

# GLIF: A Framework for Prototyping Symbolic Natural Language Understanding

Jan Frederik Schaefer

FAU Erlangen-Nürnberg

**Prospects of Formal Mathematics – Bridging between informal and formal**

Hausdorff Research Institute for Mathematics

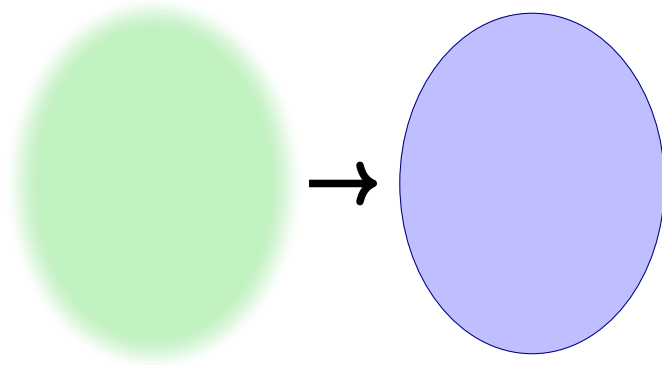
Bonn

July 9, 2024

# Method of Fragments

Natural Language

Logic



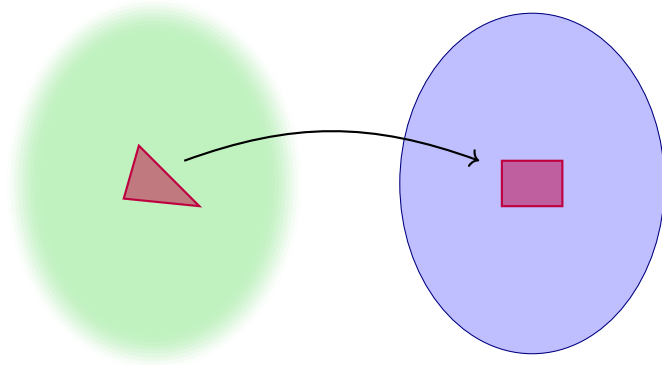
How do we get from messy language to formal logic?

*Montague* [Mon70]: Look at a “nice” subset and map into logic.

# Method of Fragments

Natural Language

Logic



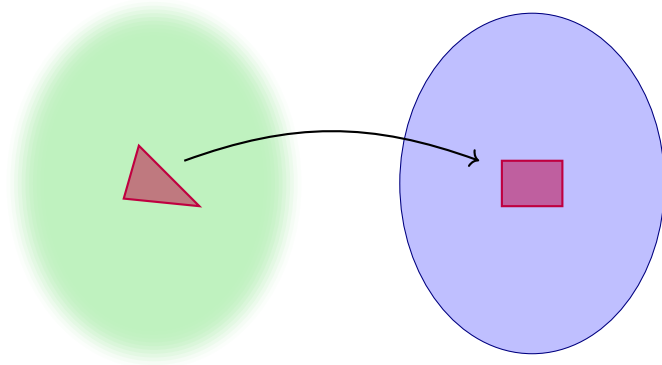
How do we get from messy language to formal logic?

*Montague* [Mon70]: Look at a “nice” subset and map into logic.

# Method of Fragments

Natural Language

Logic



*"Ahmed paints and Berta is quiet."*

*"Ahmed doesn't paint."*

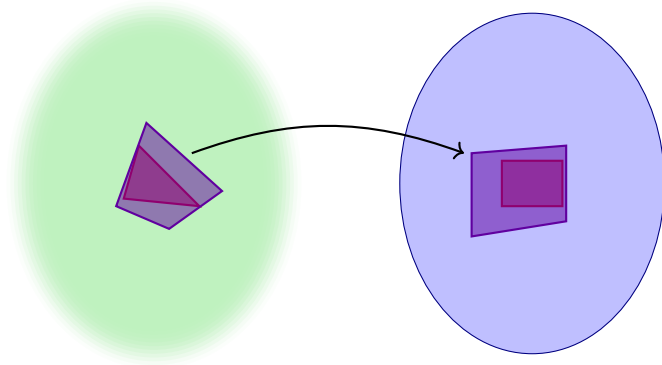
$p(a) \wedge q(b)$

$\neg p(a)$

# Method of Fragments

Natural Language

Logic



*"Every student paints and is quiet."*

*"Nobody paints."*

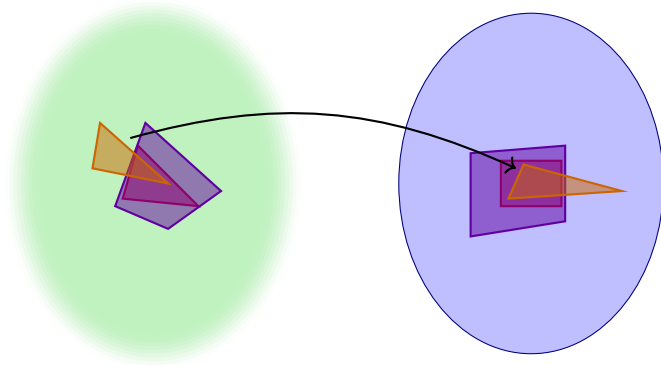
$\forall x.s(x) \Rightarrow (p(x) \wedge q(x))$

$\neg \exists x.p(x)$

# Method of Fragments

Natural Language

Logic



*"Ahmed isn't allowed to paint."*

*"Ahmed and Berta must paint."*

$\neg \Diamond p(a)$

$(\Box p(a)) \wedge \Box p(b)$

# Method of Fragments

Hand-waving is problematic:

“Ahmed paints. He is quiet.”  $\overset{?}{\rightsquigarrow} p(a) \wedge q(a)$

Montague: Specify

- grammar,
- target logic,
- semantics construction.

