"Introduction to Deep Inference and Proof Nets"

P. Bruscoli & L. Straßburger

(U. Bath) (INRIA - Futurs)

$\boxed{\text{LECTURE 2: Complexity}}$

Content:

- Quick background on proof complexity
- sequent calculus vs. deep inference
  - analitic 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 Co-NP = NP

  if NO then Co-NP ≠ NP ; hence P ≠ NP.

More background on proof Complexity

- Study the relative strength of proof systems

- p-simulation:

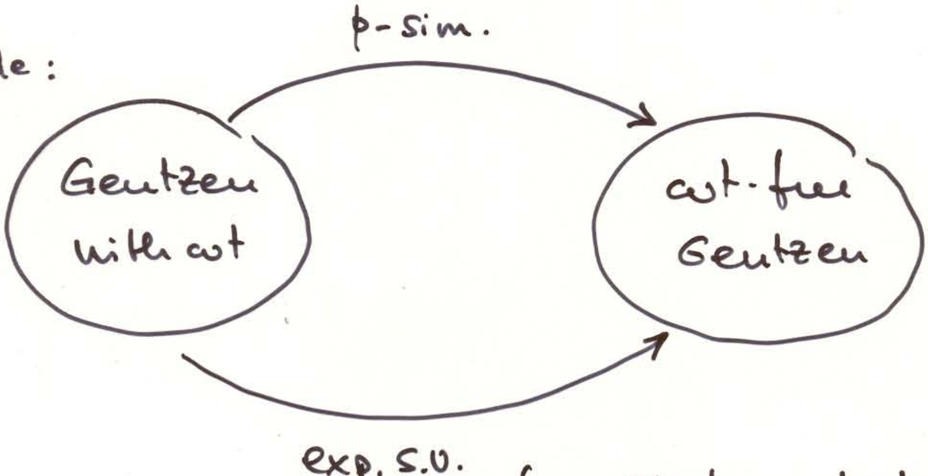
  Proof system U p-simulates proof system V iff proofs in V can be converted into proofs in U efficiently, i.e. in polynomial time (on their size)

  $U \longrightarrow V$

- 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:



(over Statman's tautologies, pigeon-hole-principle ....)

- p-equivalence

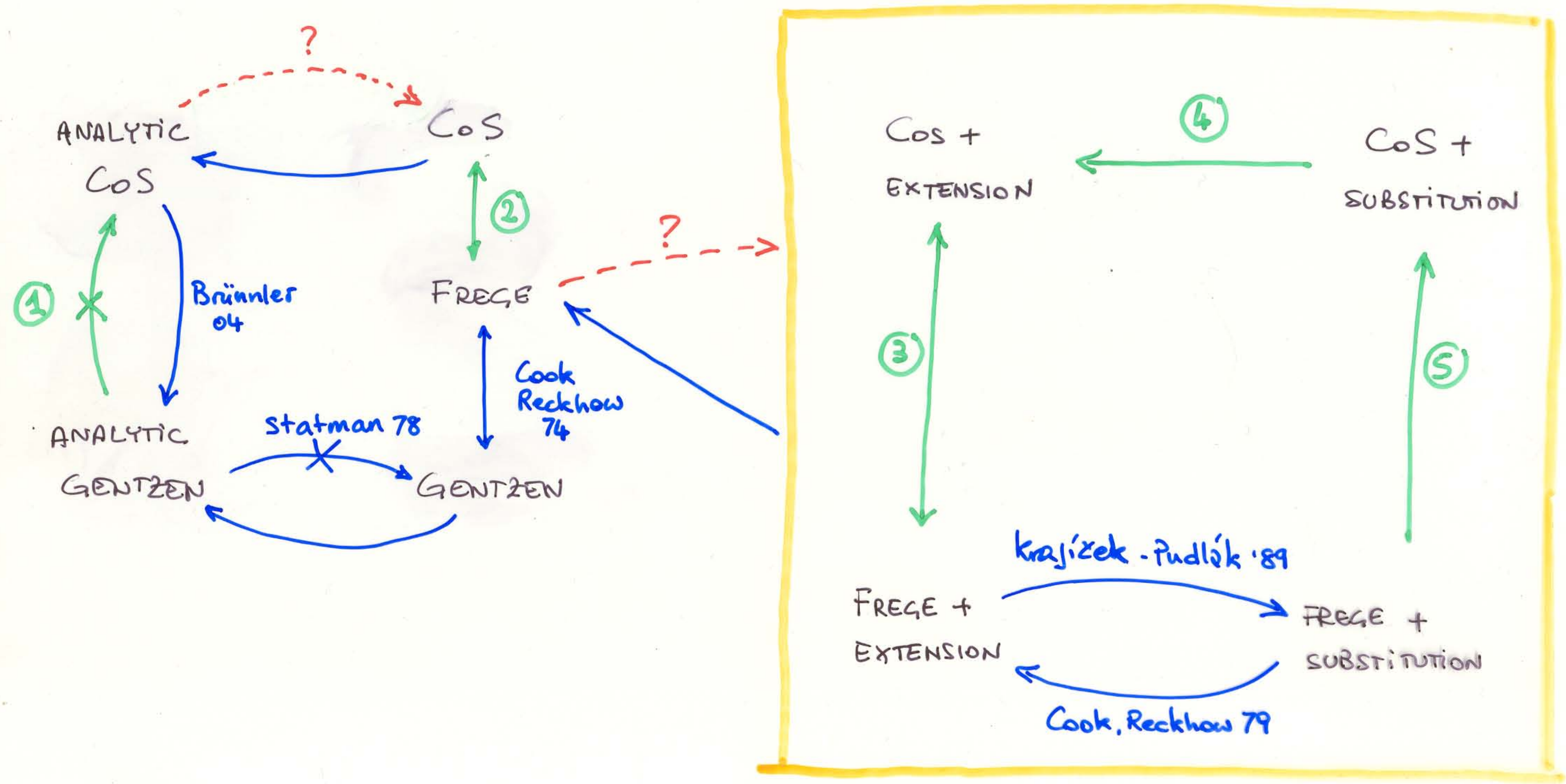
  p-simulation in both directions.

