for unifying the syntactic types used in different FNs and for the concrete implementation of frames
  - FrameNet allows for abstracting over RGL

Use case (1)

- Provide a semantic API on top of RGL to facilitate the development of GF application grammars
  - In combination with the syntactic API of RGL
  - Hiding the comparatively complex construction of verb phrases

```
mkCl person (mkVP (mkVP live_V) (mkAdv in_Prep place))
    -- mkCl : NP -> VP -> Cl
    -- mkVP : V -> VP
    -- mkVP : VP -> Adv -> VP
    -- mkAdv : Prep -> NP -> Adv
```

```
Residence                       -- Residence : NP -> Adv -> V -> Cl
    person                      -- NP (Resident)
    (mkAdv in_Prep place)       -- Adv (Location)
    live_V_Residence            -- V (LU)
```

Use case (2)

- FN-annotated knowledge bases → multilingual verbalization

		Time	Place	Relatives	**Child**	
Being_born	dzimt.v	1933. gada 3. maijs	Slokas pagasts	zvejnieka ģimene	Imants Ziedonis	
		Institution	**Subject**	Time	Place	**Student**
Education_teaching	absolvēt.v	Tukuma 1. vidusskola		1952. gads	Tukums	Imants Ziedonis
Education_teaching	beigt.v	Latvijas Universitāte	vēsture un filoloģija	1959. gads		Imants Ziedonis
Education_teaching	beigt.v	Augstākais literārais [..]		1964. gads	Maskava	Imants Ziedonis
		Employer	**Place_of_employment**	**Position**	Time	**Employee**
Being_employed	redaktors.n	izdevniecība Liesma		> redaktors		Imants Ziedonis
Being_employed	sekretārs.n	Latvijas rakstnieku [..]		> sekretārs		Imants Ziedonis
Being_employed	loceklis.n	AP tautas izglītības [..]		> loceklis		Imants Ziedonis
Being_employed	loceklis.n	Latvijas Institūts		> loceklis	1998. gads	Imants Ziedonis
Being_employed	padomnieks.n			> padomnieks	1997. gads	Imants Ziedonis
Being_employed	skolotājs.n	Jūrmalas 1. vidusskola		> skolotājs		Imants Ziedonis
		Time	**Prize**	Rank	Competition	**Competitor**
Win_prize	apbalvot.v	1983. gads	Tautu draudzības [..]			Imants Ziedonis
Win_prize	piešķirt.v	1972. gads	Nopelniem bagātais [..]			Imants Ziedonis
Win_prize	piešķirt.v	1977. gads	Tautas dzejnieka goda [..]			Imants Ziedonis
Win_prize	saņemt.v		1991. gada barikāžu [..]			Imants Ziedonis

Imants Ziedonis ir dzimis 1933. gada 3. maijā Slokas pagastā.
Imants Ziedonis was born in Sloka parish on 3 May 1933.

FrameNet-based grammar: abstract

- **Frame** valence patterns are represented by functions
  - Taking one or more core FEs (A-Z) and one LU as arguments
  - Returning an object of type *Clause* whose linearization type is `{np: NP; vp: VP}`

```
fun Desiring_V       : Experiencer_NP -> Focal_participant_Adv -> V -> Clause
fun Desiring_V2      : Experiencer_NP -> Focal_participant_NP -> V2 -> Clause
fun Desiring_V2_Pass : Experiencer_NP -> Focal_participant_NP -> V2 -> Clause
fun Desiring_VV      : Event_VP -> Experiencer_NP -> VV -> Clause
```

- **FE**s are declared as semantic categories subcategorized by the syntactic RGL types
  - *NP*, *VP*, *Adv* (includes prepositional objects), *S* (embedded sentences), *QS*

```
cat Event_VP
cat Experiencer_NP
cat Focal_participant_NP
cat Focal_participant_Adv
```

FrameNet-based grammar: concrete

- The mapping from the semantic FrameNet types to the syntactic RGL **types** is shared for all languages
  - Linearization types are of type *Maybe* to allow for optional (empty) FEs

```
lincat Focal_participant_NP  = Maybe NP
lincat Focal_participant_Adv = Maybe Adv
```

- To implement the frame functions, RGL **constructors** are applied to the arguments depending on their types and syntactic roles, and the voice

```
lin Desiring_V2 experiencer focal_participant v2 = {
  np = fromMaybe NP experiencer ;
  vp = mkVP v2 (fromMaybe NP focal_participant)
}

lin Desiring_V2_Pass experiencer focal_participant v2 = {
  np = fromMaybe NP focal_participant ;
  vp = mkVP (passiveVP v2) (mkAdv by8agent_Prep (fromMaybe NP experiencer))
}
```

FrameNet-based API to GF Resource Grammar Library

A tool for cross-lingual comparison of FrameNet-annotated corpora

Frames | Verbs

vers. 0.9.7

REMU

Court_examination (1)
Create_physical_artwork (1)
Create_representation (2)
Creating (2)
Criminal_investigation (1)
Cutting (2)
Damaging (2)
Daring (2)
Death (1)
Deciding (1)
Delimitation_of_diversity (1)
Delivery (3)
Deny_permission (1)
Departing (1)
Deserving (1)
Desiring (3)
Destroying (4)
Detaching (2)
Detaining (2)
Differentiation (3)
Discussion (1)
Dispersal (1)
Distinctiveness (1)
Dodging (1)
Dominate_competitor (1)
Dominate_situation (1)
Dressing (2)
Drop_in_on (1)
Dunking (1)
Duplication (2)
Duration_relation (2)

Desiring

Desiring_V : Experiencer_NP → Focal_participant_Adv → V → Clause

▶ **Eng**: [They]$_{Experiencer}$ [ASPIRED] [towards the Chelsea shore , where , in the early 1960s many thousands lived with sensible occupations and adequate amounts of money]$_{Focal\_participant}$

▶ **Swe**: [Roberte]$_{Experiencer}$ [LÄNGTADE] [hem till Tyskland]$_{Focal\_participant}$

Desiring_V2 : Experiencer_NP → Focal_participant_NP → V2 → Clause

▼ **Eng**:

▶ **covet**_V2_Desiring : V2

▶ **crave**_V2_Desiring : V2

▼ **desire**_V2_Desiring : V2

[I]$_{Experiencer}$ neither [DESIRE] [this house]$_{Focal\_participant}$

▶ **fancy**_V2_Desiring : V2

▶ **feel_like**_V2_Desiring : V2

▶ **want**_V2_Desiring : V2

▶ **yearn**_V2_Desiring : V2

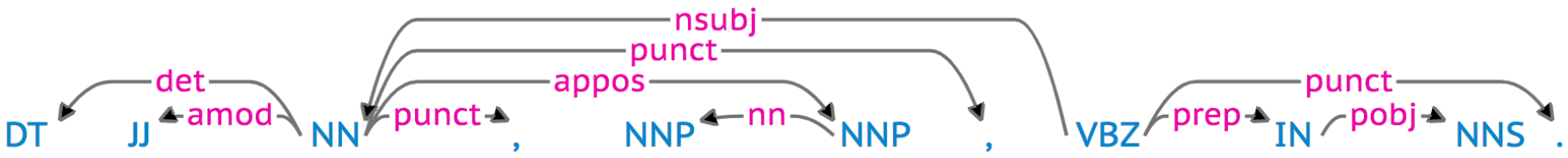
▶ **Swe**: [Jag]$_{Experiencer}$ [KÄNNER FÖR] [en tur på landet]$_{Focal\_participant}$

