"Introduction to Deep Inference and Proof Nets"

P. Braschi & L. Straßburger

(U. Bath) (INRIA - Futur)

LECTURE 2:

Content:

- Quick background on proof complexity
- Sequent calculus vs. deep inference
  - Analytic calculi (cut-free)
- Frege systems vs. deep inference
Quick Background on Proof Complexity

- We deal with propositional classical logic.
- The validity problem is **CONP-complete**, i.e.: a certificate stating that a formula is not valid can be checked in polynomial time on the size of the formula.

- What about certifying validity?
  - We need proofs.
  - We need proof systems, i.e. algorithms that check proofs in polynomial time on the size of the proof:
    - Gentzen systems (sequent calculus)
    - Frege systems (Hilbert systems)
    - resolution
    - Tableaux
    - ...!

  *NOT PROOF NETS!!*

- Can we check validity in time that is polynomial on the size of the formula?
  - If YES then \( \text{Co-NP} = \text{NP} \)
  - If NO then \( \text{Co-NP} \neq \text{NP} \); hence \( P \neq \text{NP} \).
More background on proof complexity

- Study the relative strength of proof systems
  - \( p \)-simulation:
    Proof system \( U \) \( p \)-simulates proof system \( V \) if
    proofs in \( V \) can be converted into proofs in \( U \)
    efficiently, i.e. in polynomial time (on their size)
    \( U \rightarrow V \)
  - Exponential speed-up:
    Proof system \( U \) has an exponential speed-up over
    proof system \( V \) if some proofs in \( U \) are
    exponentially shorter than the best proofs in \( V \),
    for some set of tautologies.

Example:

\[
\begin{align*}
\text{Gentzen with cut} & \quad \text{p-sim.} & \quad \text{cut-free Gentzen} \\
\text{exp. s.u.} & \quad \text{(over Sauer-Wu's tautologies, pigeon-hole-principle)} & \quad \ldots
\end{align*}
\]

- \( p \)-equivalence
  - \( p \)-simulation in both directions.
- Size of a formula/derivation: the number of units/atom-vars
  occurrences in there. \( \| \cdot \| \)
Legend:

- **CoS**: calculus of structures (deep inference)
- \( F \rightarrow G \): proof system \( F \) polynomially simulate proof system \( G \)
- \( F \not\rightarrow G \): it doesn't happen.

Diagram:

- Analytic CoS
- Cos
- Brünner 04
- Statman 78
- Gentzen
- Frege
- Cos + Extension
- Cook, Reckhow 74
- \( \rightarrow \)
- Krajířek - Pudlák '89
- Frege + Substitution
- \( \rightarrow \)
- Cook, Reckhow 79
- Frege + Extension

Arrows indicate relationships between the systems and theorems.
Measures in deep inference

- Formulae/structures in d.i. are modulo an equality relation. The same in derivations/proofs. How much complexity is hidden behind $=$?

Quadratic (polynomial)

- In formulae:
  Decide if $\alpha = \beta$ in polynomial time by reducing $\alpha$ and $\beta$ to some canonical form and comparing them.

Ex: canonical form:
- remove as many units as possible
- order units, atoms and variables
- normal form for assoc + unit: left to right.

$$[	ext{bvd}] \lor [\text{cva}] \sim$$
$$[	ext{cva}] \lor [	ext{bvd}] \sim$$
$$[	ext{avc}] \lor [	ext{bvd}] \sim$$
$$[	ext{av} [\text{cv} [\text{bvd}]]] .$$

If the size of the formula is $n$, repeat the process $n$ times.
• Derivations in CoS: (modify)

\[ \frac{\alpha_0}{\alpha_1} \frac{\alpha_1}{\alpha_2} \cdots \frac{\alpha_{k-1}}{\alpha_k} \]

where \( \eta_i \) alternate between = and any rule of the system.

• Recall:

A CoS proof system is implicationally complete if for every valid implication \( \alpha \rightarrow \beta \) there is a derivation with premises \( \alpha \) and conclusion \( \beta \).

Sks, SkSg are implicationally complete.

• What is the relative complexity of proofs in Sks wrt those in SkSg?

