A SHORT PROOF OF THE PIGEON HOLE PRINCIPLE USING EXTENDED RESOLUTION

Stephen A. Cook

For each integer $n > 1$, the pigeon hole principle can be formulated as a set of propositional clauses as follows: Introduce atoms P_{ij}, $1 \le i \le n$, $1 \le j \le n-1$, where P_{ij} says intuitively that i gets mapped to j. Then the set of clauses

$$S_n \begin{cases} P_{i1} \vee P_{i2} \vee \dots \vee P_{i,n-1} & 1 \le i \le n \\ \neg P_{ik} \vee \neg P_{jk}, & 1 \le i < j \le n, \quad 1 \le k \le n-1 \end{cases}$$

asserts intuitively that there is a one-one map from $\{1,2,\dots,n\}$ to $\{1,2,\dots,n-1\}$. Hence, by the pigeon-hole principle, this set of clauses must be inconsistent. Several years ago, Dick Karp (for one) noticed that there didn't seem to be any short (i.e. polynomial in n) resolution refutation of the set S_n, and posed the problem of trying to prove this. In fact, I believe the shortest resolution refutation known for S_n has $(n-1)(n+2)2^{n-3}$ clauses, but no one has been able to prove a non-polynomial lower bound on an arbitrary resolution refutation of S_n.

After reading Tseitin's paper [1] describing extended resolution (ER), the question arose whether there exists a short ER refutation of S_n. It turns out that such a short refutation does exist, and it is the purpose of this note to describe it and show briefly how it motivated my paper [2] on feasibly constructive proofs.

The normal proof of the pigeon hole principle proceeds by induction on n. If ϕ is a one-one map from $\{1,2,\ldots,n\}$ to $\{1,2,\ldots,n-1\}$, then we can define a one-one map ϕ' from $\{1,2,\ldots,n-1\}$ to $\{1,2,\ldots,n-2\}$ by setting

$$\phi'(i) = \begin{cases} \phi(i) \text{ if } \phi(i) \neq n-1 \\ \phi(n) \text{ if } \phi(i) = n-1 \end{cases}, \quad 1 \leq i \leq n-1$$

Proceeding in this way, we can eventually produce the absurdity of a one-one map from $\{1,2\}$ to $\{1\}$.

The ER refutation of the clauses S_n reflects the above argument very closely. We simply introduce new proposition atoms Q_{ij}, $1 \leq i \leq n-1$, $1 \leq j \leq n-2$, by the extension rule so that $\{Q_{ij}\}$ describes the map ϕ' above, assuming $\{P_{ij}\}$ describes the map ϕ. Thus we want $Q_{ij} \equiv (P_{ij} \vee (P_{i,n-1} \& P_{nj}))$. The four clauses which specify this equivalence are

(1) $Q_{ij} \vee \neg P_{ij}$

(2) $Q_{ij} \vee \neg P_{i,n-1} \vee \neg P_{nj}$

(3) $\neg Q_{ij} \vee P_{ij} \vee P_{i,n-1}$

(4) $\neg Q_{ij} \vee P_{ij} \vee P_{nj}$

(If we stick strictly to Tseitin's formulation of the extension rule, then two intermediate atoms must be introduced before the above four clauses can be derived by a few resolutions). From the $O(n^3)$ clauses (1)-(4), $1 \leq i \leq n-1$, $1 \leq j \leq n-2$, and from the clauses in S_n, one can derive the clauses

$$S_{n-1}\begin{cases} Q_{i1} \vee \dots \vee Q_{i,n-2}, & 1 \le i \le n-1 \\ \neg Q_{ik} \vee \neg Q_{jk}, & 1 \le i < j \le n-1, \quad 1 \le k \le n-2 \end{cases}$$

using $O(n^3)$ resolutions. It is an easy exercise to do this, provided one keeps in mind the intuitive reason that ϕ' is a one-one map from $\{1,2,\dots,n-1\}$ to $\{1,2,\dots,n-2\}$ given that ϕ is a one-one map from $\{1,2,\dots,n\}$ to $\{1,2,\dots,n-1\}$. Proceeding in this way n-2 times, we eventually produce the set of clauses $S_2 = \{R_{11}, R_{21}, \neg R_{11} \vee \neg R_{21}\}$ from which the empty clause can be derived in two resolutions. Thus the total length of the ER refutation of S_n is $O(n^4)$, or $O(N^{4/3})$, where N is the number of clauses in S_n.