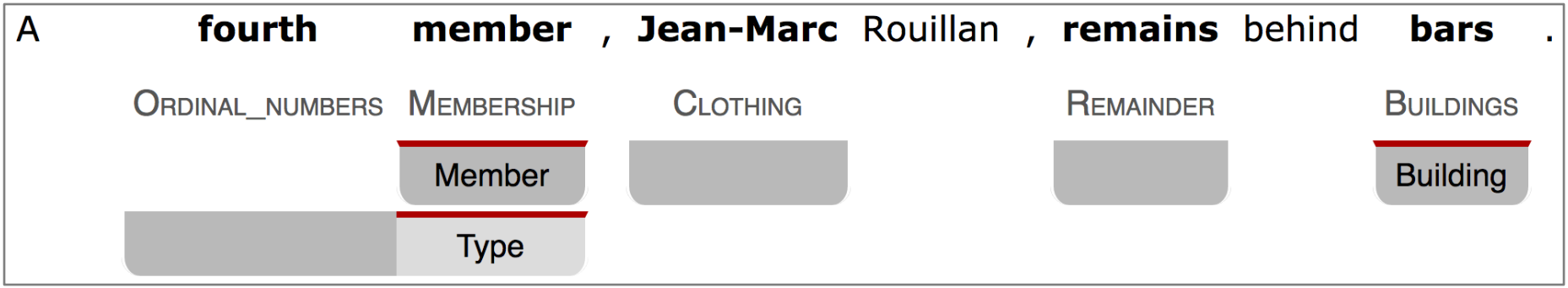
Desiring_VV : Event_VP → Experiencer_NP → VV → Clause

▶ **Eng**: [He]$_{Experiencer}$ ground his teeth together , [LUSTING] [to tear the alien apart and eat of its lurid vitals , so as to comprehend something of its strange nature]$_{Event}$

http://grammaticalframework.org/framenet/

Semantic Role Labelling (SRL)





FrameNet

TurboParser + SEMAFOR: http://demo.ark.cs.cmu.edu/parse

	A	fourth	member	,	Jean-Marc	Rouillan	,	remains	behind	bars	.
member.01											
remain.01	A1								A3		

PropBank

LTH parser: http://barbar.cs.lth.se:8081/

PropBank (http://propbank.github.io)

remain-v; 3 senses

Sense Number 1: remain in a certain state

Examples:
The President is expected to remain firm on his veto of the bill.
In this game you must remain down to the count of ten.

Mappings:
FrameNet: State_continue
PropBank: remain.01
WordNet 3.0 Sense Numbers: 1
WordNet Verb Particle Constructions, Multiword Expressions:
remain_down 1
remain_firm 1

Sense Number 2: be left; stay behind; continue in a place, position, or situation

Examples:
She remained with the child until he fell asleep.
If you subtract ten from twelve, two remain.

Mappings:
FrameNet: Existence,Remainder
PropBank: remain.01
WordNet 3.0 Sense Numbers: 2, 3, 4

Predicate: *remain*

Roleset id: remain.01 , *be left behind* , Source: vncls: 47.1-1 , framnet: Remainder , Left_to_do , State_continue , Existence

Roles:

Arg1-PPT: *Thing remaining* (vnrole: 47.1-1-Theme)
Arg2-GOL: *benefactive, entity who gets the remainder*
Arg3-PRD: *attribute of arg1*

Example: All alone

Only one local ringer remains.

Arg1: Only one local ringer
Rel: remains

Example: With PP

Exports will remain under a government target.

Arg1: Experts
Rel: remain
Arg3: under a government target

Example: remain on

Five countries will remain on the so-called priority watch list

Arg1: Five countries
Rel: remain
Arg3: on the so-called priority watch list

Example: With AdjP

Their influence will remain subordinate to Japan's.

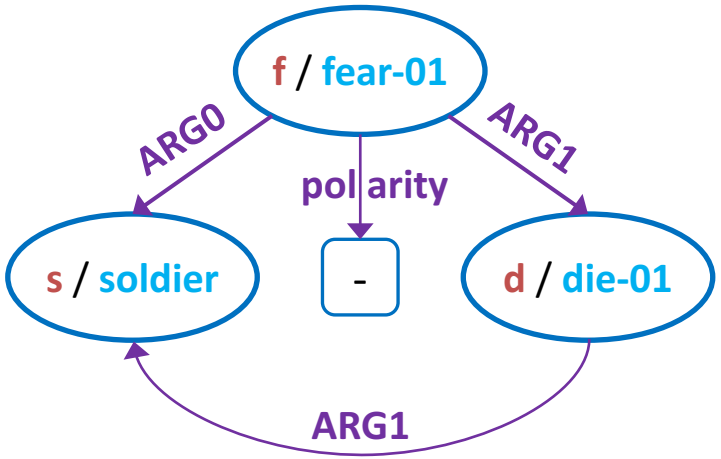
Arg1: Their influence
Rel: remain
Arg3: subordinate to Japan's

AMR (Abstract Meaning Representation)

- From SRL to **whole-sentence** meaning representation
  - Incl. **PropBank** SRL, NER and NEL, treatment of modality, negation, etc.
- Simple and **compact** data structure
  - PENMAN notation: directed labeled graph encoded in a tree-like form
  - Easy to read and write (for a human), and traverse (for a program)
  - Langkilde and Knight (1998) → Banarescu et al. (2013)
- Aimed at **large-scale** human annotation and semantic parsing
  - **Practical**, replicable amount of abstraction
  - An actual **sembank** of 40K+ sentences
- Captures many aspects of meaning
  - Aims to **abstract** away from (English) syntax

AMR (Abstract Meaning Representation)

- Nodes are **variables** labelled by **concepts**
  - Entities, events, states, properties
  - `s` / `soldier`: `s` is an instance of `soldier`
- Edges are semantic **relations**
- AMR abstracts in numerous ways by assigning the same conceptual structure to different surface realizations