- size of a formula / derivation: the number of units/atom/vars occurrences in there. $|\alpha|$

Legenda: CoS : calculus of Structures (deep inference)

$\mathcal{F} \longrightarrow \mathcal{G}$: proof system $\mathcal{F}$ polynomially simulate proof system $\mathcal{G}$

$\mathcal{F} \not\longrightarrow \mathcal{G}$: it doesn't happen.



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 α and β 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.

$$[b \vee d] \vee [c \vee a] \rightsquigarrow$$

$$[c \vee a] \vee [b \vee d] \rightsquigarrow$$

$$[a \vee c] \vee [b \vee d] \rightsquigarrow$$

$$[a \vee [c \vee [b \vee d]]] .$$

If the size of the formula is n, repeat the process n times.

- Derivations in CoS: (modify)

$$\nu_1 \frac{\alpha_0}{\alpha_1}$$
$$\vdots$$
$$\nu_k \frac{\alpha_{k-1}}{\alpha_k}$$

where ν_i alternate between $=$ and any rule of the system.

- Recall:

A CoS proof system is implicationally complete if for every valid implication $\alpha \to \beta$ there is a derivation with premiss α and conclusion β.

SKS, SKSg are implicationally complete.

- What is the relative complexity of proofs in SKS wrt those in SKSg?

Remember the transformation from

c↓ to {ac↓, m}
w↓ to {aw↓, s}
i↓ to {ai↓, s}
.....

They are all quadratic:

- Thm: KS and KSg are p-equivalent

SKS and SKSg are p-equivalent

CoS vs Sequent Calculus (Gentzen systems)

- (Robustness thm) All systems in Sequent calculus are p-equivalent, pairwise.

- We have seen how to translate $GS \wedge p$: what complexity? (Quadratic, of course)

Thm: For every Gentzen derivation Δ with premisses $\phi_1 \dots \phi_h$ and conclusion ψ there is a derivation Φ:

$$\begin{array}{c} (\varphi_1 \dots \varphi_h) \\ \Phi \Big\| \, SKSg \\ \psi \end{array} ;$$

if n is the size of Δ, then

$O(n^2)$ is the size of Φ.