*fixes NL subset*

*maps parse trees to logic*

*Example from [Mon74]*

- |  |
|--|
| <p>T11. If <math>\phi, \psi \in P_I</math> and <math>\phi, \psi</math> translate into <math>\phi', \psi'</math> respectively, then <math>\phi</math> <b>and</b> <math>\psi</math> translates into <math>[\phi \wedge \psi]</math>, <math>\phi</math> <b>or</b> <math>\psi</math> translates into <math>[\phi \vee \psi]</math>.</p> <p>T12. If <math>\gamma, \delta \in P_{IV}</math> and <math>\gamma, \delta</math> translate into <math>\gamma', \delta'</math> respectively, then <math>\gamma</math> <b>and</b> <math>\delta</math> translates into <math>\hat{x}[\gamma'(x) \wedge \delta'(x)]</math>, <math>\gamma</math> <b>or</b> <math>\delta</math> translates into <math>\hat{x}[\gamma'(x) \vee \delta'(x)]</math>.</p> <p>T13. If <math>\alpha, \beta \in P_T</math> and <math>\alpha, \beta</math> translate into <math>\alpha', \beta'</math> respectively, then <math>\alpha</math> <b>or</b> <math>\beta</math> translates into <math>\hat{P}[\alpha'(P) \vee \beta'(P)]</math>.</p> |
|--|

Claim: That doesn't scale well  $\rightsquigarrow$  **We need prototyping!**

# NLU Prototyping

```
> my-translate "Every student paints and is quiet."
```

```
 $\forall x.s(x) \Rightarrow (p(x) \wedge q(x))$ 
```

```
> my-answer "Every student is quiet. John is a student. Is John quiet?"
```

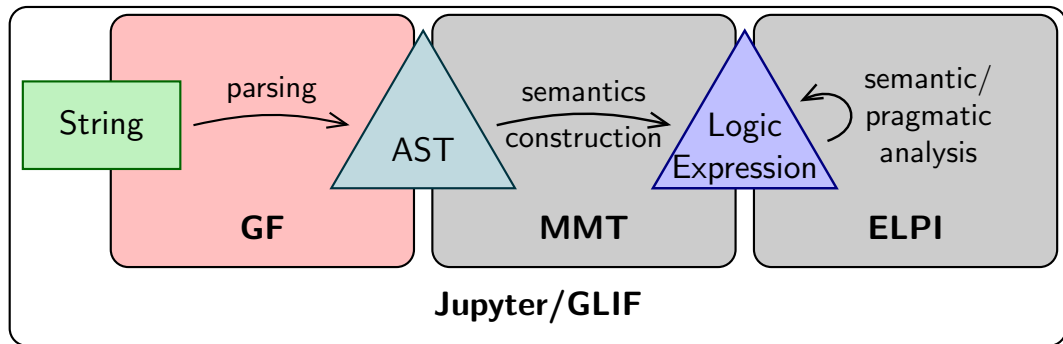
```
 $\forall x.s(x) \Rightarrow q(x), s(j) \vdash? q(j)$ 
```

```
yes
```

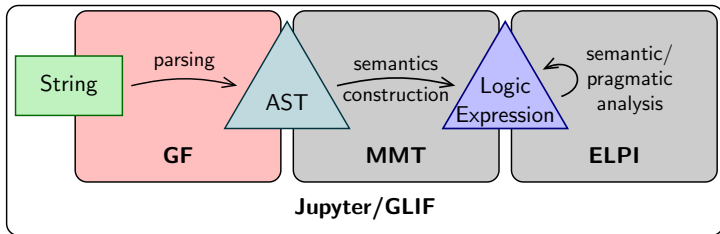
- Traditionally done in Prolog/Haskell
  - requires a lot of work
- A dedicated framework might be better
  - only partial solutions exist
- Can we combine existing partial solutions?
  - ↪ GLIF



## Components of GLIF: GF



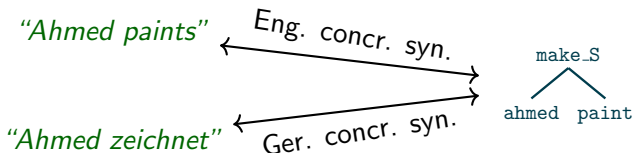
# Components of GLIF: GF



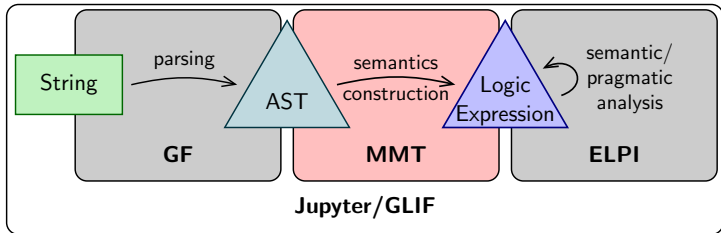
# Components of GLIF: Grammatical Framework [GF]

- Specialized for developing natural language grammars
- Separates abstract and concrete syntax

```
make_S : NP -> VP -> S;                                abstract  
make_S np vp = np.s ++ vp.s!np.n;                        concrete
```
- Abstract syntax based on LF
- Comes with large library  $\geq 36$  *languages*



# Components of GLIF: MMT



# Components of GLIF: MMT

- Modular logic development and knowledge repr.
- Not specialized in one logical framework *we use LF*
- We will use MMT to:
  - ① represent abstract syntax
  - ② specify target logic and discourse domain theory
  - ③ specify semantics construction

# Components of GLIF: MMT

- Modular logic development and knowledge repr.
- Not specialized in one logical framework *we use LF*
- We will use MMT to:
  - ① **represent abstract syntax**
  - ② specify target logic and discourse domain theory
  - ③ specify semantics construction