Remember the transformation from

- \( \mathfrak{c}t \rightarrow \{ \text{fact, m}_f \} \)
- \( \mathfrak{w}t \rightarrow \{ \text{awt, s} \} \)
- \( \mathfrak{i}t \rightarrow \{ \text{ait, s} \} \)

They are all quadratic:

• Thus: ks and ksg are \( \beta \)-equivalent

Sks and SkSg are \( \beta \)-equivalent
(Robustness Theorem) All systems in sequent calculus are $\beta$-equivalent, pairwise.

We have seen how to translate CoS$\varphi$: what complexity? (Quadratic, of course)

Thm: For every Gentzen derivation $\Delta$ with premises $\phi_1, \ldots, \phi_n$ and conclusion $\psi$, there is a derivation $\Xi$: $(\phi_1, \ldots, \phi_n)$

\[
\phi_1 \rightarrow \cdots \rightarrow \phi_n \rightarrow \psi
\]

if $n$ is the size of $\Delta$, then

$O(n^2)$ is the size of $\Xi$.

Moreover,

if $\Delta$ is analytic (cut-free) then $\Xi$ is in $kS\varphi$

Cor: $S\varphi kS\varphi$ $\beta$-simulates Gentzen

$kS\varphi$ $\beta$-simulates analytic Gentzen
Analytic Gentzen vs Analytic CoS

- Statman tautologies
  - formulas grow polynomially
  - their proofs grow exponentially in analytic Gentzen systems
  - and polynomially in analytic CoS.
  
  (they grow polynomially in Gentzen + cut)

- The proof of these tautologies in CoS looks very different than in Gentzen:
  deep inference has a key role for
  - speed-up
  - structuring the proof in a more intuitive way (semantics...)

Statman's tautologies

\[ G_0 = (c_0 \lor d_0) \]
\[ \Rightarrow (c_0 \lor d_0) \]

\[ G_1 = (c_1 \lor d_1) \]
\[ \land ((c_1 \lor d_1) \Rightarrow c_0) \]
\[ \lor ((c_1 \lor d_1) \Rightarrow d_0)) \]
\[ \Rightarrow (c_0 \lor d_0) \]

\[ G_2 = (c_2 \lor d_2) \]
\[ \land (((c_2 \lor d_2) \Rightarrow c_1) \]
\[ \lor (c_2 \lor d_2) \Rightarrow d_1) \]
\[ \land (((c_2 \lor d_2) \land (c_1 \lor d_1)) \Rightarrow c_0) \]
\[ \lor ( ((c_2 \lor d_2) \land (c_1 \lor d_1)) \Rightarrow d_0)) \]
\[ \Rightarrow (c_0 \lor d_0) \]

...
Statman's tautologies

\[ G_0 = (c_0 \lor d_0) \quad \iff \quad [(c_0 \overline{d_0}) \quad \overline{d_0}] \]

\[ G_1 = (c_1 \lor d_1) \quad \iff \quad [(\overline{c_1} \overline{d_1}) \quad (c_1 \lor d_1)] \quad (\overline{c_0} \ \overline{d_0}) \]

\[ G_2 = (c_2 \lor d_2) \quad \iff \quad [(\overline{c_2} \overline{d_2}) \quad (c_2 \lor d_2)] \quad (\overline{c_0} \ \overline{d_0}) \]

notation:  \[ [a \ b \ c] = abvbc \]
\[ (a \ b \ c) = a \land bnc \]
\[ \overline{a} = \overline{\overline{a}} \]

(Statman) formulae grow linearly, their proofs grow exponentially
Gentzen system (one sided)

\begin{align*}
\text{id} & \quad \frac{A \rightarrow A}{A} \\
\text{\wedge} & \quad \frac{A \rightarrow \frac{A \rightarrow C}{C}}{A \rightarrow (A \wedge C)} \\
\text{\vee} & \quad \frac{A \rightarrow \frac{A \rightarrow B}{B}}{A \rightarrow (A \wedge B)} \\
\text{c} & \quad \frac{A \rightarrow \frac{A \rightarrow B \rightarrow B}{B}}{A \rightarrow (A \rightarrow B)}
\end{align*}

Rules can be applied at the root only!
(non-deep inference)