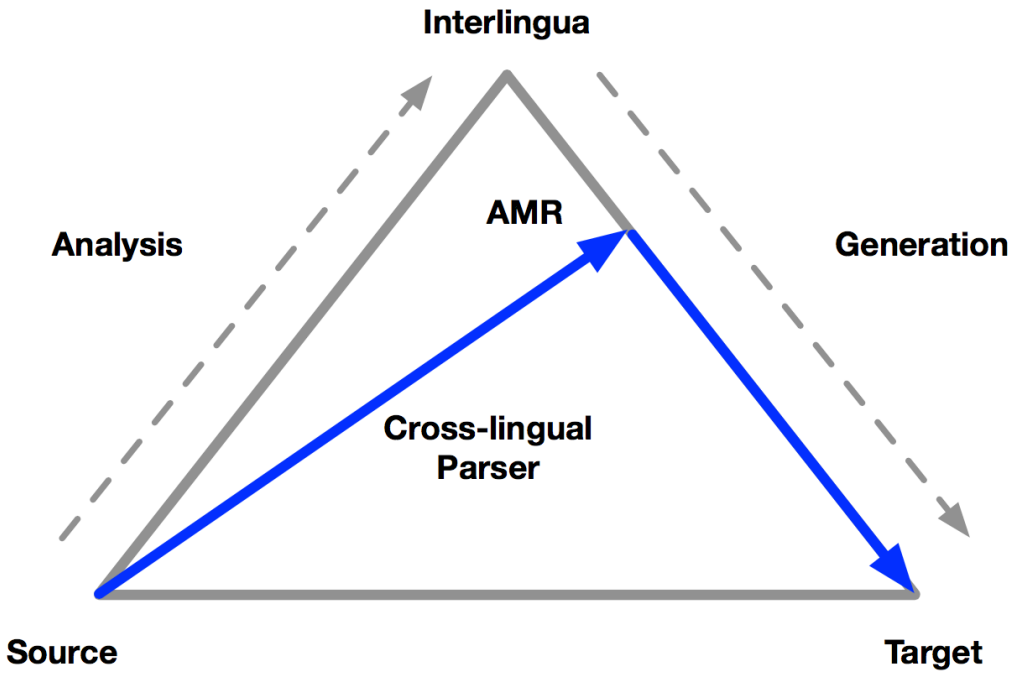

```
(f / fear-01
  :polarity "-"
  :ARG0 ( s / soldier )
  :ARG1 ( d / die-01
          :ARG1 s ))
```

The soldier was not afraid of dying.
The soldier was not afraid to die.
The soldier did not fear death.

(Pust et al., 2015)

AMR (Abstract Meaning Representation)

- AMR is still **biased** towards English or other source languages
- Meanwhile, AMR is **agnostic** about how to derive meanings from strings, and vice versa
- Xue N., Bojar O., Hajič J., Palmer M., Uresova Z., Zhang X. ***Not an Interlingua, but close****: Comparison of English AMRs to Chinese and Czech.* LREC 2014



Schneider N., Flanigan J., O'Gorman T.
AMR Tutorial at NAACL 2015
https://github.com/nschneid/amr-tutorial/

Text-to-AMR: human annotation

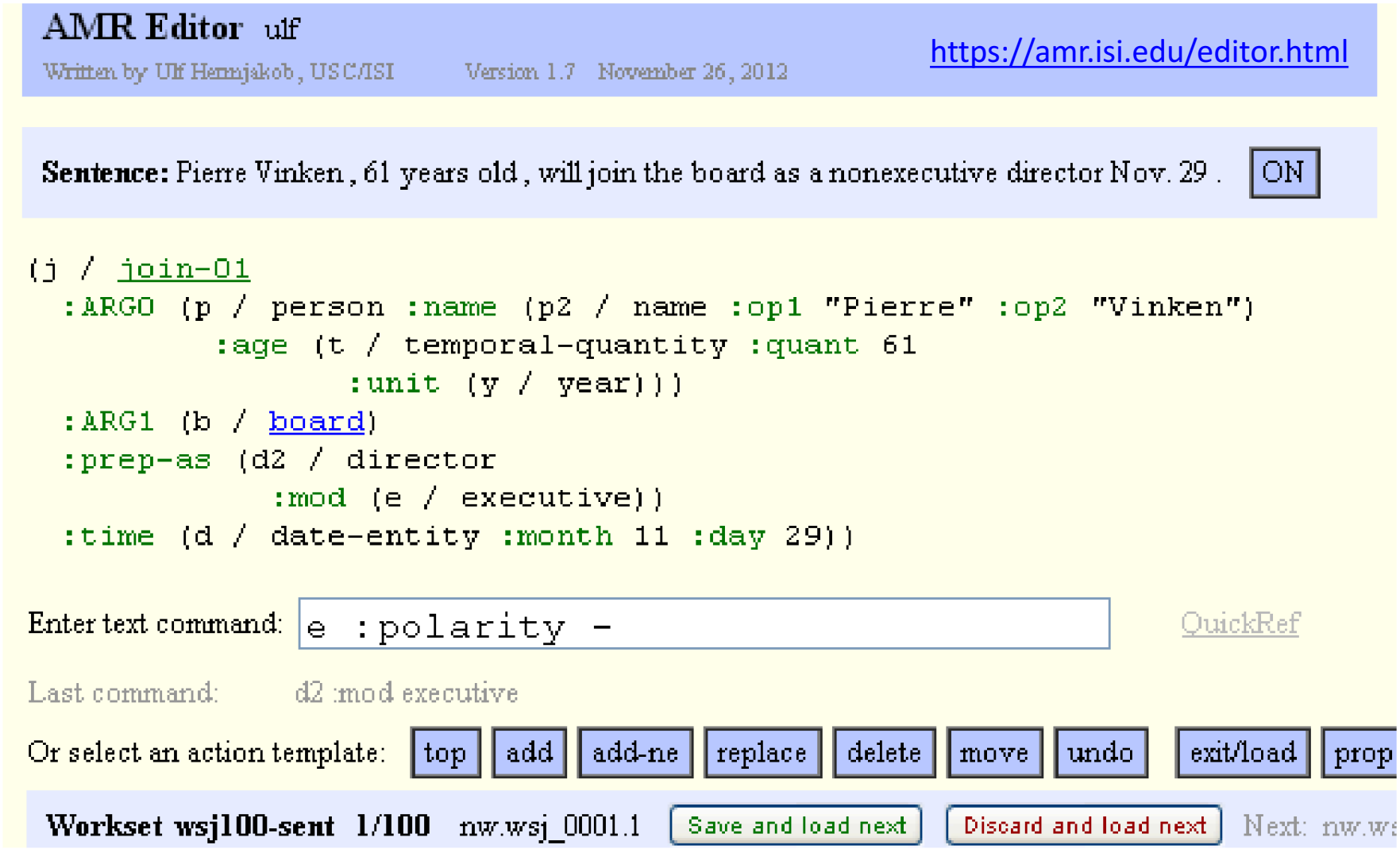
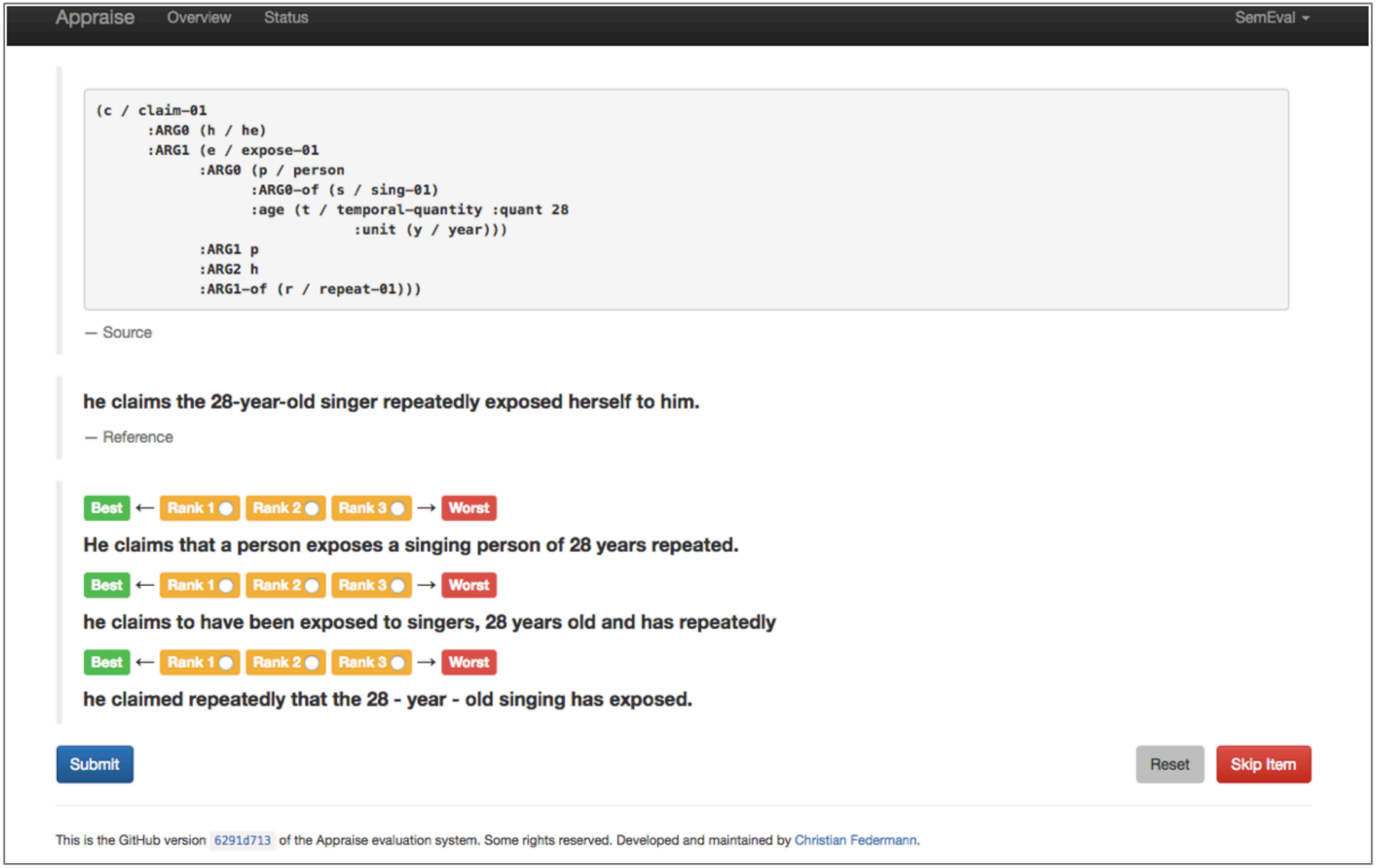


