Greedy Pebbling: Towards Proof Space Compression

Andreas Fellner * and Bruno Woltzenlogel Paleo **

fellner.a@gmail.com bruno@logic.at
Theory and Logic Group
Institute for Computer Languages
Vienna University of Technology

Abstract. This paper describes algorithms and heuristics for playing a *Pebbling Game*. Playing the game with a small number of pebbles is analogous to checking a proof with a small amount of available memory. Here this analogy is exploited: new pebbling algorithms are conceived and evaluated on the task of compressing the space of thousands of propositional resolution proofs generated by SAT- and SMT-solvers.

1 Pebbling as a Satisfiability Problem

To find the pebble number of a proof, the question whether the proof can be pebbled using no more than k pebbles can be encoded as a propositional satisfiability problem. In this chapter let φ be a proof with nodes $v_1, \dots, v_n$ and let v_n be its the root node. Due to rule **??** of the Static Pebbling Game, the number of moves that pebble nodes is exactly n and due to theorem **??** determining the order of these moves is enough to define a strategy. For every $x \in \{1, \dots, k\}$, every $j \in \{1, \dots, n\}$ and every $t \in \{0, \dots, n\}$ there is a propositional variable $p_{x,j,t}$. The variable $p_{x,j,t}$ being mapped to $\top$ by a valuation is interpreted as the fact that in the t'th round of the game node v_j is marked with pebble x. Round 0 is interpreted as the initial setting of the game before any move has been done. The following constraints, combined conjunctively, are satisfiable *iff* there is a pebbling strategy σ for φ with pebbling number smaller or equal k. In case the formula is satisfiable, a pebbling strategy can be read off from any satisfying assignment.

1. The root is pebbled in the last round

$$\Psi_1 = \bigvee_{x=1}^{k} p_{x,n,n}$$

* Supported by the Google Summer of Code 2013 program.
** Supported by the Austrian Science Fund, project P24300.

2. No node is pebbled initially

$$\Psi_2 = \bigwedge_{x=1}^{k} \bigwedge_{j=1}^{n} (\neg p_{x,j,0})$$

3. A pebble can only be on one node in one round

$$\Psi_3 = \bigwedge_{x=1}^{k} \bigwedge_{j=1}^{n} \bigwedge_{t=1}^{n} \left(p_{x,j,t} \rightarrow \bigwedge_{i=1, i \neq j}^{n} \neg p_{x,i,t} \right)$$

4. For pebbling a node, its premises have to be pebbled the round before and only one node is being pebbled each round.

$$\Psi_4 = \bigwedge_{x=1}^{k} \bigwedge_{j=1}^{n} \bigwedge_{t=1}^{n} \Bigg(\left(\neg p_{x,j,t} \wedge p_{x,j,(t+1)}\right) \rightarrow \\ \left(\bigwedge_{i \in P_j^{\varphi}} \bigvee_{y=1, y \neq x}^{k} p_{y,i,t} \right) \wedge \left(\bigwedge_{i=1}^{n} \bigwedge_{y=1, y \neq x}^{k} \neg \left(\neg p_{y,i,t} \wedge p_{y,i,(t+1)}\right) \right) \Bigg)$$

The sets A_{φ} and P_j^{φ} are interpreted as sets of indices of the respective nodes. This encoding is polynomial, both in n and k. However constraint 4 accounts to $O(n^3 * k^2)$ clauses. Even small resolution proofs have more than 1000 nodes and pebble numbers bigger than 100, which adds up to 10^{13} clauses for constraint 4 alone. Therefore, although theoretically possible to play the pebbling game via SAT-solving, this is practically infeasible for compressing proof space. The following theorem proves the correctness of the encoding.

Theorem 1 (Correctness of SAT encoding of pebbling).
$\Psi = \Psi_1 \wedge \Psi_2 \wedge \Psi_3 \wedge \Psi_4$ is satisfiable iff *there exists a pebbling strategy using no more than k pebbles*

