4.8 Pigeonhole and resolution

In this section we will present a result of a different nature – we will use the pigeonhole principle to show that the principle *itself* is hard to deduce in a classical proof system, known as Resolution.

4.8.1 Resolution refutation proofs

The resolution proof system was introduced by Blake (1937) and has been made popular as a theorem-proving technique by Davis and Putnam (1960) and Robinson (1965). This system operates with clauses, i.e., with Or's of literals, where each literal is either a variable x_i or its negation $\overline{x}_i$. A *truth-assignment* is an assignment of constants 0 and 1 to all the variables. Such an assignment *satisfies* (*falsifies*) a clause if it evaluates at least one (respectively, none) of its literals to 1. A set of clauses is satisfiable if there is an assignment which satisfies all its clauses.

Let F be a set of clauses and suppose that F is not satisfiable. A *resolution refutation proof* for F is a sequence of clauses $\mathcal{R} = (C_1, \ldots, C_t)$ where C_t is the empty clause (which, by definition, is satisfied by no assignment) and each intermediate clause C_i either belongs to F or is derived from some previous two clauses using the following *resolution rule*:

The clause $C \vee C'$ can be inferred from two clauses $C \vee x_i$ and $C' \vee$

In this case one also says that the variable x_i was *resolved* to derive the clause $C \vee C'$. The *length* of such a proof is equal to the number t of clauses in the derivation.

Observe that the resolution rule is *sound* in the following sense: if some assignment (of constants to all the variables) falsifies the derived clause $C \vee C'$, then it must falsify at least one of the clauses $C \vee x_i$ and $C' \vee \overline{x}_i$ from which it was derived. It is also known (and easy to show) that Resolution is *complete*: every unsatisfiable set of clauses has a resolution refutation proof.

What about the length of such derivations? Due to its practical importance, this question bothered complexity theoreticians and logicians for a long time.

The first exponential lower bound for the length of regular resolution was proved by Tseitin (1968) already 30 years ago. (These are resolution proofs with the additional restriction that along every path every particular variable x_i can be resolved at most once; a path in a derivation is just a sequence of clauses, each of which is one of the two hypotheses from which the next clause is derived.) However, despite its apparent simplicity, the first lower bounds for non-regular resolution were only proved in 1985 by Haken. These bounds were achieved for the set of clauses PHP_n^{n+1} formalizing the pigeonhole principle. Subsequently, Haken's argument was refined and extended to other principles as well as to proof systems generalizing Resolution.

One may also consider the generalized pigeonhole principle PHP_n^m saying that m pigeons ($m \geqslant n+1$) cannot sit in n holes so that every pigeon is alone in its hole. The larger the difference $m - n$, the "more true" is the principle itself, and its proof might be shorter. Buss and Pitassi (1998) have proved that, for $m \geqslant 2^{\sqrt{n \log n}}$, PHP_n^m has a resolution proof of length polynomial in m. But for a long time, no non-trivial lower bound was known for $m > n^2$. Overcomming this "n square" barrier was one of the most challenging open problems about the power of Resolution. This problem was recently resolved by Ran Raz (2001) who proved that for any $m \geqslant n+1$, any Resolution proof of PHP_n^m requires length 2^{n^ϵ}, where $\epsilon > 0$ is an absolute constant.

4.8.2 Haken's lower bound

Recall that the pigeonhole principle states that n pigeons cannot sit in $n-1$ holes so that every pigeon is alone in its hole. To formalize the principle, let us introduce boolean variables $x_{i,j}$ interpreted as:

$x_{i,j} = 1$ if and only if the ith pigeon sits in the jth hole.

Let PHP_{n-1}^n denote the set of clauses:

(i) $x_{i,1} \vee x_{i,2} \vee \cdots \vee x_{i,n-1}$ for each $i = 1, \ldots, n$;
(ii) $\overline{x}_{i,k} \vee \overline{x}_{j,k}$ for each $1 \leqslant i \neq j \leqslant n$ and $1 \leqslant k \leqslant n-1$.

Note that the And of all clauses of the first sort is satisfiable if and only if every pigeon sits in at least one hole, whereas the And of the clauses of the second sort can be satisfied if and only if no two pigeons sit in the same hole. Thus, by the pigeonhole principle(!), the And of all clauses in PHP_{n-1}^n is not satisfiable.

Theorem 4.13 (Haken 1985). *For a sufficiently large n, any Resolution proof of PHP_{n-1}^n requires length $2^{\Omega(n)}$.*

Originally, Haken's proof used the so-called "bottleneck counting" argument and was quite involved. Here we present a new and simple proof of his result found by Beame and Pitassi (1996).