Figure 1: Screenshot of the AMR Editor when entering a text command, showing the core portion of the main window.

AMR-to-text: human evaluation



Sample AMR (1)

::snt *A fourth member, Jean-Marc Rouillan, remains behind bars.*

```
(r / remain-01
    :ARG1 (p / person
        :wiki -
        :name (n / name
            :op1 "Jean-Marc" :op2 "Rouillan")
        :mod (p2 / person
            :ARG0-of (h / have-org-role-91
                :ARG2 (m / member))
            :ord (o / ordinal-entity
                :value 4)))
    :ARG3 (b / behind
        :op1 (b2 / bar)))
```

Sample AMR (1)

::snt *A fourth member, Jean-Marc Rouillan, remains behind bars.*

```
(r / remain-01
    :ARG1 (p / person
        :wiki -
        :name (n / name
            :op1 "Jean-Marc" :op2 "Rouillan")
        :mod (p2 / person
            :ARG0-of (h / have-org-role-91
                :ARG2 (m / member))
            :ord (o / ordinal-entity
                :value 4)))
    :ARG3 (b / behind
        :op1 (b2 / bar)))
```

Remaining members person 4 jean-marc rouillan – behind bar.

Jean-Marc Rouillan, that is the 4th member, is remained behind a bar.

Sample AMR (1)

::snt *A fourth member, Jean-Marc Rouillan, remains behind bars.*

```
(r / remain-01
    :ARG1 (p / person
        :wiki -
        :name (n / name
            :op1 "Jean-Marc" :op2 "Rouillan")
        :mod (p2 / person
            :ARG0-of (h / have-org-role-91
                :ARG2 (m / member))
            :ord (o / ordinal-entity
                :value 4)))
    :ARG3 (b / behind
        :op1 (b2 / bar)))
```

Remaining members person 4 jean-marc rouillan – behind bar. JAMR

Jean-Marc Rouillan, that is the 4th member, is remained behind a bar. GF

Sample AMR (1)

::snt *A fourth member, Jean-Marc Rouillan, remains behind bars.*

```
(r / remain-01
    :ARG1 (p / person
        :wiki -
        :name (n / name
            :op1 "Jean-Marc" :op2 "Rouillan")
        :mod (p2 / person
            :ARG0-of (h / have-org-role-91
                :ARG2 (m / member))
            :ord (o / ordinal-entity
                :value 4)))
    :ARG3 (b / behind
        :op1 (b2 / bar)))
```

Remaining members person 4 jean-marc rouillan – behind bar. JAMR

Jean-Marc Rouillan, that is the 4th member, is remained behind a bar. GF

Sample AMR (2)

::snt *They should have been expelled from school at a minimum.*

```
(r / recommend-01
    :ARG1 (e / expel-01
        :ARG1 (t / they)
        :ARG2 (s / school)
        :degree (a / at-a-minimum)))
```

Should they at-a-minimum expel school. JAMR

It is recommended that they are expelled to a school at a minimum. GF

expel-01

ARG0=PAG (prototypical agent)
ARG1=PPT (prototypical patient)
ARG2=DIR (direction) → DIR_Prep → to_Prep

ToDo: based on **statistics** from PropBank and FrameNet **corpora**,
"reconstruct" Prep-s, depending on frame/verb valency, ARG role, or NP head

Sample AMR (3)

::snt *Texas criminal courts and prosecutors do not coddle to anyone.*

```
(c / coddle-01
    :polarity -
    :ARG0 (a / and
        :op1 (c2 / court
            :ARG0-of (c4 / criminal-03)
            :location (s / state
                :wiki "Texas"
                :name (n / name :op1 "Texas")))
        :op2 (p / person
            :ARG0-of (p2 / prosecute-01)
            :location s))
    :ARG1 (a2 / anyone))
```

No texas texas criminal court and prosecutors coddle anyone. JAMR

A criminal court in Texas and a person that prosecutes do not coddle anyone. GF

person that prosecutes → prosecutor
organization that governs → government

Sample AMR (4)

::snt *How Long are We Going to Tolerate Japan?*

```
(t / tolerate-01
     :ARG0 (w / we)
     :ARG1 (c / country
          :wiki "Japan"
          :name (n / name :op1 "Japan"))
     :duration (a / amr-unknown))
```