Proof. Suppose Ψ is satisfiable and let $\mathcal{I}$ be a satisfying variable assignment interpreted as the set of true variables. We will use $P(x, j, t)$ as an abbreviation for $p_{x,j,(t-1)} \notin \mathcal{I}$ and $p_{x,j,t} \in \mathcal{I}$. Since $\mathcal{I}$ satisfies Ψ_3, in $P(x, j, t)$ x is uniquely defined by j and t and we can write $P(j, t)$ instead. First we will prove the following assertion. For every $t \in \{1, \ldots, n\}$ there exists exactly one $j \in \{1, \ldots, n\}$ such that $P(j, t)$. Ψ_1 states that the root v_n has to be pebbled in the last round and Ψ_2 states that no node is pebbled initially. So for n there has to be a $t \in \{1, \ldots, n\}$ such that $P(n, t)$. $\mathcal{I}$ satisfies Ψ_4, therefore for every predecessor of v_j of v_n there exists $x \in \{1, \ldots, k\}$ such that $p_{x,j,(t-1)}$. Using the same argument for v_j like for v_n there has to be a $t' \in \{1, \ldots, (t-1)\}$ such that $P(j, t')$. Every node of the proof is a recursive ancestor of the root, therefore for every $j \in \{1, \ldots, n\}$ there exists at least one $t \in \{1, \ldots, n\}$ such that $P(n, t)$. For every $t \in \{1, \ldots, n\}$ Ψ_4 ensures that if $P(n, t)$ is true $\mathcal{I}$ then there is no $i \in \{1, \ldots, n\}, i \neq j$ such that $P(i, t)$, which proves the assertion. The assertion implies the existence of a bijection $\tau : \{1, \ldots, n\} \to \{v_1, \ldots, v_n\}$ such that $\tau(n) = v_n$ and $\tau(t) = j$ *iff* $P(j, t)$. Therefore $\sigma := \{\tau(1), \ldots, \tau(n)\}$ is well defined. σ is a pebbling strategy, because $\tau(n) = v_n$, rule **??** is obeyed because of Ψ_4, rule **??** is obeyed, because unpebbling moves are given implicit (see Theorem **??**) and rule **??** is obeyed because τ is a bijection. Ψ_3 being satisfied ensures that σ uses no more than k pebbles.

Suppose there is a pebbling strategy σ using no more than k pebbles. Let the function free : $\{1, \ldots, n\} \to 2^{\{1,\ldots,k\}} \setminus \emptyset$ be defined recursively as follows and $\mathrm{peb}(t) = \min(\mathrm{free}(t))$.

$$\mathrm{free}(t) = \begin{cases} 1 & : t = 1 \\ \mathrm{free}(t-1) \setminus \{\mathrm{peb}(t-1)\} \quad \cup & : \textit{otherwise} \\ \Big\{ \mathrm{peb}(s) \mid \sigma_s \in P^{\varphi}_{\sigma_{t-1}}, s \in \{1, \ldots, t-2\} \text{ and for all } v \in C^{\varphi}_{\sigma_s} & \\ \ \text{there exists } r \in \{1, \ldots, t-1\} : \sigma_r = v \Big\} & \end{cases}$$

Intuitively, free(.) keeps track of the unused pebbles in each round. If a pebble is placed on a node, it is not free anymore. Pebbles are made free again by the implicit unpebbling moves, which correspond to the second set in the recursive definition of free(.). Since σ uses no more than k pebbles, free(.) is well defined.

Let $\mathcal{I}$ be a set of variables of Ψ defined as follows. $p_{x,j,t} \in \mathcal{I}$ *iff* $t > 0$ and there exists $s \in \{1, \ldots, t\}$ such that $\mathrm{peb}(s) = x$, $\sigma_s = v_j$ and for all $r \in \{s+1, \ldots, t\}$: $x \notin \mathrm{free}(r)$.

$\mathcal{I}$ is a satisfying assignment for Ψ. Ψ_1 is satisfied, because $\sigma_n = v_n$, therefore trivially $p_{\mathrm{peb}(n),n,n} \in \mathcal{I}$. Clearly Ψ_2 is satisfied by $\mathcal{I}$ as no variables with $t = 0$ are included in $\mathcal{I}$. To see that Ψ_3 is satisfied, suppose there exist x, t, i, j such that $i \neq j$ and $\{p_{x,j,t}, p_{x,i,t}\} \subseteq \mathcal{I}$. Then by definition of $\mathcal{I}$ there exist unique t_1 and t_2 such that $\mathrm{peb}(t_1) = x, \sigma_{t_1} = v_j$ and $\mathrm{peb}(t_2) = x, \sigma_{t_2} = v_i$. From $i \neq j$ follows $v_i \neq v_j$, therefore $t_1 \neq t_2$ w.l.o.g. suppose $t_1 > t_2$. From $\mathrm{peb}(t_2) = x$, $p_{x,i,t} \in \mathcal{I}$ and $t \geq t_1 > t_2$ follows $x \notin \mathrm{free}(t_1)$, which is a contradiction to $\mathrm{peb}(t_1) = x$. Let $P(x, j, t)$ be defined as above. Then from $P(x, j, t)$ follows $\mathrm{peb}(t) = x$ and $\sigma_t = v_j$. Rule **??** of the static pebbling game ensures that there if $P(x, j, t)$ is true, then there exists a $y\{1, \ldots, k\}\setminus\{x\}$ such that $p_{y,i,t-1} \in \mathcal{I}$. Suppose $P(x, j, t)$ and $P(y, i, t)$ both hold for some t, $x \neq y$ and $i \neq j$, then $y = \mathrm{peb}(t) = x$ and $v_j = \sigma_t = v_i$ are both contradictions. Therefore also Ψ_4 is satisfied by $\mathcal{I}$. □

2 Conclusions

References