**GF**

```
cat
  NP; VP; S;
fun
  make_S :
    NP -> VP -> S;
```



**MMT**

```
NP : type
VP : type
S  : type
make_S :
  NP → VP → S
```

# Components of GLIF: MMT

- Modular logic development and knowledge repr.
- Not specialized in one logical framework *we use LF*
- We will use MMT to:
  - ① represent abstract syntax
  - ② **specify target logic and discourse domain theory**
  - ③ specify semantics construction

## Logic Syntax

```
o : type //propositions
¬ : o → o
∧ : o → o → o
∨ : o → o → o

ι : type //individuals
∀ : (ι → o) → o
∃ : (ι → o) → o
```

## Discourse Domain

```
paint : ι → o
quiet : ι → o
ahmed : ι
berta : ι
```

idea:  $\forall f$  or  $\forall \lambda x.f(x)$   
instead of  $\forall x.f(x)$

# Components of GLIF: MMT

- Modular logic development and knowledge repr.
- Not specialized in one logical framework *we use LF*
- We will use MMT to:
  - ① represent abstract syntax
  - ② specify target logic and discourse domain theory
  - ③ **specify semantics construction**

## Semantics Construction

*map symbols in abstract syntax to terms in logic/domain theory*

Simple setting

S	$\mapsto$ o
NP	$\mapsto$ ι
VP	$\mapsto$ ι $\rightarrow$ o
make_S	$\mapsto$ $\lambda n. \lambda v. v$ n
ahmed	$\mapsto$ ahmed

More advanced

NP	$\mapsto$ (ι $\rightarrow$ o) $\rightarrow$ o
sentence	$\mapsto$ $\lambda n. \lambda v. n$ v
everyone	$\mapsto$ $\lambda p. \forall \lambda x. p$ x
berta	$\mapsto$ $\lambda p. p$ berta



# Example: Parsing + Semantics Construction

*“Ahmed and Berta paint”*

↓ parsing

make\_S (andNP ahmed berta) paint

↓ semantics construction

$(\lambda n. \lambda v. n \ v) \ ((\lambda a. \lambda b. \lambda p. a \ p \ \wedge \ b \ p) \ (\lambda p. p \ \text{ahmed}) \ (\lambda p. p \ \text{berta})) \ \text{paint}$

↓  $\beta$ -reduction

$\text{paint} \ \text{ahmed} \ \wedge \ \text{paint} \ \text{berta}$

## Example: Input Language for SageMath

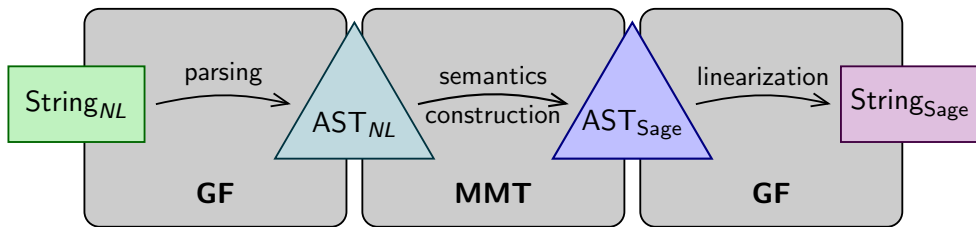
- Can we make a natural input language for SageMath?

*WolframAlpha-like*

```
sage: g = AlternatingGroup(5)
sage: g.cardinality()
60
```

*“Let  $G$  be the alternating group on 5 symbols. What is the cardinality of  $G$ ?”*

## Example: Input Language for SageMath



## Example: Input Language for SageMath

> Let  $G$  be the alternating group on 5 symbols.

```
# G = AlternatingGroup(5)
```

> Let  $|H|$  be a notation for the cardinality of  $H$ .

```
# def bars(H): return H.cardinality()
```

> What is  $|G|$ ?

```
# print(bars(G))
```

```
60
```

> Let  $A_N$  be a notation for the alternating group on  $N$  symbols.

```
# def A(N): return AlternatingGroup(N)
```

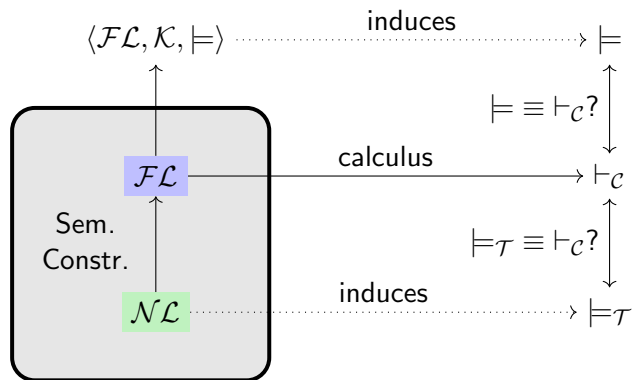
> What are the cardinalities of  $A_4$  and  $A_5$ ?

```
# print(A(4).cardinality()); print(A(5).cardinality())
```

```
12
```

```
60
```

# Levels of inference



- 1 Test: Does "*Ahmed and Berta paint.*"  $\models_{\mathcal{T}}$  "*Berta paints.*"?
- 2 Model prediction: Yes, because  $p(a) \wedge p(b) \vdash_c p(b)$ .
- 3 Correct result: Ask people.