Of course provable:

\[c_2 \land ((c_2 \lor d_2) \rightarrow c_1) \land (((c_2 \lor d_2) \land (c_1 \lor d_1)) \rightarrow d_0) \rightarrow c_0 \land d_0\]

\[
\left[ \bar{c}_2 \left( [c_2 d_2] \bar{c}_1 \right) \left( [c_2 d_2] \bar{c}_1 \bar{d}_0 \right) c_0 \land d_0 \right]
\]

\[
\left[ A_2 A_1 B_0 c_0 \land d_0 \right]
\]

\[
\left[ A_2 (A_1 B_1) (A_0 B_0) c_0 \land d_0 \right] \quad \left[ B_2 (A_1 B_1) (A_0 B_0) c_0 \land d_0 \right]
\]

Theorem (Statman '78)

Every proof of $G_k$ in Gentzen system has size $O(2^k)$. 

\[ t \]
\[ [(\bar{c}_0 \bar{d}_0) c_0 d_0] = G_0 \]
\[ 2x_i \]
\[ [(c_1 d_1) (\bar{c}_i \bar{d}_i)] \bar{c}_0 [(c_1 d_1) (\bar{c}_i \bar{d}_i)] \bar{d}_0 c_0 d_0 \]
\[ [(\bar{c}_i \bar{d}_i) (\bar{c}_i \bar{d}_i)] \]
\[ [(c_1 d_1) \bar{c}_0 [c_1 d_1] \bar{d}_0] c_0 d_0 \]
\[ 1xc_i \]
\[ [(\bar{c}_i \bar{d}_i)] \]
\[ ([c_1 d_1] \bar{c}_0 [c_1 d_1] \bar{d}_0) c_0 d_0 \]
\[ = G_1 \]
\[ 4x_i \]
\[ [[[c_2 d_2] (\bar{c}_2 \bar{d}_2)] \bar{c}_1 [[[c_2 d_2] (\bar{c}_2 \bar{d}_2)] \bar{d}_1] \]
\[ ([c_1 d_1] \bar{c}_0 [[[c_2 d_2] (\bar{c}_2 \bar{d}_2)] [c_1 d_1] \bar{d}_0] c_0 d_0 \]
\[ 4xs \]
\[ (\bar{c}_2 \bar{d}_2) (\bar{c}_2 \bar{d}_2) (\bar{c}_2 \bar{d}_2) (\bar{c}_2 \bar{d}_2) \]
\[ ([c_2 d_2] \bar{c}_1 [c_2 d_2] \bar{d}_1) \]
\[ ([c_2 d_2] [c_1 d_1] \bar{c}_0 [c_2 d_2] [c_1 d_1] \bar{d}_0) c_0 d_0 \]
\[ 3xc_i \]
\[ (\bar{c}_2 \bar{d}_2) \]
\[ ([c_2 d_2] \bar{c}_1 [c_2 d_2] \bar{d}_1) \]
\[ ([c_2 d_2] [c_1 d_1] \bar{c}_0 [c_2 d_2] [c_1 d_1] \bar{d}_0) c_0 d_0 \]
\[ = G_2 \]

Observation: Given \( G_k \), the size of the proof is \( O(k^2) \)
Frege systems vs. CoS

- Frege systems: requirement implicationally complete
- Robustness: all Frege systems over the same language mutually \( \equiv \)-simulate each other (\( \equiv \)-equivalent).

Axioms:

\[
F_1 \equiv A \rightarrow (B \rightarrow (A \land B))
\]

\[
F_2 \equiv (A \land B) \rightarrow A
\]

\[
F_3 \equiv (A \land B) \rightarrow B
\]

\[
F_4 \equiv A \rightarrow (A \lor B)
\]

\[
F_5 \equiv B \rightarrow (A \lor B)
\]

\[
F_6 \equiv \forall A \rightarrow A
\]

\[
F_7 \equiv A \rightarrow \forall A
\]

\[
F_8 \equiv A \rightarrow (B \rightarrow A)
\]

\[
F_9 \equiv \forall A \rightarrow (A \rightarrow B)
\]

\[
F_{10} \equiv (A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))
\]

\[
F_{11} \equiv (A \rightarrow C) \rightarrow ((B \rightarrow C) \rightarrow (A \land B) \rightarrow C)
\]

\[
F_{12} \equiv (A \rightarrow (B \rightarrow C)) \rightarrow (B \rightarrow (A \rightarrow C))
\]

\[
F_{13} \equiv (A \rightarrow B) \rightarrow (\forall B \rightarrow \forall A)
\]

\[
F_{14} \equiv f \rightarrow (A \land \forall A)
\]

\[
F_{15} \equiv (A \land \forall A) \rightarrow f
\]

\[
F_{16} \equiv t \rightarrow (A \lor \forall A)
\]

\[
F_{17} \equiv (A \lor \forall A) \rightarrow t
\]