Proof. We will concentrate on a particular subset of truth assignments. Look at the set of underlying variables $X = \{x_{i,j} : 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant n-1\}$ as an $n \times (n-1)$ matrix. Say that a truth assignment $\alpha: X \to \{0,1\}$ to the underlying variables $x_{i,j}$ is *critical* if it defines a one-to-one map from $n-1$ pigeons to $n-1$ holes, with the remaining pigeon not mapped to any hole. A critical assignment, where i is the left-out pigeon, is called *i-critical* (see Fig. 4.2). In what follows we will be interested only in these critical truth assignments. (How many such assignments do we have?)

Take an arbitrary resolution refutation proof $\mathcal{R} = (C_1, \ldots, C_t)$ for PHP_{n-1}^n. As a first step, we get rid of negations: we replace each clause C in $\mathcal{R}$ by a *positive* clause C^+, i.e., by a clause without negated variables.

$$\begin{array}{c|ccccc|}
 & 1 & 0 & 0 & 0 & 0 \\
 & 0 & 1 & 0 & 0 & 0 \\
 & 0 & 0 & 1 & 0 & 0 \\
i & 0 & 0 & 0 & 0 & 0 \\
 & 0 & 0 & 0 & 1 & 0 \\
 & 0 & 0 & 0 & 0 & 1
\end{array}
\qquad
\begin{array}{c|ccccc|}
 & 0 & 1 & 0 & 0 & 0 \\
 & 0 & 0 & 0 & 1 & 0 \\
 & 1 & 0 & 0 & 0 & 0 \\
i & 0 & 0 & 0 & 0 & 0 \\
 & 0 & 0 & 0 & 0 & 1 \\
 & 0 & 0 & 1 & 0 & 0
\end{array}$$

Fig. 4.2. Two i-critical truth assignments for $i = 4$ in the case of 6 pigeons and 5 holes

The idea of this transformation is due to Buss (1987) and works as follows: replace each occurrence of $\overline{x}_{i,j}$ in the clause C by the Or

$$C_{i,j} \rightleftharpoons x_{1,j} \vee \cdots \vee x_{i-1,j} \vee x_{i+1,j} \vee \cdots \vee x_{n,j}$$

of all the variables, corresponding to the jth hole, except $x_{i,j}$. The resulting sequence of positive clauses $\mathcal{R}^+ = (C_1^+, \ldots, C_t^+)$ is no longer a valid resolution refutation proof – it is just a *sequence* which we will call a *positive pseudo-proof* of PHP^n_{n-1}. For the rest of the proof it will only be important that this sequence has the property that, with respect to critical assignments, the rules in it are still sound. That is, if C is derived from C_1 and C_2 in the original proof $\mathcal{R}$ then, for every critical α,

$$C_1^+(\alpha) \cdot C_2^+(\alpha) \leqslant C^+(\alpha).$$

This is an immediate consequence of the following claim.

Claim 4.14. *For every critical truth assignment* α, $C^+(\alpha) = C(\alpha)$.

Proof. Suppose there is a critical assignment α such that $C^+(\alpha) \neq C(\alpha)$. This can only happen if C contains a literal $\overline{x}_{i,j}$ such that $\overline{x}_{i,j}(\alpha) \neq C_{i,j}(\alpha)$. But this is impossible, since α has precisely one 1 in the jth column. □

We will use this property (the soundness with respect to critical assignments) to show that the pseudo-proof $\mathcal{R}^+$ (and, hence, also the original proof $\mathcal{R}$) must be long, namely – that $t \geqslant 2^{n/32}$. For the sake of contradiction, assume that we have fewer than $2^{n/32}$ clauses in $\mathcal{R}^+$. Say that a clause is *long* if it has at least $n^2/8$ variables, i.e., if it includes more that 1/8 fraction of all $n(n-1)$ possible variables. Let ℓ be the number of long clauses in $\mathcal{R}$; hence

$$\ell < 2^{n/32}.$$

Since each long clause has at least a 1/8 fraction of all the variables, there must be (by the pigeonhole principle!) a variable $x_{i,j}$ which occurs in at least $\ell/8$ of the long clauses. Set this variable to 1, and at the same time set to 0 all the variables $x_{i,j'}$ and $x_{i',j}$ for all $j' \neq j, i' \neq i$ (see Fig. 4.3). After this setting, all the clauses containing $x_{i,j}$ will disappear from the proof (they all get the value 1) and the variables which are set to 0 will disappear from the remaining clauses.

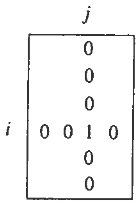

Fig. 4.3. Setting of constants to eliminate long clauses containing $x_{i,j}$

Applying this restriction to the entire proof $\mathcal{R}^+$ leaves us with a new positive pseudo-proof of PHP^{n-1}_{n-2}, where the number of long clauses is at most $\ell(1-1/8)$. Continue in this fashion until we have set all long clauses to 1. Applying this argument iteratively $d = 8\ln \ell$ many times, we are guaranteed to have knocked out all long clauses, because

$$\ell(1-1/8)^d < \mathrm{e}^{\ln \ell - d/8} = 1.$$

Thus, we are left with a positive pseudo-proof $\mathcal{R}'$ of PHP^m_{m-1}, where $m = n - 8\ln \ell$, and where *no* clause is long, i.e., has length at least $n^2/8$. But this contradicts the following claim which states that such a pseudo-proof must have a clause of size

$$2m^2/9 = 2(n - 8\ln \ell)^2/9 > 2(n - n/4)^2/9 = n^2/8.$$

So, it remains to prove the claim.