We have tolerated the japan amr-unknown. JAMR

How long do we tolerate Japan? GF

```
if ':mode expressive'           in amr: amr =           amr.replace(':mode expressive',           ' ') + ' !'
if ':mode imperative'           in amr: amr =           amr.replace(':mode imperative',           ' ') + ' !'
if ':mode interrogative'        in amr: amr =           amr.replace(':mode interrogative',        ' ') + ' ?'
if 'cause-01:ARG0(amr-unknown)' in amr: amr = 'why '   + amr.replace('cause-01:ARG0(amr-unknown)', ' ') + ' ?'
if ':location(amr-unknown)'     in amr: amr = 'where ' + amr.replace(':location(amr-unknown)',     ' ') + ' ?'
if ':ARG1(amr-unknown)'         in amr: amr = 'who '   + amr.replace(':ARG1(amr-unknown)',         ' ') + ' ?'
if ':mod(amr-unknown)'          in amr: amr = 'what '  + amr.replace(':mod(amr-unknown)',          ' ') + ' ?'
if ':duration(amr-unknown)'     in amr: amr = 'how '   + amr.replace(':duration(amr-unknown)',     ' ') + ' ?'
if 'amr-unknown'                in amr: amr = 'what '  + amr.replace('amr-unknown',               ' ') + ' ?'
```

Sample AMR (5)

::snt *Xinhua News Agency, Tokyo, September 1st, by reporter Yiguo Yu*

```
(b / byline-91
    :ARG0 (p2 / publication
        :name (n / name
            :op1 "Xinhua" :op2 "News" :op3 "Agency"))
    :ARG1 (p / person
        :name (n2 / name :op1 "Yiguo" :op2 "Yu")
        :ARG0-of (r / report-01))
    :location (c2 / city
        :name (n3 / name :op1 "Tokyo"))
    :time (d / date-entity
        :month 9
        :day 1))
```

Xinhua news agency has reported yiguo yu byline in a tokyo 1 9. JAMR

Xinhua News Agency by Yiguo Yu on 1 September in Tokyo. GF

Sample AMR (6)

::snt *Alliot-Marie arrived on Sunday.*

```
(a / arrive-01
    :ARG1 (p / person :name (n / name :op1 "Alliot-Marie"))
    :time (d / date-entity :weekday (s / sunday)))
```

Sunday 's arrival of alliot-marie michèle_alliot-marie. JAMR

```
unknown qualified constant L.arrive_V2
```
GF

```
(a / arrive-01
    :ARG0 (p / person :name (n / name :op1 "Alliot-Marie"))
    :time (d / date-entity :weekday (s / sunday)))
```

Alliot-marie michèle_alliot-marie arrived sunday. JAMR

Alliot-Marie arrives on Sunday. GF

SemEval 2017: Task 9

- Subtask 1: Parsing Biomedical Data
- Subtask 2: AMR-to-English Generation

UL / IMCS / LETA
JAMR (5-grams)
Tranducer → lin
SMT: AMR → Eng
AMR→(U)D→lin

SemEval

ACL

Approaches:
- "grammar-based"
- SMT/NMT
- end-to-end

	Win	Win+Tie	Trueskill	BLEU
RIGOTRIO	**54.91**	**81.49**	**1.07**	18.82
CMU	50.36	72.48	0.85	**19.01**
FORGe	43.64	57.43	0.45	4.74
ISI	26.05	38.39	-1.19	10.92
Sheffield	8.38	21.16	-2.20	3.32

Table 3: Main generation results: The three manually-derived metrics agree on the systems' relative rankings.

	Win	Win+Tie	Trueskill
RIGOTRIO	**53.00**	**79.98**	**1.03**
CMU	50.02	71.91	0.819
FORGe	44.49	58.57	0.458
ISI	26.40	38.60	-1.172
Sheffield	9.46	22.84	-2.132

Table 4: Human judgments of generation results after self-judgments are removed: The results are fundamentally the same

RIG-GOT-RIO → Trio from Riga with regards to GOT & RIO ;)

RIGOTRIO at SemEval-2017 Task 9: Combining Machine Learning and Grammar Engineering for AMR Parsing and Generation

Normunds Gruzitis[1], Didzis Gosko[2] and Guntis Barzdins[1]

[1]University of Latvia, IMCS / Rainis blvd. 29, Riga, Latvia
normunds.gruzitis@lumii.lv, guntis.barzdins@lumii.lv
[2]LETA / Marija street 2, Riga, Latvia
didzis.gosko@leta.lv

"We made the following resources available to participants: [..] The **JAMR** (Flanigan et al., 2016) generation system, as a **strong generation baseline**. [..]" (May & Priyadarshi, 2017)

Abstract

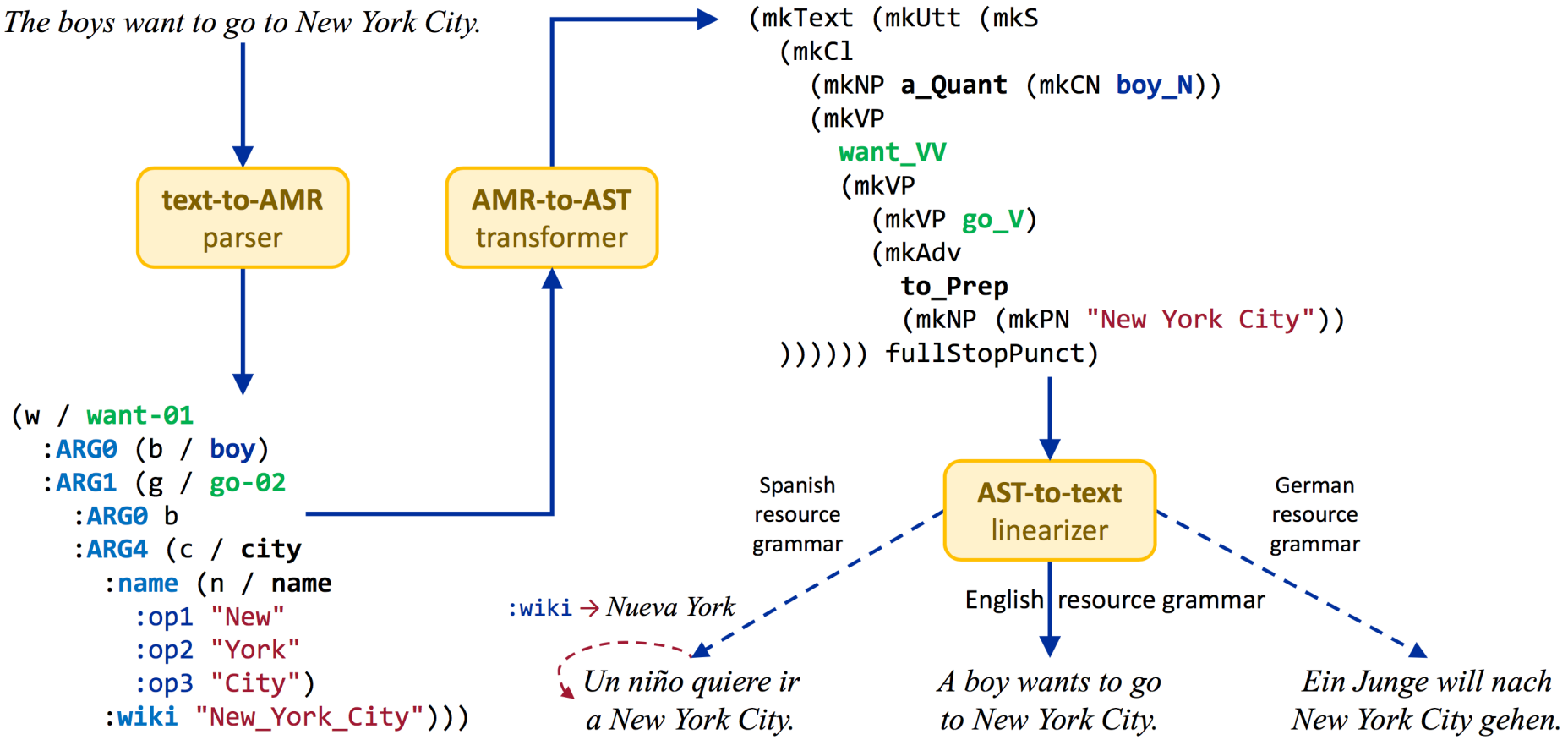
By addressing both text-to-AMR parsing and AMR-to-text generation, SemEval-2017 Task 9 established AMR as a powerful semantic interlingua. We strengthen the interlingual aspect of AMR by applying the multilingual Grammatical Framework (GF) for AMR-to-text generation. Our current rule-based GF approach completely covered only 12.3% of the test AMRs, therefore we combined it with state-of-the-art JAMR Generator to see if the combination increases or decreases the overall performance. The combined system achieved the automatic BLEU score of 18.82 and the human Trueskill score of 107.2, to be compared to the plain JAMR Generator results. As for AMR mon NER tools that are often restricted to types "person", "organization", "location", etc.