## Natural deduction in MMT: “*Judgments as types*”

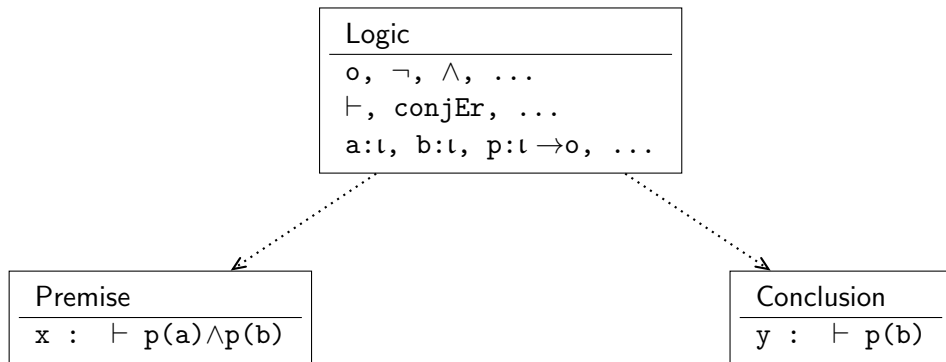
$\vdash : o \rightarrow type$

$s1 : \vdash p(a) \wedge p(b)$

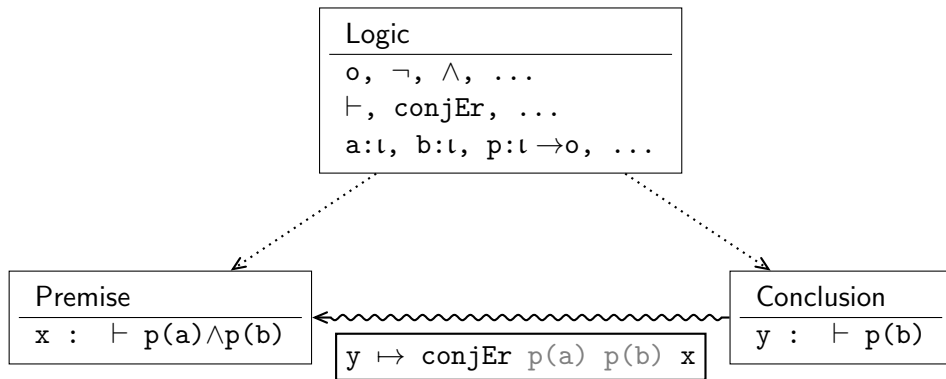
$conjEr : \{A:o\} \rightarrow \{B:o\} \rightarrow \vdash A \wedge B \rightarrow \vdash B$

$s2 : \vdash p(b)$   
 $\quad = conjEr\ p(a)\ p(b)\ s1$

$\llbracket \text{"Ahmed and Berta paint"} \rrbracket \vdash_{\mathcal{ND}} \llbracket \text{"Berta paints"} \rrbracket$

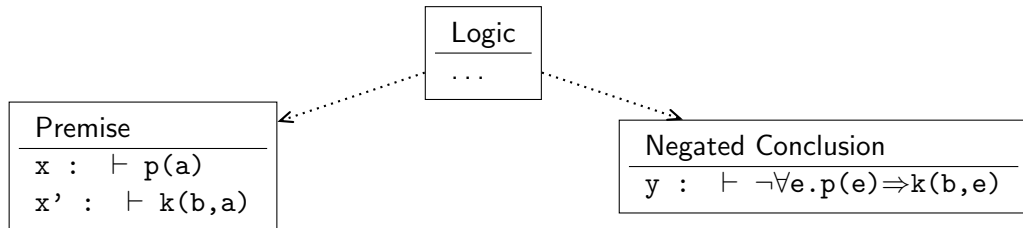


$\llbracket \text{"Ahmed and Berta paint"} \rrbracket \vdash_{\mathcal{ND}} \llbracket \text{"Berta paints"} \rrbracket$

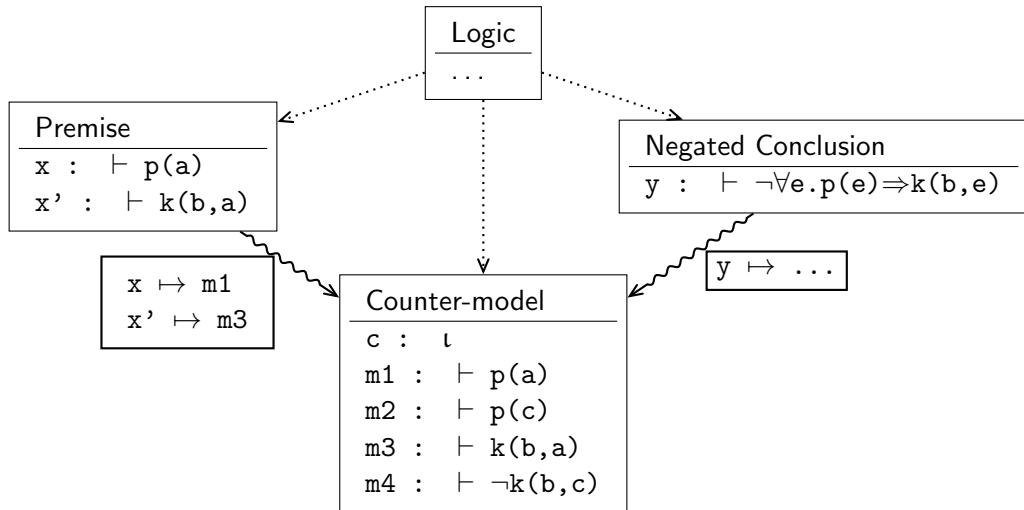




$\llbracket \text{"Ahmed paints"} \rrbracket, \llbracket \text{"Berta knows Ahmed"} \rrbracket \not\models_{\mathcal{ND}} \llbracket \text{"Berta knows everyone who paints"} \rrbracket$



$\llbracket \text{"Ahmed paints"} \rrbracket, \llbracket \text{"Berta knows Ahmed"} \rrbracket \not\models_{ND} \llbracket \text{"Berta knows everyone who paints"} \rrbracket$