Moreover,
if Δ is analytic (cut-free) then Φ is in kSg

- Cor: SKSg p-simulates Gentzen

  kSg p-simulates analytic Gentzen

Analytic Gentzen vs Analytic CoS

- Statman tautologies

  - formulae 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

$$G_0 = (c_0 \vee d_0)$$
$$\Rightarrow (c_0 \vee d_0)$$

$$G_1 = (\ (c_1 \vee d_1)$$
$$\wedge\ (((c_1 \vee d_1) \Rightarrow c_0)$$
$$\vee ((c_1 \vee d_1) \Rightarrow d_0)))$$
$$\Rightarrow (c_0 \vee d_0)$$

$$G_2 = ((c_2 \vee d_2)$$
$$\wedge\ (((c_2 \vee d_2) \Rightarrow c_1)$$
$$\vee (c_2 \vee d_2) \Rightarrow d_1))$$
$$\wedge\ ((((c_2 \vee d_2) \wedge (c_1 \vee d_1)) \Rightarrow c_0)$$
$$\vee (((c_2 \vee d_2) \wedge (c_1 \vee d_1)) \Rightarrow d_0)))$$
$$\Rightarrow (c_0 \vee d_0)$$

...

statman) formulae grow linearly, their proofs grow exponentially

Statman's tautologies

$$G_0 = (c_0 \vee d_0) \Rightarrow (c_0 \vee d_0) \quad \overset{\text{de Morgan}}{\Longleftrightarrow} \quad [(\bar{c}_0\ \bar{d}_0)\ c_0\ d_0]$$

$$G_1 = \big((c_1 \vee d_1) \wedge (((c_1 \vee d_1) \Rightarrow c_0) \vee ((c_1 \vee d_1) \Rightarrow d_0))\big) \Rightarrow (c_0 \vee d_0) \quad \Longleftrightarrow \quad [(\bar{c}_1\ \bar{d}_1)\ ([c_1\ d_1]\ \bar{c}_0\ [c_1\ d_1]\ \bar{d}_0)\ c_0\ d_0]$$

$$G_2 = \Big((c_2 \vee d_2) \wedge (((c_2 \vee d_2) \Rightarrow c_1) \vee ((c_2 \vee d_2) \Rightarrow d_1)) \wedge ((((c_2 \vee d_2) \wedge (c_1 \vee d_1)) \Rightarrow c_0) \vee (((c_2 \vee d_2) \wedge (c_1 \vee d_1)) \Rightarrow d_0))\Big) \Rightarrow (c_0 \vee d_0)$$

$$\Longleftrightarrow \quad [(\bar{c}_2\ \bar{d}_2)\ ([c_2\ d_2]\ \bar{c}_1\ [c_2\ d_2]\ \bar{d}_1)\ ([c_2\ d_2][c_1\ d_1]\ \bar{c}_0\ [c_2\ d_2][c_1\ d_1]\ \bar{d}_0)\ c_0\ d_0]$$

...

notation: $[a\ b\ c] = a \vee b \vee c$
$(a\ b\ c) = a \wedge b \wedge c$
$\bar{a} = \neg a$

(Statman) formulae grow linearly, their proofs grow exponentially

Gentzen system (one sided)

$$\mathsf{id}\ \frac{}{[A\ \bar{A}]} \qquad \wedge\ \frac{[A\ C] \quad [B\ C]}{[(A\ B)\ C]} \qquad \vee\ \frac{A}{[A\ B]} \qquad c\ \frac{[A\ B\ B]}{[A\ B]}$$

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

of course provable

$$\vdots$$

$$c_2 \wedge ((c_2 \vee d_2) \Rightarrow c_1) \wedge (((c_2 \vee d_2) \wedge (c_1 \vee d_1)) \Rightarrow d_0)) \Rightarrow c_0 \vee d_0$$

$$[\bar{c}_2\ ([c_2\ d_2]\ \bar{c}_1)([\bar{c}_2\ d_2]\ [c_1\ d_1]\ \bar{d}_0)\ c_0\ d_0]$$

$$\parallel$$

$$[A_2\quad A_1\quad B_0\quad c_0\ d_0]$$

$$\vdots$$

$$\frac{[A_2\ (A_1\ B_1)(A_0\ B_0)\ c_0\ d_0] \qquad [B_2\ (A_1\ B_1)(A_0\ B_0)\ c_0\ d_0]}{[(\underbrace{\bar{c}_2}_{A_2}\ \underbrace{\bar{d}_2}_{\wedge\, B_2})\ (\underbrace{[c_2\ d_2]\ \bar{c}_1}_{A_1}\ \ \underbrace{[c_2\ d_2]\ \bar{d}_1}_{\wedge\, B_1})\ (\underbrace{[c_2\ d_2][c_1\ d_1]\ \bar{c}_0}_{A_0}\ \ \underbrace{[c_2\ d_2][c_1\ d_1]\ \bar{d}_0}_{\wedge\, B_0})\ c_0\ d_0]}$$