Inference rule:  

\[
\text{mp} \quad A \quad A \rightarrow B \quad \frac{}{B}
\]
Here on Frege systems

- Frege proof systems:
  finite collection of sound inference rules, each of which is a type of $n > 0$ formulae $s.t. \ n-1$ are premises, and from these 1 conclusion follows.

  Inference rules with 0 premises are called axioms.

- Frege derivation of length $l$ with premises $\beta_1, \ldots, \beta_n$ and conclusion $\beta_i$ is a sequence of formulae $\beta_1, \ldots, \beta_i$ s.t. each $\beta_i$ either belongs to $\{\beta_1, \ldots, \beta_n\}$ or is the conclusion of an instance of an inference rule, whose premises belong to $\beta_1, \ldots, \beta_{i-1}$, where $1 \leq i \leq l$.

- Frege proof of $\beta$ is a Frege derivation with no premises and conclusion $\beta$.

- Derivations: $\Gamma$

- Size of $\Gamma$: $|\Gamma|$ number of unit/atom/vars occurrences therein.
**Translating Frege into CoS**

- Frege formulae $\alpha \rightarrow \beta$ are translated into $SKSg$ formulae $[\alpha \beta]

- The cut rule of $SKSg$ easily simulate modus ponens.

- Thm: For every Frege derivation $T$ with premises $\alpha_1, ..., \alpha_n$, $n \geq 0$, and conclusion $\beta$, there is a derivation $\Phi$ $(\alpha_1; ...; \alpha_n) \vdash_{SKSg} \beta$; if $M$ and $n$ are respectively lengths and site of $T$, then the length and size of $\Phi$ are respectively $O(M)$ and $O(n^2)$.

**Proof:**

- Frege axioms are tautologies, so each $F_i$ admits a proof $\Phi_i$ in $SKSg$, $1 \leq i \leq 17$.