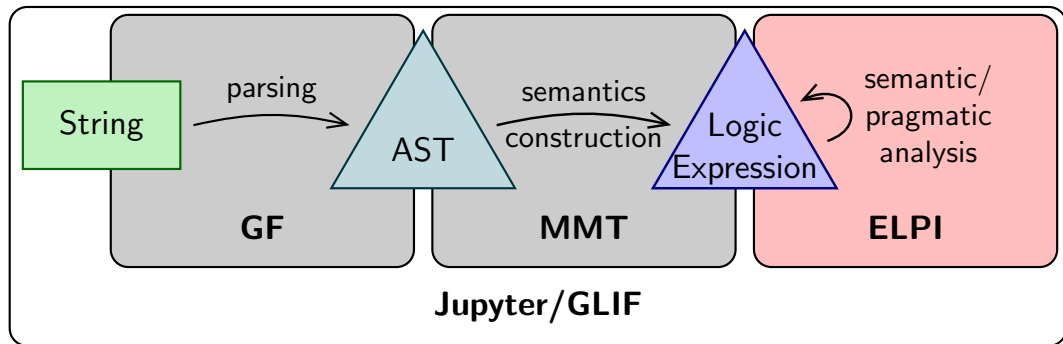


## Mini summary

- Parsing with GF
- Logic syntax in MMT
- Semantics construction in MMT
- (Manual) inference in MMT

*“Bring your own logic”*

## Components of GLIF: ELPI



# Components of GLIF: ELPI

- Implementation and extension of  $\lambda$ Prolog
- MMT can generate logic signatures
- Generic inference/reasoning step after semantics construction

$\approx$  *Prolog + HOAS*

## MMT

$o : \text{type} \text{ // } \textit{propositions}$

$\neg : o \rightarrow o$

$\wedge : o \rightarrow o \rightarrow o$

$\vee : o \rightarrow o \rightarrow o$

$\iota : \text{type} \text{ // } \textit{individuals}$

$\forall : (\iota \rightarrow o) \rightarrow o$

$\exists : (\iota \rightarrow o) \rightarrow o$

## ELPI

`kind o type.`

`not : o -> o.`

`and : o -> o -> o.`

`or : o -> o -> o.`

`kind i type.`

`type forall (i -> o) -> o.`

`type exists (i -> o) -> o.`

## Example: Discard wrong readings in controlled natural language

*"the ball has a mass of 5kg"*  $\rightarrow$  AST  $\longrightarrow$  `mass(theball, quant(5, kilo gram))`

## Example: Discard wrong readings in controlled natural language

*"the ball has a mass of 5kg"*  $\rightarrow$  AST  $\longrightarrow$  `mass(theball, quant(5, kilo gram))`

*"a kinetic energy of 12mN"*  $\nearrow$  AST<sub>1</sub>  $\longrightarrow$   `$\lambda x.E_{\text{kin}}(x, \text{quant}(12, \text{milli Newton}))$`

$\searrow$  AST<sub>2</sub>  $\longrightarrow$   `$\lambda x.E_{\text{kin}}(x, \text{quant}(12, \text{meter} \cdot \text{Newton}))$`

## Example: Discard wrong readings in controlled natural language

*"the ball has a mass of 5kg"*  $\rightarrow$  AST  $\longrightarrow$  `mass(theball, quant(5, kilo gram))`

*"a kinetic energy of 12mN"*  $\rightarrow$  AST<sub>1</sub>  $\longrightarrow$   ~~`$\lambda x.E_{kin}(x, \text{quant}(12, \text{milli Newton}))$`~~

*"a kinetic energy of 12mN"*  $\rightarrow$  AST<sub>2</sub>  $\longrightarrow$   `$\lambda x.E_{kin}(x, \text{quant}(12, \text{meter} \cdot \text{Newton}))$`



```
In [20]: 1 parse "the ball has a mass of 5 k g and a kinetic energy of 12 m N" |
          2 construct
```

(mass theball (quant 5 kilo gram))  $\wedge$  (ekin theball (quant 12 milli Newton))  
(mass theball (quant 5 kilo gram))  $\wedge$  (ekin theball (quant 12 meter·Newton))

```
In [21]: 1 parse "the ball has a mass of 5 k g and a kinetic energy of 12 m N" |
          2 construct | filter -predicate=filter pred
```

(mass theball (quant 5 kilo gram)) ^ (ekin theball (quant 12 meter·Newton))

## Example: ForTheL

```
parse -cat=DefinitionStatement "a subset of S is a set T such that every element of T belongs to S"
```

```

$$\forall[V\_T:\mathbb{1}](\text{subset } V\_T V\_S) \Leftrightarrow (\text{set } V\_T) \wedge \forall[V\_new:\mathbb{1}](\text{element } V\_new V\_T) \wedge T \Rightarrow (\text{belongsTo } V\_new V\_S) \wedge T$$

```

```
parse -cat=Statement "there exists an empty set" | construct -v semantics/forthelUnsortedSem
```

```

$$\exists[V\_new:\mathbb{1}]((\text{empty } V\_new) \wedge (\text{set } V\_new)) \wedge T$$

```

```
parse -cat=Statement "S is a subset of every set iff S is empty" | construct -v semantics/forthelUn
```

```

$$(\forall[V\_new:\mathbb{1}](\text{set } V\_new) \wedge T \Rightarrow (\text{subset } V\_S V\_new) \wedge T) \Leftrightarrow (\text{empty } V\_S)$$

```

## Example: “pairwise disjoint”

*“A, B and C are pairwise disjoint”*

$\text{disjoint}(A, B) \wedge \text{disjoint}(A, C) \wedge \text{disjoint}(B, C)$

- **Approach 1**

Semantics construction with lots of  $\lambda$ s:

$\text{disjoint}(A, B) \wedge \text{disjoint}(A, C) \wedge \top \wedge \text{disjoint}(B, C) \wedge \top \wedge \top \wedge \top$

*difficult!*

Simplify with ELPI:

$\text{disjoint}(A, B) \wedge \text{disjoint}(A, C) \wedge \text{disjoint}(B, C)$

