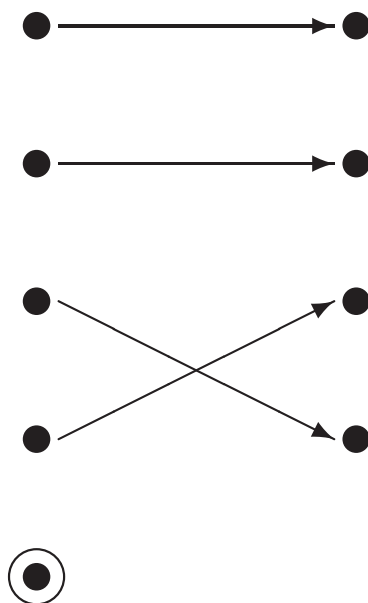


Theorem. Any **RR** of PHP_{n-1}^n requires at least $2^{n/20}$ clauses.

Proof.

A truth assignment (ta) σ is *i-critical* if σ leaves pigeon i out, and maps the remaining $(n - 1)$ pigeons to holes bijectively.



Two clauses C_1, C_2 are **equivalent wrt critical ta's** if $\sigma \models C_1 \iff \sigma \models C_2$, for all critical σ .

A **R** inference $C_1, C_2 \vdash C_3$ is **sound wrt critical ta's** if whenever a critical σ satisfies C_1, C_2 , it also satisfies C_3 .

Consider a **RR** \mathcal{R} of PHP_{n-1}^n , and in every clause of \mathcal{R} replace \bar{P}_{ik} by the literals $\{P_{lk} \mid l \neq i\}$, obtaining thus $\hat{\mathcal{R}}$.

All clauses in $\hat{\mathcal{R}}$ are **monotone**, equivalent to the corresponding original clauses wrt critical ta's, and all inferences in $\hat{\mathcal{R}}$ are sound wrt critical ta's. *But $\hat{\mathcal{R}}$ is not a **RR** per se !*

Define a clause to be **large** if it contains at least $n^2/10$ -many variables.

Let $S =$ *the number of large clauses in $\hat{\mathcal{R}}$.*

Claim. There is a P_{ij} contained in at least $S/10$ -many large clauses of $\hat{\mathcal{R}}$.

Proof of Claim. There are $< n^2$ variables, and suppose that each of them is in less than $S/10$ many large clauses. Now what is the largest collection of clauses we can form where each clause is required to be large and we have $< S/10$ “copies” of each of the n^2 variables?

$$< \frac{n^2 \cdot S/10}{n^2/10} = S$$

Contradiction, since we *do* have S large clauses by assumption.

We derive a contradiction from $S < 2^{n/20}$.

Choose such a P_{ij} , set it to 1, and set to 0 all P_{il} and $P_{l'j}$, where $l \neq j$ and $l' \neq i$.

Apply restriction to $\hat{\mathcal{R}}$, to obtain a monotone refutation $\hat{\mathcal{R}}'$ of PHP_{n-2}^{n-1} .

Technical point:

$$\begin{array}{ccc}
 \mathcal{R} \text{ for } \text{PHP}_{n-1}^n & \xrightarrow{\bar{P}_{ik} \rightsquigarrow \{P_{lk} | l \neq i\}} & \hat{\mathcal{R}} \text{ for } \text{PHP}_{n-1}^n \\
 \downarrow & & \downarrow P_{ij}=1, \forall l \neq j, l' \neq i, P_{il}=0, P_{l'j}=0 \\
 \mathcal{R}' \text{ for } \text{PHP}_{n-2}^{n-1} & \longrightarrow & \hat{\mathcal{R}}' \text{ for } \text{PHP}_{n-2}^{n-1}
 \end{array}$$

i.e., *commutative diagram*.

The number of large clauses in $\hat{\mathcal{R}}'$ for PHP_{n-2}^{n-1} is at most $9S/10$.

Repeat this $\log_{10/9} S$ many times until we knock out all large clauses.

So we end up with a monotone **RR** of $\text{PHP}_{n'-1}^{n'}$ where

$$n' \geq n - \log_{10/9} S > n(1 - (\log_{10/9} 2)/20) > 0.671n$$

and where there are no large (i.e., of size $\geq n^2/10$) clauses^a.

Lemma. Any $\hat{\mathcal{R}}$ for PHP_{n-1}^n must have a clause with at least $2n^2/9$ literals (which by definition are all positive variables!).

Contradiction!

$$2(n')^2/9 > n^2/10$$

^aActually, after $\log_{10/9} S$ many times you may still have one more large clause left, so really $n' > 0.671n - 1$. But the argument can still be made to work by taking n sufficiently large.

Proof of Lemma. For each C in $\hat{\mathcal{R}}$, let $\text{Complex}(C)$ be the minimum number of clauses in PHP_{n-1}^n that imply C on all critical ta's.

Note that only “pigeon” clauses $\{P_{i1}, P_{i2}, \dots, P_{i(n-1)}\}$ are included in a minimal set implying C (as we restrict ourselves to critical ta's and clauses $\{\bar{P}_{ik}, \bar{P}_{jk}\}$ are always satisfied by critical ta's).

The complexity of “pigeon” clauses is 1, and of \square is n .

If $C_1, C_2 \vdash C_3$, then $\text{Complex}(C_3) \leq \text{Complex}(C_1) + \text{Complex}(C_2)$.

Claim. $\exists C, n/3 < \text{Complex}(C) \leq 2n/3$.

Proof of Claim. Take C to be a clause of complexity $> n/3$ (such a clause exists as $\text{Complex}(\square) = n$). If both parents are of complexity $\leq n/3$, C is our clause. Otherwise, one parent is of complexity $> n/3$. Take that parent, and repeat. This process must end since initial clauses are of complexity 0 or 1.

Let P be a subset of “pigeon” clauses that minimally implies C , $|P| = m = \text{Complex}(C)$.

Claim. $|C| \geq (n - m)m$.

Note that $(n - m)m \geq 2n^2/9$, as it takes its min when at the extremes, i.e., when $m = n/3$ or $m = 2n/3$.

Proof of Claim. For $i \in P$ (short-hand for $\{P_{i_1}, \dots, P_{i_{(n-1)}}\} \in P$), consider an i -critical α such that $\alpha \not\models C$.

(If every i -critical α satisfied C , we would not need $i \in P$, since i is needed in P only if some i -critical α falsifies C .)

For each $j \notin P$ (remember $\text{Complex}(C) \leq 2n/3$), let α' be α except $\alpha'(i) = \alpha(j)$, and α' is j -critical.

$\alpha' \models C$.

Since $j \notin P$, and α' is j -critical, α' satisfies P , so it must satisfy C .

But α and α' are the same on all variables, except on P_{jl} and P_{il} .

Since $\alpha \not\models C$ but $\alpha' \models C$, it follows that $P_{il} \in C$.

Using the same α on all $j \notin P$ we get $(n - m)$ *distinct* variables $P_{il_1}, P_{il_2}, \dots, P_{il_{n-m}}$ in C .

Repeating the whole argument for each $i \in P$, gives us $(n - m)m$ variables in C .