The paper starts with NER extensions used for the Biomedical AMR parsing subtask, followed by a novel approach of using Grammatical Framework for AMR generation, and concludes with a brief analysis of our SemEval results.

2 Text-to-AMR parsing

Only two adaptations to the AMR parser from SemEval-2016 (Barzdins and Gosko, 2016) were implemented: it was retrained on the union of LDC2015E86, LDC2016E25, LDC2016E33 and Bio AMR Corpus, and a gazetteer was added to extend the NER coverage to organic compound names found in the Bio AMR Corpus (e.g. "B-Raf enzyme", "dabrafenib small-molecule", etc.). The gazetteer was generalized w.r.t. numbers used

Under the hood

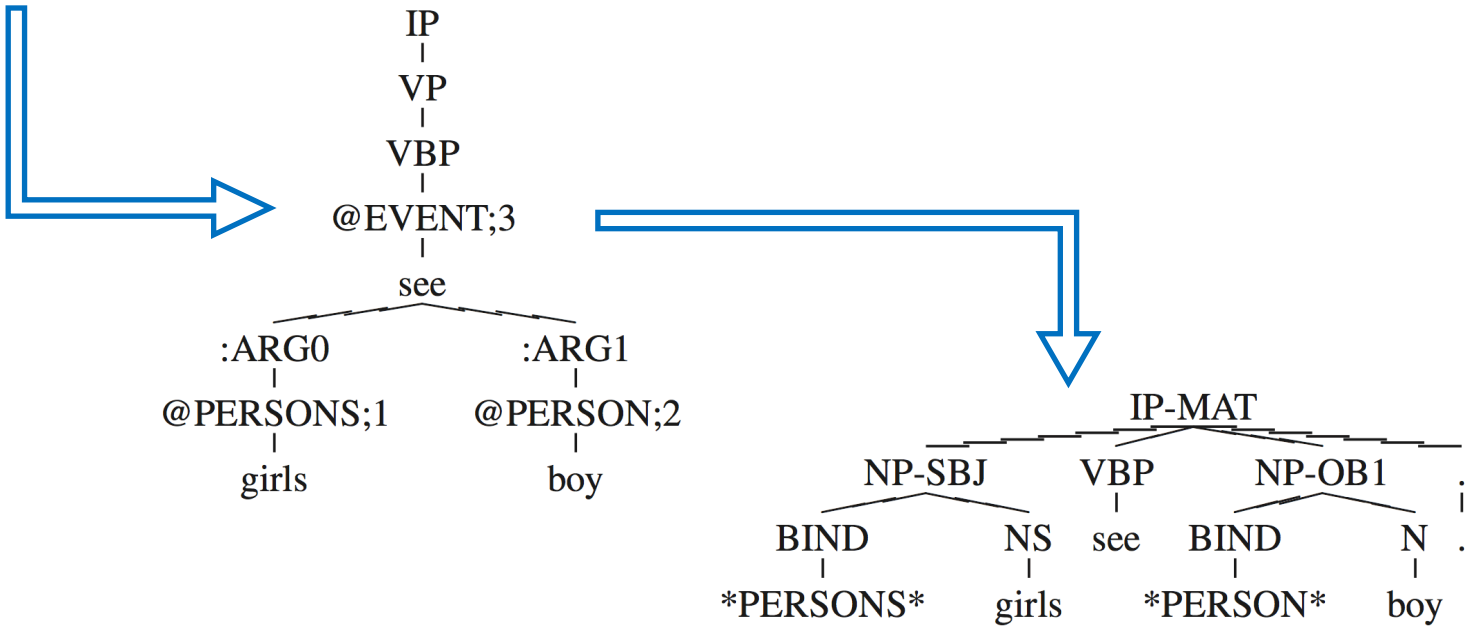



```
# mkVP : VV -> VP -> VP
# (frame1 (:ARG1 (var frame2))) => (frame1 (mkVP frame2))
/VV_FRAME/ < (/:ARG1/=vp < (/VAR/=var < /FRAME/=v))        # Tregex
[move v >1 vp] [relabel vp /^.+$/mkVP/] [delete var]       # Tsurgeon
```

Inspired by Butler (2016)

- Alastair Butler. **Deterministic natural language generation from meaning representations for machine translation**. NAACL Workshop on Semantics-Driven Machine Translation, 2016

```
( EVENT-3 / see
      :ARG0 ( PERSONS-1 / girls)
      :ARG1 ( PERSON-2 / boy))
```



Multilingual AMR-to-Text: experiment

TestTrees: t01_girls_see_a_boy
TestTreesEng: a girl sees a boy .
TestTreesLav: meitene redz zēnu .
TestTreesRus: девочка видит мальчика .

TestTrees: t04_two_pretty_girls_see_a_boy
TestTreesEng: 2 pretty girls see a boy .
TestTreesLav: 2 jaukas meitenes redz zēnu .
TestTreesRus: 2 хорошенькие девочки видят мальчика .

TestTrees: t21_girls_who_see_the_game_like_the_boys_who_play
TestTreesEng: a girl that sees a game likes a boy that plays .
TestTreesLav: meitenei , kas redz spēli , patīk zēns , kas spēlē .
TestTreesRus: девочка , которая видит игру нравдит мальчика , которого играет .