## Example: “pairwise disjoint”

*“A, B and C are pairwise disjoint”*

$\text{disjoint}(A, B) \wedge \text{disjoint}(A, C) \wedge \text{disjoint}(B, C)$

- **Approach 1**

Semantics construction with lots of  $\lambda$ s:

$\text{disjoint}(A, B) \wedge \text{disjoint}(A, C) \wedge \top \wedge \text{disjoint}(B, C) \wedge \top \wedge \top \wedge \top$

*difficult!*

Simplify with ELPI:

$\text{disjoint}(A, B) \wedge \text{disjoint}(A, C) \wedge \text{disjoint}(B, C)$

- **Approach 2**

Semantics construction creates preliminary expression:

$\text{relNT disjoint (cons A (cons B (cons C nil)))}$

Convert with ELPI:

$\text{disjoint}(A, B) \wedge \text{disjoint}(A, C) \wedge \text{disjoint}(B, C)$

*easier*

## Example: Epistemic Q&A

*John knows that Mary or Eve knows that Ping has a dog.* ( $S_1$ )

*Mary doesn't know if Ping has a dog.* ( $S_2$ )

*Does Eve know if Ping has a dog?* ( $Q$ )

$$S_1 = \Box_{john}(\Box_{mary}hd(ping) \vee \Box_{eve}hd(ping))$$

$$S_2 = \neg(\Box_{mary}hd(ping) \vee \Box_{mary}\neg hd(ping))$$

$$Q = \Box_{eve}hd(ping) \vee \Box_{eve}\neg hd(ping)$$

$$S_1, S_2 \vdash_{S5_n} Q \quad \rightsquigarrow \quad \text{yes}$$

$$S_1, S_2 \vdash_{S5_n} \neg Q \quad \rightsquigarrow \quad \text{no}$$

$$\text{else} \quad \rightsquigarrow \quad \text{unknown}$$

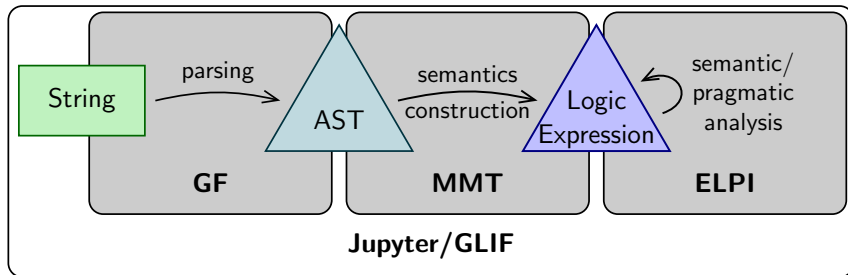
# Conclusion

## Summary:

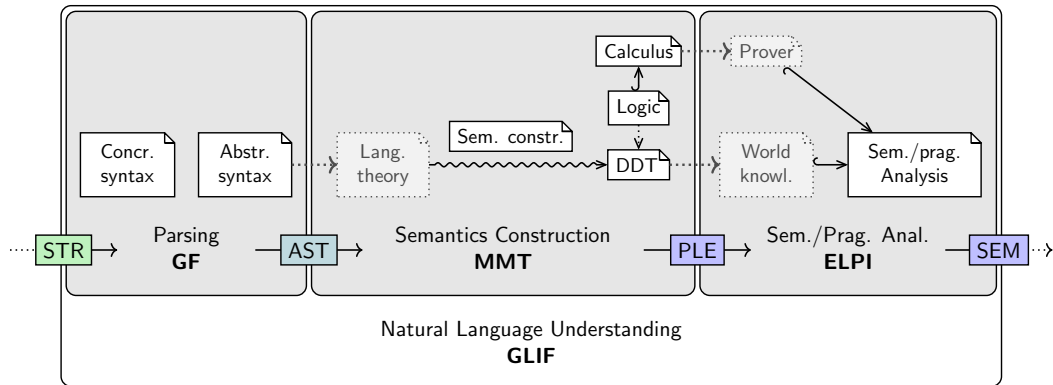
- $GLIF = GF + MMT + ELPI$
- Prototyping natural language understanding
- We use it for teaching

## Examples:

- ① *“What is the cardinality of G?”*
- ② *“a kinetic energy of 12mN”*
- ③ *“A, B and C are pairwise disjoint”*
- ④ *“John knows that Eve has a dog”*



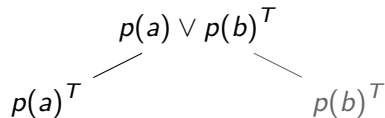
# Pipeline Specification



## Example: Tableaux Machine [**KohKol:ramgpm03**]

- Can use tableaux for model generation
- Tableau machine: pick “best” branch as model and continue there with next sentence *like a human?*

*“Ahmed or Berta paints”*



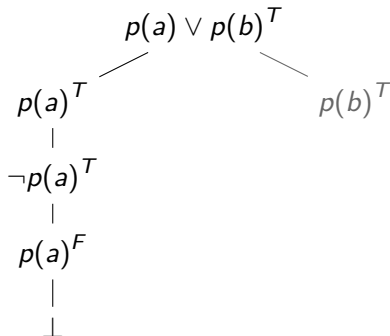


## Example: Tableaux Machine [KohKol:ramgpm03]

- Can use tableaux for model generation
- Tableau machine: pick “best” branch as model and continue there with next sentence  
*like a human?*

*“Ahmed or Berta paints”*

*“Ahmed doesn’t paint”*

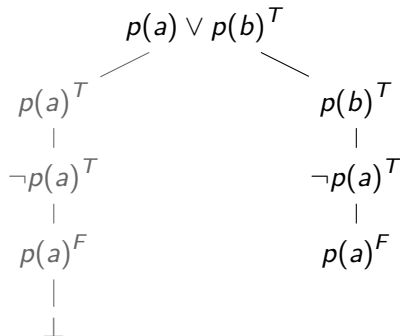


## Example: Tableaux Machine [**KohKol:ramgpm03**]

- Can use tableaux for model generation
- Tableau machine: pick “best” branch as model and continue there with next sentence  
*like a human?*