$F_{10} \equiv (A \rightarrow (B \rightarrow c)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow c))$

```
\[\begin{align*}
\text{i\&} & \hfill t \\
\downarrow & \hfill [(A (B C)) [\bar{A} [\bar{B} C]]] \\
\equiv & \hfill [(A (B C)) [\bar{A} ((B t) C)] ] \\
\downarrow & \hfill [(A (B C)) [\bar{A} ([B [A \bar{A}]) C]] ] \\
\equiv & \hfill [(A (B C)) [\bar{A} [(\bar{B} [A \bar{A}]) C]] ] \\
\end{align*}\]
```

```math
\[\begin{align*}
\text{cd\&} & \hfill [(A (B C)) [(A \bar{B}) [\bar{A} A C]]] \\
\downarrow & \hfill [(A (B C)) [(A \bar{B}) [\bar{A} C]]] \\
\end{align*}\]```
By induction on the length of \( \Gamma = \beta_1 \ldots \beta_k, \beta \) we prove the existence of a derivation \( \Phi' \)

\[
\frac{\alpha_1 \ldots \alpha_n}{(\beta_1 \ldots \beta_k) \lambda \beta}
\]

- Base case \( k = 0 \)
  
  (i) if \( \beta \) is a premise then \( \Phi' = \beta \)
  
  (ii) if \( \beta \in F_i; \sigma \) for some axiom scheme \( i \) and instance \( \sigma \), then \( \Phi' = F_i; \sigma \)

- Inductive step

We have \( \Phi_k = \beta_1 \ldots \beta_k \) and

\[
\frac{\gamma_k}{(\beta_1 \ldots \lambda \beta_k) \lambda \beta}
\]

\( \gamma_k \) is the conjunction of premises of \( \Phi_k \).

Possible cases:

(i) \( \beta \) is a premise: \( \Phi' = \frac{\gamma_k}{\lambda \beta} \phi_k \lambda \beta \)

(ii) \( \beta \in F_i; \sigma \) for some \( i, \sigma \):

\[
\frac{\Phi_k}{(\beta_1 \ldots \lambda \beta_k) \lambda t}
\]

\[
(\beta_1 \ldots \lambda \beta_k) \lambda \sigma \frac{skss}{}
\]

\[
(\beta_1 \ldots \lambda \beta_k) \lambda \beta
\]
(iii) \( \beta \) is the conclusion of an instance \( \text{mp} \ \frac{\beta_{k'}, \beta_{k''}}{\beta} \)

where \( \beta_{k''} \equiv \beta_{k'} \to \beta \) and \( 1 \leq k', k'' \leq k \):

\[
\gamma_k
\]
\[
\Phi_k \parallel \text{skel}
\]
\[
c \uparrow
\]
\[
(\beta_{1} \land \ldots \land \beta_{k''} \land \ldots \land \beta_{k})
\]
\[
e \uparrow
\]
\[
(\beta_{1} \land \ldots \land (\beta_{k'} \land \beta_{k''}) \land \ldots \land \beta_{k})
\]
\[
= \uparrow
\]
\[
(\beta_{1} \land \ldots \land \beta_{k'}) \land (\beta_{k'} [\overline{\beta_{k'}}, \nu \beta])
\]
\[
s \uparrow
\]
\[
(\beta_{1} \land \ldots \land \beta_{k'}) \land [\overline{\beta_{k'}}, \nu \beta]
\]
\[
= \uparrow
\]
\[
(\beta_{1} \land \ldots \land \beta_{k'}) \land [\overline{\nu \beta}]
\]
\[
= \uparrow
\]
\[
(\beta_{1} \land \ldots \land \beta_{k'}) \land \beta
\]

(wlog, \( k' < k'' \))

- At every inductive step the length of \( \Phi' \) is increased by \( O(1) \)
  inference steps.
  \( \text{skel} \)
- From \( \Phi' \parallel \) obtain \( \overline{\Phi} \parallel \) by applying \( \text{mp} \).
- So, length \( \overline{\Phi} \) is \( O(k) \)
- \( |\Phi| \in O(k^2 m) \), where \( m \) is the max size of the formula in \( \Gamma \); so \( |\Phi| \in O(n^2) \) where \( |\Gamma| = n \)
From Cos to Frege

- It requires many more technicalities, because we need to simulate
- deep inference
- the amount of =

- Sketch of the argument:

1. In sks: \[ \alpha \rightarrow \Gamma \beta, \quad \forall \alpha, \beta, \forall \{\Gamma\} \]

   Correspondingly, in Frege: there is a derivation

   \[ (\alpha \rightarrow \beta, \quad \vdots, \quad \forall \{\alpha\} \rightarrow \forall \{\beta\}) \]

   whose length is \(O(m)\), size is \(O(n^2)\) where

   \[ m = |\forall \{\Gamma\}|, \quad n = |\forall \{\alpha\} \rightarrow \forall \{\beta\}| \]

2. In sks: \( \alpha = \beta \)

   Corresponding in Frege there is a derivation with

   premises \( \alpha \), conclusion \( \beta \), length \(O(n^3)\), size \(O(n^4)\)

   where \( n = |\alpha| + |\beta| \)
For every inference step \( \frac{\alpha}{\beta} \) where \( \nu \) is a rule of \( \text{SkSg} \), there is a \( \text{FREGE} \) derivation with
premise \( \alpha \), conclusion \( \beta \), length \( O(n) \), size \( O(n^2) \) where \( n = |\alpha| + |\beta| \).

- Each inference rule in \( \text{SkSg} \) \( \frac{R}{T} \) is turned into a tautology \( R \rightarrow T \), so it requires a constant-size proof in \( \text{FREGE} \).
  But this can be in context: use lemma 41.

**Theorem**: For every derivation \( \Phi \models_{\text{SkSg}} \Phi \) there is a
\( \text{FREGE} \) derivation \( \Gamma \), with
premise \( \alpha \), conclusion \( \beta \). If \( |\Phi| = n \) then
the length and size of \( \Gamma \) are respectively
\( O(n^4) \) and \( O(n^5) \)

**Conclusion**

\[ \text{SkS} \quad \text{SkSG} \quad \text{\( \text{p-simulate FREGE} \)} \]

\[ \text{FREGE} \quad \text{\( \text{p-simulate SKS} \quad \text{SKSG} \)} \]