In contemplating the above argument several years ago, I was struck by how naturally and smoothly the ER refutation followed the original mathematical proof of the pigeon hole principle. This lead to a more general question: Suppose T_n is a set of clauses for each n, and suppose one can prove informally that T_n is inconsistent for every n. Under what conditions can one follow the informal proof to find a short ER refutation for each set T_n? The key requirement of the informal proof seems to be that it be highly constructive. For example, the proof by induction of the pigeon-hole principle provides directly a means of transforming to an absurdity each potential one-one map from $\{1,2,\dots,n\}$ to $\{1,2,\dots,n-1\}$, and the whole transformation process requires time polynomial in n. It is this "feasibly constructive" property of the proof that allows simulation by extended resolution. On the other hand, one can construct a set of clauses T_n which is satisfiable if and only if there is some proof in Peano number theory of $0 = 1$ of length n or less, and the number of clauses in T_n is bounded by a polynomial in n. The consistency of Peano number theory implies the inconsistency of T_n, but none of the usual consistency proofs for number theory help in finding a polynomial bounded ER refutation of T_n.

The system PV in [2] is the most general formal system I could think of whose proofs have this feasibly constructive property. Briefly, formulas in PV are equations $f(x) = g(x)$ between functions (or terms built from such functions), where f and g are any natural number functions which can be computed in time bounded by a polynomial in the length of their arguments (in binary notation). For each fixed n, we can introduce propositional atoms $P_0^x,\ldots,P_{n-1}^x$ for the binary digits of x (assume x has length not exceeding n), and atoms $P_0^{f(x)},\ldots,P_{m-1}^{f(x)}$ and $P_0^{g(x)},\ldots,P_{m-1}^{g(x)}$ for the bits of $f(x)$ and $g(x)$, for suitable m. Furthermore, we can introduce clauses (in the spirit of the extension rule) which define the bits $P_i^{f(x)}$ and $P_i^{g(x)}$ in terms of the bits P_i^x, in the sense that any truth assignment satisfying these clauses must assign values to the $P_i^{f(x)}$ and $P_i^{g(x)}$ which give the binary notation for $f(x)$ and $g(x)$ when x has the binary notation given by the P_i^x.

It turns out that the total number of such clauses is bounded by a polynomial in n for a fixed equation f(x) = g(x). One of the main results in the paper states that if the equation f(x) = g(x) is provable in the system PV, then the above system of clauses together with the clauses stating $\neg(P_0^{f(x)} \equiv P_0^{g(x)} \,\&\, \ldots \,\&\, P_{m-1}^{f(x)} \equiv P_{m-1}^{g(x)})$ has an ER refutation bounded by a polynomial in n. This theorem is supposed to capture the intention that every feasibly constructive proof can be simulated by a short ER refutation.

The question of whether there is a polynomial p(n) such that every inconsistent set of n clauses has an ER refutation of length p(n) or less remains open. For that matter, it is not known whether any proof system (suitably defined) for the propositional calculus always has polynomially bounded proofs. This question is important, since it is equivalent to the question of whether *NP* is closed under complements.

References

[1] G. S. Tseitin, "On the complexity of derivation in propositional calculus". Studies in Mathematics and Mathematical Logic, Part II, A.O. Slisenko, ed. (Translated from Russian).

[2] S.A. Cook, "Feasibly constructive proofs and the propositional calculus". Preliminary Version. Proceedings 7th Annual ACM STOC, May, 1975, pp.83-97.