*“Ahmed or Berta paints”*

*“Ahmed doesn’t paint”*



# Example: Tableaux Machine

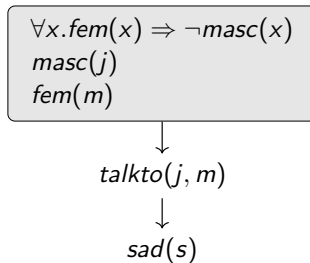
Background Knowledge

*"John talks to Mary."*

*talkto(j, m)*

*"Sasha is sad."*

*sad(s)*



# Example: Tableaux Machine

Background Knowledge

*"John talks to Mary."*

$talkto(j, m)$

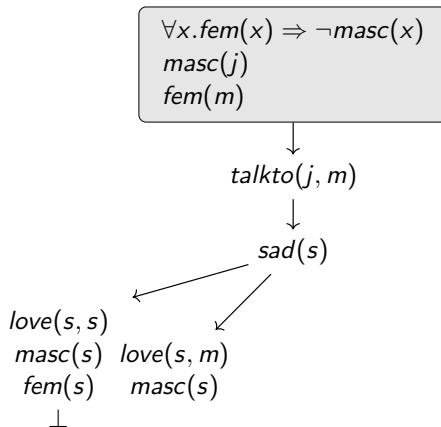
*"Sasha is sad."*

$sad(s)$

*"He loves her."*

$\exists X.masc(X) \wedge$

$\exists Y.fem(Y) \wedge love(X, Y)$



# Example: Tableaux Machine

Background Knowledge

*"John talks to Mary."*

$talkto(j, m)$

*"Sasha is sad."*

$sad(s)$

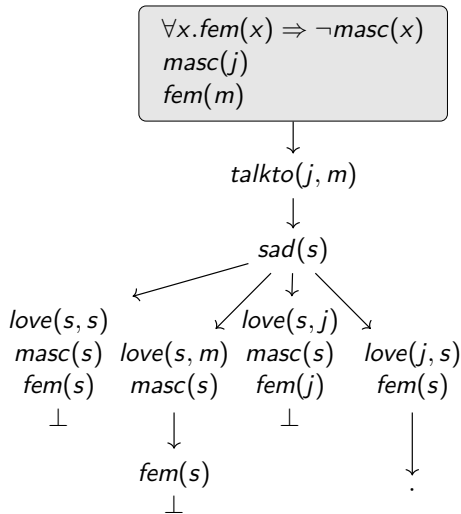
*"He loves her."*

$\exists X.masc(X) \wedge$

$\exists Y.fem(Y) \wedge love(X, Y)$

*"Sasha is a woman."*

$fem(s)$



# Example: Tableaux Machine

Background Knowledge

*"John talks to Mary."*

$talkto(j, m)$

*"Sasha is sad."*

$sad(s)$

*"He loves her."*

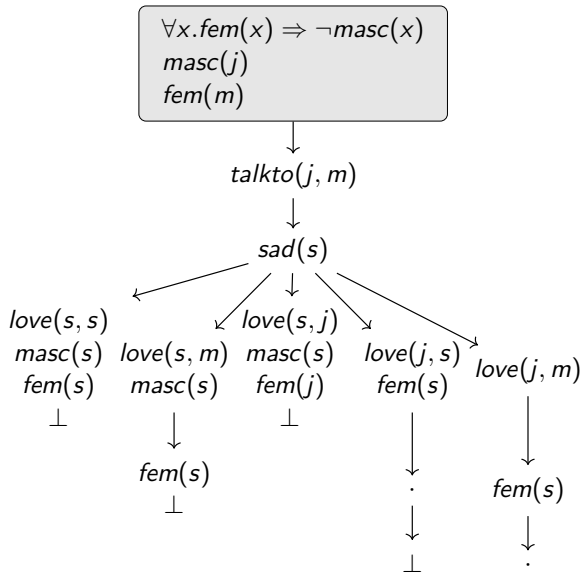
$\exists X.masc(X) \wedge$   
 $\exists Y.fem(Y) \wedge love(X, Y)$

*"Sasha is a woman."*

$fem(s)$

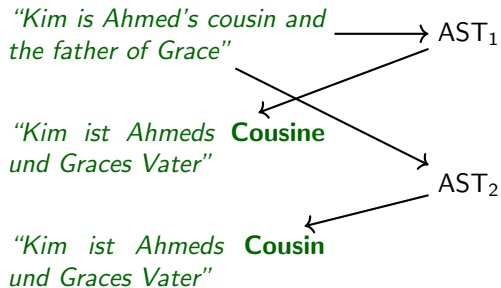
*"John doesn't love Sasha."*

$\neg love(j, s)$



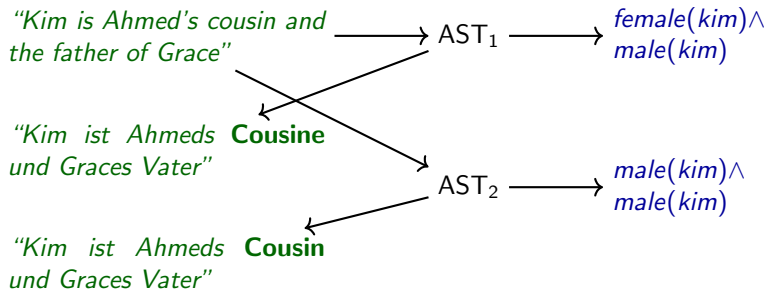
## Example: Translation

- Two German words for “*cousin*”, depending on the gender
- Two entries in abstract syntax: `cousin_female` and `cousin_male`
- Use inference to discard ASTs