Theorem (Statman '78)

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

$$
\begin{array}{lll}
 & t & \\ \cline{2-2}
\mathsf{i}\!\downarrow & [(\bar c_0\ \bar d_0)\ c_0\ d_0] & = G_0 \\ \cline{2-2}
2\times\mathsf{i}\!\downarrow & [([[c_1\ d_1]\ (\bar c_1\ \bar d_1)]\ \bar c_0\ [[c_1\ d_1]\ (\bar c_1\ \bar d_1)]\ \bar d_0)\ c_0\ d_0] & \\ \cline{2-2}
2\times\mathsf{s} & [(\bar c_1\ \bar d_1)\ (\bar c_1\ \bar d_1) & \\
 & \quad ([c_1\ d_1]\ \bar c_0\ [c_1\ d_1]\ \bar d_0)\ c_0\ d_0] & \\ \cline{2-2}
1\times\mathsf{c}\!\downarrow & [(\bar c_1\ \bar d_1) & \\
 & \quad ([c_1\ d_1]\ \bar c_0\ [c_1\ d_1]\ \bar d_0)\ c_0\ d_0] & = G_1 \\ \cline{2-2}
4\times\mathsf{i}\!\downarrow & [([[c_2\ d_2]\ (\bar c_2\ \bar d_2)]\ \bar c_1\ [[c_2\ d_2]\ (\bar c_2\ \bar d_2)]\ \bar d_1) & \\
 & \quad ([[c_2\ d_2]\ (\bar c_2\ \bar d_2)][c_1\ d_1]\ \bar c_0\ [[c_2\ d_2]\ (\bar c_2\ \bar d_2)][c_1\ d_1]\ \bar d_0)\ c_0\ d_0] & \\ \cline{2-2}
4\times\mathsf{s} & [(\bar c_2\ \bar d_2)\ (\bar c_2\ \bar d_2)\ (\bar c_2\ \bar d_2)\ (\bar c_2\ \bar d_2) & \\
 & \quad ([c_2\ d_2]\ \bar c_1\ [c_2\ d_2]\ \bar d_1) & \\
 & \quad ([c_2\ d_2][c_1\ d_1]\ \bar c_0\ [c_2\ d_2][c_1\ d_1]\ \bar d_0)\ c_0\ d_0] & \\ \cline{2-2}
3\times\mathsf{c}\!\downarrow & [(\bar c_2\ \bar d_2) & \\
 & \quad ([c_2\ d_2]\ \bar c_1\ [c_2\ d_2]\ \bar d_1) & = G_2 \\
 & \quad ([c_2\ d_2][c_1\ d_1]\ \bar c_0\ [c_2\ d_2][c_1\ d_1]\ \bar d_0)\ c_0\ d_0] &
\end{array}
$$

Observation: Given G_k the size of the proof is $O(k^3)$

Frege systems vs. CoS

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

Axioms:

$$F_1 \equiv A \to (B \to (A \wedge B))$$

$$F_2 \equiv (A \wedge B) \to A$$

$$F_3 \equiv (A \wedge B) \to B$$

$$F_4 \equiv A \to [A \vee B]$$

$$F_5 \equiv B \to [A \vee B]$$

$$F_6 \equiv \neg\neg A \to A$$

$$F_7 \equiv A \to \neg\neg A$$

$$F_8 \equiv A \to (B \to A)$$

$$F_9 \equiv \neg A \to (A \to B)$$

$$F_{10} \equiv (A \to (B \to C)) \to ((A \to B) \to (A \to C))$$

$$F_{11} \equiv (A \to C) \to ((B \to C) \to ([A \vee B] \to C))$$

$$F_{12} \equiv (A \to (B \to C)) \to (B \to (A \to C))$$

$$F_{13} \equiv (A \to B) \to (\neg B \to \neg A)$$

$$F_{14} \equiv f \to (A \wedge \neg A)$$

$$F_{15} \equiv (A \wedge \neg A) \to f$$

$$F_{16} \equiv t \to [A \vee \neg A]$$

$$F_{17} \equiv [A \vee \neg A] \to t$$

Inference rule:

$$\mathsf{mp}\ \frac{A \qquad A \to B}{B}$$

More on Frege systems

- Frege proof system:
  finite collection of SOUND inference rules, each of which is a tuple of $n > 0$ formulae s.t. $n-1$ are premisses, and from these 1 conclusion follows.