Claim 4.15. *Any positive pseudo-proof of* PHP^m_{m-1} *must have a clause with at least* $2m^2/9$ *variables.*

Proof. Let $\mathcal{R}'$ be a positive pseudo-proof of PHP^m_{m-1}. Recall that $\mathcal{R}'$ contains no negated literals and that the rules in $\mathcal{R}'$ are sound with respect to critical assignments. This implies that for every clause C in $\mathcal{R}'$ there is a set of clauses $\mathcal{W}$ from PHP^m_{m-1} whose conjunction implies C on all critical truth assignments. That is, every critical assignment satisfying all the clauses in $\mathcal{W}$ must also satisfy the clause C. We call such a set of clauses $\mathcal{W}$ a *witness* of C. One clause C may have several witnesses. We define the *weight* of C as the minimal number of clauses in its witness.

Let us make several observations about this measure. Since we are considering only critical truth assignments, only the "pigeon" clauses of type (i), saying that some pigeon must be mapped to a hole, will be included in a minimal witness, just because all other clauses are satisfied by every critical assignment (no column has two 1's). The weight of these initial "pigeon" clauses is 1, and the weight of the final clause is m (since this clause outputs 0 for *all* critical assignments). Since (by soundness) the weight of a clause is at most the sum of weights of the two clauses from which it is derived, there must exist a clause C in the proof whose weight s is between $m/3$ and $2m/3$.

(This is a standard and useful trick, and we address it in Exercise 4.14.) We will prove that this clause C must contain at least $2m^2/9$ variables.

To show this, let $\mathcal{W} = \{C_i : i \in S\}$, where $|S| = s$, be a minimal set of pigeon clauses

$$C_i = x_{i,1} \vee x_{i,2} \vee \cdots \vee x_{i,m-1}$$

in PHP^m_{m-1} whose conjunction implies C. We will show that C has at least

$$(m-s)s \geqslant 2m^2/9$$

distinct literals.

Take an $i \in S$, and let α be an i-critical truth assignment falsifying C. (Such an assignment exists by the minimality of $\mathcal{W}$; check this!) For each $j \notin S$, consider the j-critical assignment α' obtained from α by replacing i by j. This assignment differs from α only in two places: if α mapped the pigeon j to the hole k, then α' maps the pigeon i to this hole k (see Fig. 4.4).

$$\begin{array}{c|ccccc|} & & k & & & \\ \hline & 1 & 0 & 0 & 0 & 0 \\ j & 0 & \textcircled{1} & 0 & 0 & 0 \\ & 0 & 0 & 1 & 0 & 0 \\ i & 0 & \textcircled{0} & 0 & 0 & 0 \\ & 0 & 0 & 0 & 1 & 0 \\ & 0 & 0 & 0 & 0 & 1 \\ \hline \end{array} \longrightarrow \begin{array}{c|ccccc|} & & k & & & \\ \hline & 1 & 0 & 0 & 0 & 0 \\ j & 0 & \textcircled{0} & 0 & 0 & 0 \\ & 0 & 0 & 1 & 0 & 0 \\ i & 0 & \textcircled{1} & 0 & 0 & 0 \\ & 0 & 0 & 0 & 1 & 0 \\ & 0 & 0 & 0 & 0 & 1 \\ \hline \end{array}$$

Fig. 4.4. Assignment α' is obtained from α by interchanging the ith and jth rows.

Since $j \notin S$, the assignment α' satisfies all the clauses of the witness $\mathcal{W}$ of C, and hence, must satisfy the clause C. Since $C(\alpha) = 0$ and the assignments α,α' differ only in the variables $x_{i,k}$ and $x_{j,k}$, this can only happen when C contains the variable $x_{i,k}$ (remember that the clause C has no negated literals). Running the same argument over all $m-s$ pigeons $j \notin S$ (using the same α), it follows that C must contain at least $m-s$ distinct variables $x_{i,k_1}, x_{i,k_2}, \ldots, x_{i,k_{m-s}}$ corresponding to the ith pigeon. Repeating the argument for all pigeons $i \in S$ shows that C contains at least $(m-s)s$ variables, as claimed.

This completes the proof of the claim, and thus, the proof of the theorem. □

Exercises

4.1.$^-$ Suppose five points are chosen inside an equilateral triangle with side-length 1. Show that there is at least one pair of points whose distance apart is at most $1/2$. *Hint*: Divide the triangle into four suitable boxes.