TestTrees: t27_they_are_thugs_and_deserve_a_bullet
TestTreesEng: they are a thug and it deserves a bullet .
TestTreesLav: viņi ir slepkava , un pelna lodi .

Under the hood

The overall AMR-to-text process:

1. The input AMR is rewritten from the **PENMAN** notation to the the **LISP-like** bracketing tree syntax.
2. In case of a **multi-sentence** AMR, the graph is split into two or more graphs to be processed separately.
3. For each AMR, a **sequence of tree pattern-matching transformation rules** is applied (*Tregex* + *Tsurgeon*), acquiring a fully or partially converted GF abstract syntax tree (AST).
4. In case of a **partially converted** AST, the pending subtrees are pruned.**
5. The resulting ASTs are passed to the GF interpreter for RGL-based **linearization**.
6. Since RGL supports many more languages (30+), this approach can be extended to **multilingual** AMR-to-text generation, given a large translation lexicon (15+).

Under the hood

** **Our SemEval submission**:

Because the coverage of our hand-crafted AMR-to-AST transformation rules is currently far from complete, we used **JAMR** Generator (Flanigan et al., 2016) as a "fall-back" option for AMRs that are not fully covered by the current rule set (~200).

However, we applied heuristic **post-processing** rules to the JAMR output, which might have influenced the human judgements:

- Adding a full-stop, or question mark, or exclamation mark at the end of the sentence, or a wh-word at the beginning, based on the AMR constructs.
- Removing the remaining (unresolved) AMR constructs and concepts.
- Converting large numbers into words, adding some prepositions, etc.

..

amrs	AMR-to-text: preparing for SemEval	9 months ago
lexicons	AMR-to-text: an attempt to distinguish frame_VV and frame_VS	7 months ago
rules	AMR-to-text: non-core roles (source, destination, instrument, accompa...	a month ago
tregex	AMR-to-text: non-core roles (source, destination, instrument, accompa...	a month ago
README.md	AMR-to-text: license	a month ago
workflow.png	AMR-to-text generation: documentation	10 months ago

README.md

AMR-to-text generation via GF abstract syntax trees

The initial proposal is concisely described in a position paper by Grūzītis & Bārzdiņš (2016). Details of the current implementation can be found in a system description paper by Grūzītis et al. (2017). See the Publication section below.

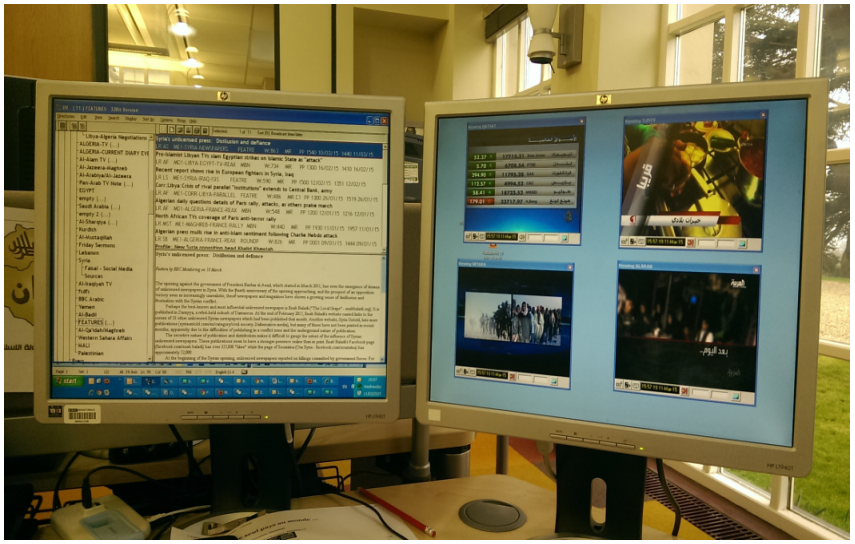
Outline: for a given AMR graph, represented as a tree in the PENMAN notation, transform it to a GF abstract syntax tree (AST), and linearize the AST in the target language. The output sentence is in general a paraphrase of the input sentence represented by the AMR graph.

Structure

1. `lexicons`: monolingual and multilingual GF lexicons – extensions to the wide coverage lexicons provided by the GF resource grammar library (RGL).

The Role of CNL and AMR in Scalable Abstractive Summarization for Multilingual Media Monitoring

Large-scale media monitoring



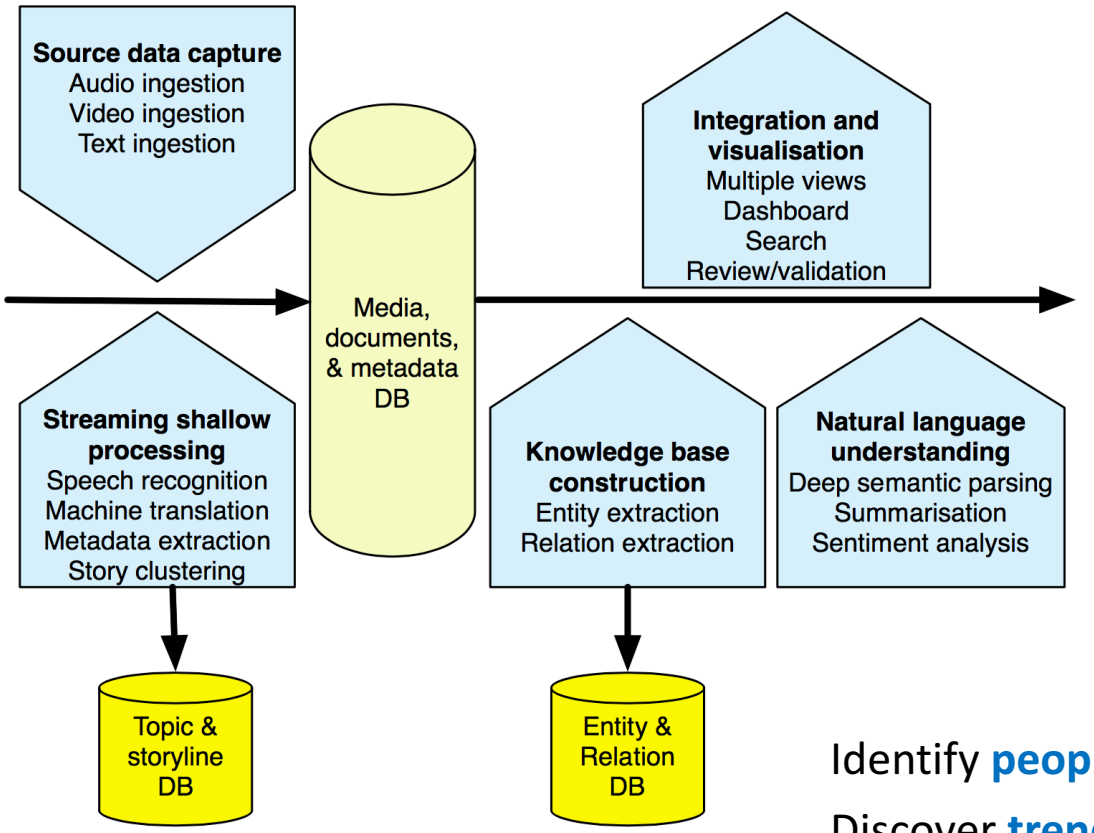
BBC monitoring journalists translate **from 30 languages into English**, follow **400 social media accounts** every day.