  Inference rules with 0 premisses are called axioms
- Frege <u>derivation</u> of <u>length</u> l with premisses $\alpha_1 \dots \alpha_h$ and conclusion β_l is a <u>sequence</u> of formulae $\beta_1 \dots \beta_l$ s.t. each β_i either belongs to $\{\alpha_1 \dots \alpha_h\}$ or is the conclusion of an instance of an inference rule, whose premisses belong to $\beta_1 \dots \beta_{i-1}$, where $1 \leq i \leq l$
- Frege <u>proof</u> of β is a Frege derivation with no premisses and conclusion β.
- Derivations: Υ
- size of Υ: $|\Upsilon|$ number of unit/atom/vars occurences therein.

Translating FREGE into CoS

- FREGE formulae $\alpha \to \beta$ are translated into SKSg formulae $[\bar{\alpha}\ \beta]$
- The cut rule of SKSg easily simulate modus ponens.
- **Thm:** For every FREGE derivation Υ with premisses $\alpha_1 \dots \alpha_h$, $h \geq 0$, and conclusion β, there is a derivation Φ

$$\begin{array}{c}(\alpha_1 \dots \alpha_h)\\ \Phi \Big\Vert \mathsf{SKSg} \\ \beta\end{array}$$

; if l and n are respectively length and size of Υ, then the length and size of Φ are respectively $O(l)$ and $O(n^2)$.

Proof:

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

$$F_{10} \equiv (A \to (B \to C)) \to ((A \to B) \to (A \to C))$$

$$
\begin{array}{ll}
 & t \\
\mathsf{i}\downarrow & \overline{[(A\ (B\bar{C}))\ \ [\bar{A}\ [\bar{B}\ C]]\ ]} \\
= & \overline{[(A\ (B\bar{C}))\ \ [\bar{A}\ [(\bar{B}\ t)\ C]]\ ]} \\
\mathsf{i}\downarrow & \overline{[(A\ (B\bar{C}))\ \ [\bar{A}\ [(\bar{B}\ [A\ \bar{A}])\ C]]\ ]} \\
\mathsf{s} & \overline{[(A\ (B\bar{C}))\ \ [\bar{A}\ [[(\bar{B}A)\ \bar{A}]\ C]]\ ]} \\
= & \overline{[(A\ (B\bar{C}))\ \ [(A\bar{B})\ [[\bar{A}\ \bar{A}]\ C]]\ ]} \\
\mathsf{c}\downarrow & [(A\ (B\bar{C}))\ \ [(A\bar{B})\ [\bar{A}\ C]]\ ]
\end{array}
$$

- By induction on the length of $\gamma = \beta_1, \dots \beta_k, \beta$ we prove the existence of a derivation Φ'

$$\begin{array}{c} \alpha_1 \dots \alpha_n \\ \Phi' \Big\| \\ ((\beta_1 \wedge \dots \wedge \beta_k) \wedge \beta) \end{array}$$

- Base case $k = 0$

(i) if β is a premiss then $\Phi' = \beta$

(ii) if $\beta \equiv F_i \sigma$ for some axiom scheme i and instance σ, then $\Phi' = \Phi_i \sigma$

- Inductive step

We have $\Phi_k = \beta_1 \dots \beta_k$ and $\begin{array}{c} \gamma_k \\ \Phi'_k \Big\| \\ (\beta_1 \wedge \dots \wedge \beta_k) \end{array}$ where γ_k is the conjunction of premises of Φ_k.

Possible cases:

(i) β is a premiss: $\Phi' = \begin{array}{c} \gamma_k \wedge \beta \\ \Phi'_k \wedge \beta \Big\| \text{sksg} \\ (\beta_1 \wedge \dots \wedge \beta_k) \wedge \beta \end{array}$

(ii) $\beta \equiv F_i \sigma$ for some i, σ:

$$= \begin{array}{c} \gamma_k \\ \Phi'_k \Big\| \\ \dfrac{(\beta_1 \wedge \dots \wedge \beta_k)}{(\beta_1 \wedge \dots \wedge \beta_k) \wedge t} \\ (\beta_1 \wedge \dots \wedge \beta_k) \wedge \phi_i \sigma \Big\| \text{sksg} \\ (\beta_1 \wedge \dots \wedge \beta_k) \wedge \beta \end{array}$$