## Example: Translation

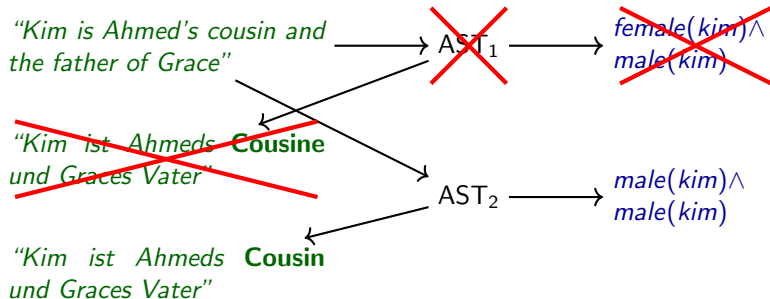
- Two German words for “*cousin*”, depending on the gender
- Two entries in abstract syntax: `cousin_female` and `cousin_male`
- Use inference to discard ASTs





## Example: Translation

- Two German words for “*cousin*”, depending on the gender
- Two entries in abstract syntax: `cousin_female` and `cousin_male`
- Use inference to discard ASTs



# Natural Deduction in MMT/LF

$$\frac{A \wedge B}{A} \wedge E$$

$$\frac{A \vee B \quad \begin{array}{c} [A]^1 \\ \vdots \\ C \end{array} \quad \begin{array}{c} [B]^1 \\ \vdots \\ C \end{array}}{C} \vee E^1$$

//  $\vdash X$  is type of proofs for  $X$  (judgments as types)  
 $\vdash : o \rightarrow type$

$\wedge E1 : \prod_{A:o} \prod_{B:o} \vdash A \wedge B \rightarrow \vdash A$

$\vee E : \prod_{A:o} \prod_{B:o} \prod_{C:o} \vdash A \vee B \rightarrow (\vdash A \rightarrow \vdash C) \rightarrow (\vdash B \rightarrow \vdash C) \rightarrow \vdash C$

# Generating Provers in ELPI

**LF rule**       $\wedge E1 : \prod_{A:o} \prod_{B:o} \vdash A \wedge B \rightarrow \vdash A$

**ELPI equivalent**

direct: `pi A \ pi B \ ded (and A B) => ded A.`

syn. sugar: `ded A :- ded (and A B).`

# Generating Provers in ELPI

**LF rule**       $\wedge E1 : \prod_{A:o} \prod_{B:o} \vdash A \wedge B \rightarrow \vdash A$

**ELPI equivalent**

direct: `pi A \ pi B \ ded (and A B) => ded A.`

syn. sugar: `ded A :- ded (and A B).`

**Example: Or-Elimination**

LF:       $\vee E : \prod_{A:o} \prod_{B:o} \prod_{C:o} \vdash A \vee B \rightarrow (\vdash A \rightarrow \vdash C) \rightarrow (\vdash B \rightarrow \vdash C) \rightarrow \vdash C$

ELPI: `ded C :- ded (or A B), ded A => ded C, ded B => ded C.`

**Example: Forall-Introduction**

LF:       $\forall I : \prod_{P:l \rightarrow o} (\prod_{x:l} \vdash P \ x) \rightarrow \vdash \forall P$

ELPI: `ded (forall P) :- pi x \ ded (P x).`

# Controlling the Proof Search

- Problem: Search diverges *searching harder than checking*
- Solution: Control search with helper predicates:  
*inspired by ProofCert project by Miller et al.*
  - Intuition: Decide whether to apply rule
  - Do not affect correctness
  - Extra argument tracks aspects of proof state

Before: `ded A :- ded (and A B) .`

Now: `ded X A :- help/andEl X A B X1, ded X1 (and A B) .`

# Helper Predicates

Name	Predicate	Argument
Iter. deepening	checks depth	remaining depth
Proof term	generates term	proof term
Product	calls other predicates	arguments for other predicates
Backchaining	Prolog's backchaining ( $\approx$ forward reasoning from axioms via $\Rightarrow/\forall$ elimination rules)	pattern of formula to be proven (e.g. a conjunction)

## Example helper: Iterative deepening

```
help/andEl (idcert  $N$ ) _ _ (idcert  $N1$ ) :-  $N > 0$ ,  $N1$  is  $N - 1$ .
```

# Tableau Provers

$$\frac{A \wedge B^F}{A^F \mid B^F} \wedge^F \qquad \frac{A \wedge B^F \quad \begin{array}{c} [A^F] \\ \vdots \\ \perp \end{array} \quad \begin{array}{c} [B^F] \\ \vdots \\ \perp \end{array}}{\perp} \wedge^F$$

LF:  $\wedge^F : \prod_{A:o} \prod_{B:o} A \wedge B^F \rightarrow (A^F \rightarrow \perp) \rightarrow (B^F \rightarrow \perp) \rightarrow \perp$

ELPI: `closed X :- help/andF X A B X1 X2 X3, f X1 (and A B),  
f/hyp A => closed X2, f/hyp B => closed X3.`

With iterative deepening we get a working prover!

→ Other helpers result in more efficient provers

# References I

- [GF] *GF - Grammatical Framework*. URL:  
<http://www.grammaticalframework.org> (visited on 09/27/2017).
- [Mon70] R. Montague. “English as a Formal Language”. In: Reprinted  
in [Thomason:fp74], 188–221. Edizioni di Comunita, Milan, 1970,  
pp. 189–224.
- [Mon74] Richard Montague. “The Proper Treatment of Quantification in Ordinary  
English”. In: *Formal Philosophy. Selected Papers*. Ed. by R. Thomason. New  
Haven: Yale University Press, 1974.