A monitoring journalist typically monitors **4 TV channels and several online sources simultaneously**. This is about the **maximum that any person can cope with** mentally and physically. The required **human effort thus scales linearly** with the number of monitored sources.

Monitoring journalists constantly need to be on the lookout for **more sources** and follow **important stories**—but as it is, they are tied down with **mundane, routine monitoring tasks**.

Monitoring **250 video channels** results in a daily buffer of **2.5TB**, a weekly buffer of **19Tb**, and an annual buffer of **1Pb**.

SUMMA –
Scalable Understanding of Multilingual MediA



BBC

Deutsche Welle

The University Of Sheffield.

LETA
nacionālā informācijas aģentūra

priberam

idiap RESEARCH INSTITUTE

H2020 grant No. 688139

Identify **people**, **places**, **events** of interest

Discover **trends**, emerging **events**, crucial new **stories**

Storyline highlights

Article 1: ··· An ongoing battle in Aleppo eventually terminated when the rebels took over the city. ··· President Assad gave a speech, denouncing the death of Syrian soldiers. ···

Article 2: ··· Syrian rebels took control of Aleppo ···

Article 3: ··· The Syrian opposition forces won the battle over Aleppo city. ··· Syrian president announced on Syrian television that such insurgence will not be tolerated. ···

Blog 1: ··· As described in this news story (*link to article 1*) our Syrian brothers are starting to make progress in their opposition to the tyrannic rule of Assad. ···

Output Summary:

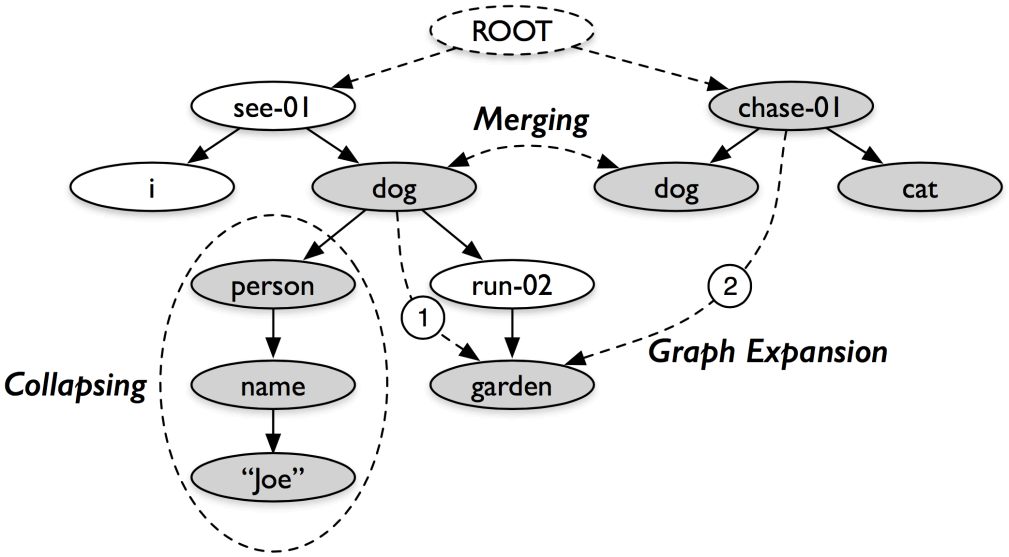
Syrian rebels took over Aleppo
Article1 Article2 Article3 Sentiment: 70% positive, Variance: High

Assad gave a speech about the battle
Article1 Article3 Sentiment: 35% positive, Variance: Low

- Event-based multi-document summarization
- Storyline highlights across a set of related stories

Abstractive text summarization

- Extractive summarization selects representative sentences from the input documents
- Abstractive summarization builds a **semantic representation** from which a summary is **generated**
- What semantic representation?
  - PropBank / FrameNet
  - AMR



Sentence A: I saw Joe's dog, which was running in the garden.
Sentence B: The dog was chasing a cat.
Summary: Joe's dog was chasing a cat in the garden.

Liu F., Flanigan J., Thomson S., Sadeh N., Smith N.A. Toward Abstractive Summarization Using Semantic Representations. NAACL 2015

SemEval 2016 Task 8 on AMR parsing

1. **Riga** (U Latvia, IMCS / LETA): 0.6196
2. **CAMR** (U Brandeis / Boulder Learning Inc. / Rensselaer Polytechnic Institute): 0.6195
3. ICL-HD (Ruprecht-Karls-Universität Heidelberg): 0.6005
4. UCL+Sheffield (University College London / U Sheffield): 0.5983
5. M2L (Kyoto University): 0.5952
6. CMU (Carnegie Mellon University / U Washington): 0.5636
7. CU-NLP (OK Robot Go Ltd. / U Colorado): 0.5566
8. UofR (U Rochester): 0.4985
9. **MeaningFactory** (U Groningen): 0.4702*
10. CLIP@UMD (U Maryland): 0.4370
11. **DynamicPower** (National Institute for Japanese Language and Linguistics): 0.3706*

* Rule/grammar-based; did not use AMR training data

Conclusion

- Unrestricted large-scale **NLU** is difficult for grammars
  - SemEval 2016: **few** grammar-based systems
  - SemEval 2017: **no** grammar-based systems (Boxer gave up…)
- **For NLG, grammar-based systems are very competitive!**
- Scaling up AMR-to-AST:
  - Add more Tregex/Tsurgeon rules
  - A more flexible and systematic graph/tree-transducer (like UD2GF)
  - Learning transformation rules (C6.0; training data?)
  - Seq-to-seq deep learning?

Publications

- Normunds Grūzītis, Pēteris Paikens, Guntis Bārzdiņš. **FrameNet Resource Grammar Library for GF**. CNL 2012

- Dana Dannélls, Normunds Grūzītis. **Extracting a bilingual semantic grammar from FrameNet-annotated corpora**. LREC 2014

- Dana Dannélls, Normunds Grūzītis. **Controlled natural language generation from a multilingual FrameNet-based grammar**. CNL 2014

- Normunds Grūzītis, Dana Dannélls, Benjamin Lyngfelt, Aarne Ranta. **Formalising the Swedish Constructicon in Grammatical Framework**. GEAF 2015

- Normunds Grūzītis, Guntis Bārzdiņš. **The role of CNL and AMR in scalable abstractive summarization for multilingual media monitoring**. CNL 2016

- Normunds Grūzītis, Dana Dannélls. **A Multilingual FrameNet-based Grammar and Lexicon for Controlled Natural Language**. Language Resources and Evaluation, 51(1), 2017

- Normunds Grūzītis, Didzis Goško, Guntis Bārzdiņš. **RIGOTRIO at SemEval-2017 Task 9: Combining Machine Learning and Grammar Engineering for AMR Parsing and Generation**. SemEval 2017