(iii) β is the conclusion of an instance mp $\dfrac{\beta_{k'} \quad \beta_{k'} \to \beta}{\beta}$

where $\beta_{k''} \equiv \beta_{k'} \to \beta$ and $1 \le k', k'' \le k$:

$$
\begin{array}{ll}
 & \gamma_k \\
 & \Phi'_k \ \Big\|\ \text{SKSg} \\
c\uparrow & \dfrac{\beta_1 \wedge \cdots \wedge \beta_{k''} \wedge \cdots \wedge \beta_k}{(\beta_1 \wedge \cdots \wedge \beta_{k'} \cdots \wedge (\beta_{k''} \wedge \beta_{k''}) \wedge \cdots \beta_k)} \\
c\uparrow & \dfrac{}{(\beta_1 \wedge \cdots \wedge (\beta_{k'}\, \beta_{k'}) \wedge \cdots (\beta_{k''}\, \beta_{k''}) \cdots \beta_k)} \\
= & \dfrac{}{(\beta_1 \wedge \cdots \wedge \beta_k) \wedge (\beta_{k'}\, [\bar{\beta}_{k'} \vee \beta])} \\
s & \dfrac{}{(\beta_1 \wedge \cdots \wedge \beta_k) \wedge [(\beta_{k'}\, \bar{\beta}_{k'}) \vee \beta]} \\
i\uparrow & \dfrac{}{(\beta_1 \wedge \cdots \wedge \beta_k) \wedge [f \vee \beta]} \\
= & \dfrac{}{(\beta_1 \wedge \cdots \wedge \beta_k) \wedge \beta}
\end{array}
$$

(wlog, $k' < k''$)

- At every inductive step the <u>length</u> of Φ' is increased by $O(1)$ inference steps.
- From $\begin{array}{c} \alpha_1 \cdots \alpha_h \\ \Phi' \ \Big\| \\ (\beta_1 \cdots \beta_k)\beta \end{array}$ obtain $\begin{array}{c} \alpha_1 \cdots \alpha_h \\ \Phi \ \Big\| \\ \beta \end{array}$ by applying $w\uparrow$.
- So, length Φ is $O(k)$
- $|\Phi| \in O(k^2 m)$, where m is the max size of the formula in Υ; so $|\Phi| \in O(n^2)$ where $|\Upsilon| = n$

From CoS to Frege

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

- Sketch of the argument:

$\boxed{1}$ In SKS:

$$\begin{array}{c} \alpha \\ \| \\ \beta \end{array} \quad \longrightarrow \quad \begin{array}{c} \xi\{\alpha\} \\ \| \\ \xi\{\beta\} \end{array} \qquad \text{for every } \alpha, \beta, \xi\{\ \}$$

Correspondingly, in FREGE: there is a derivation

$$\begin{array}{l} (\alpha \to \beta, \\ \quad \vdots \\ \xi\{\alpha\} \to \xi\{\beta\}) \end{array}$$

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

$m = |\xi\{\ \}|$ $\quad n = |\xi\{\alpha\} \to \xi\{\beta\}|$

$\boxed{2}$ In SKS: $\alpha = \beta$

Correspondingly in Frege there is a derivation with premiss α, conclusion β, length $O(n^3)$, size $O(n^4)$ where $n = |\alpha| + |\beta|$

$\boxed{3}$ For every inference step $\nu \dfrac{\alpha}{\beta}$ where ν is a rule of SKSg, there is a FREGE derivation with premiss α, conclusion β, length $O(n)$, size $O(n^2)$ where $n = |\alpha| + |\beta|$.

- each inference rule in SKSg $\frac{R}{T}$ is turned into a tautology $R \to T$, so it requires a constant-size proof in Frege.

But this can be in context : use lemma $\boxed{1}$.

$\boxed{4}$ THM For every derivation $\begin{array}{c}\alpha\\ \Phi \Big\| \text{SKSg}\\ \beta\end{array}$ there is a FREGE derivation Ψ, with premiss α, conclusion β. If $|\Phi| = n$ then the length and size of Ψ are respectively $O(n^4)$ and $O(n^5)$

CONCLUSION

SKS SKSg	p-simulate FREGE
FREGE	p-simulate SKS SKSg