Lecture 9— Pebbling tautolgies and space-length trade-offs

Massimo Lauria — lauria.massimo@gmail.com
Office 1107, Ookayama West 8th Building

Tuesday — November 24th, 2015 (This document was updated on June 21, 2017)

We use the pebbling tautologies to understand the relation between space and length in resolution. We first introduce the black and black-white pebbling games that model the space cost in deterministic and non deterministic computation; then we describe two variants of pebbling tautologies.

In the simplest variant, the cost of pebbling roughly corresponds to the literal space of the refutation of pebbling tautologies. In the more complex variant the cost of pebbling corresponds to the clause space of the refutation of pebbling tautologies.

Therefore trade-off in pebbling games translate to trade-off in refutations for pebbling tautologies.

In this lecture we consider the space of refutations in the blackboard model. We will focus on a particular set of theorems and results based on pebbling games. Jakob Norström is the major expert on this theory. He authored two surveys: a work in progress one about pebbling games[1], and one about the use of these games to obtain space lower bounds and trade-offs in proof complexity[2]. In this lecture notes we won't give full bibliographic reference for everything that is told. For more specific pointer please look at these surveys.

[1] Jakob Nordström. New wine into old wineskins: A survey of some pebbling classics with supplemental results. Manuscript in preparation., 2015

[2] Jakob Nordström. Pebble games, proof complexity and time-space trade-offs. *Logical Methods in Computer Science*, 9:15:1–15:63, September 2013

This class revolves around three concepts, each coming in a weaker and stronger form.

1. **Pebbling games on directed acyclic graphs.** A model that capture how much space is needed to perform a sequence of local computation steps.
   - *black pebbling game* models deterministic computation;
   - *black-white pebbling game* models non-deterministic computation (i.e. verifiable computation with guesses).
2. **Resolution space.** We extend the notion described last time and we consider
   - *variable space* is the number of variables in a memory configuration (counted without repetitions);
   - *clause space* is the number of clauses in a memory configuration;[3]

   and we extend these two notions to derivations and refutations and the maximum of the measure, across all configuration in a derivation.
3. **Pebbling formulas.** These are easily refutable formulas that exhibit nice relation between time and space.

[3] This is just a renaming of the notion of space we discussed last for resolution.

- $\text{Peb}(G)$ is the *simple version*. The variable space of its refutation will roughly correspond to the cost in the pebbling games for G;
- $XPebF(G)$ is the *XOR-ified version*. The clause space of its refutation will roughly correspond to the cost in the pebbling games for G.

The concept of variable space is quite unrelated with the particular proof system. At least when it comes to weak proof systems.

Exercise 1. Assume that there is a refutation of ϕ in either Pc, Pcr or cutting planes where all memory configurations have at most s variables. Show that there is resolution refutation of ϕ in variable space s.

Pebbling games

Pebbling games are defined on **directed acyclic graphs** with constant degree. We call

- *indegree* or *incoming degree* of a vertex the number of its incoming edges;
- *outdegree* or *outgoing degree* of a vertex the number of its outgoing edges;
- for every edge (u, v), u is a *predecessor* of v and v is a *successor* of u;
- *sink* a vertex with no successors;
- *source* a vertex with no predecessors.

We denote the set of predecessors and of successors of a vertex v as $\text{pred}(v)$ and $\text{succ}(v)$, respectively. In this lecture we only focus only on directed acyclic graphs with **exactly one sink** and **indegree at most two**. A pebbling game on such a graph G is a sequence of pebbling configurations. At each configuration each vertex of the graph may have either one pebble or no pebble at all, and each configuration follows from the previous one according to some legal moves. The goal is to place a pebble on the sink of the graph, having as little pebbles simultaneously on the graph as possible.

Black pebbling game A *black pebbling*[4] for a graph graph G is a sequence of configurations that starts with a configuration in which there is no pebble on any vertex. Each new configuration follows from the previous using a move which is either

[4] Michael S Paterson and Carl E Hewitt. Comparative schematology. In *Record of the Project MAC conference on concurrent systems and parallel computation*, pages 119–127. ACM, 1970; and Ravi Sethi. Complete register allocation problems. *SIAM journal on Computing*, 4(3):226–248, 1975

- (Pebble placement) if all predecessors of an unpebbled vertex v have pebbles, we can place a black pebble on v;
- (Pebble removal) remove a pebble from a pebbled vertex.

The goal of the game is to place a pebble to the sink vertex (i.e. to reach a configuration where there is a black pebble on the sink). The cost of the black pebbling is the maximum number of pebbles on G among all configurations in the pebbling. The *black pebbling number* of G, denoted as $\text{Black}(G)$, is the minimum cost among all black pebbling of G.

The intuition of a black pebbling is that G represent a sequence of partial results in a computation, where each step depends on its predecessors to be employed (e.g. the partial result in a formula evaluation or in a straight line program). The number of black pebbles are the number of partial results needed to be kept in memory simultaneously in order to complete the computation.

Black-White pebbling game. A *black-white pebbling*[5] for a graph graph G is a sequence of configurations that starts with a configuration in which there is no pebble on any vertex. Each new configuration follows from the previous using a move which is either

[5] Stephen Cook and Ravi Sethi. Storage requirements for deterministic/polynomial time recognizable languages. In *Proceedings of the sixth annual ACM symposium on Theory of computing*, pages 33–39. ACM, 1974

- (Black Pebble placement) if all predecessors of an unpebbled vertex v have pebbles, we can place a black pebble on v;
- (White Pebble placement) place a white pebble on any unpebbled vertex;
- (Black Pebble removal) remove a black pebble from any black pebbled vertex.
- (White Pebble placement) if all predecessors of a white pebbled vertex v have pebbles, we can remove the white pebble from v;

The goal of the game is to place a pebble to the sink vertex at some point and then to remove all white pebbles from the graph. The cost of the black pebbling is the maximum number of pebbles on G among all configurations in the pebbling. The *black-white pebbling number* of G, denoted as $\mathrm{BW}(G)$, is the minimum cost among all black-white pebbling of G.

The intuition of a black white pebbling is that of a non-determistic verifiable computation. A white pebble corresponds to a partial result in the computation which is not computed but it is instead just guessed (and verified at least once before being erased).

Exercise 2 (Black-white pebble reversal)**.** Consider a black white pebbling of G, and enumerate its pebbling configurations $P_1, P_2, \ldots P_\ell$. Show that $P'_\ell, P'_{\ell-1}, \ldots, P'_2, P'_1$ is a legal black white pebbling of G, where P'_i is a copy of configuration P_i where black and white pebbles are flipped.

Exercise 3. Consider a black pebbling of cost s. There is always a way to transform it in such a way that the time and space do not increase, but after every black pebble placement on sources the space of the configuration is at most $s - 1$.

Remark 4. *Of course every black pebbling is also a legal black white pebbling.*

The main goal of the lecture is to argue that the following relations hold.

Theorem 5 (Informal)**.** *These relations hold*

$$\mathrm{BW}(G) \lessapprox \textit{variable space for } \mathrm{Peb}(G) \lessapprox \mathrm{Black}(G)$$

$$\mathrm{BW}(G) \lessapprox \textit{clause space for } \mathrm{Peb}^{\oplus 2}(G) \lessapprox \mathrm{Black}(G)\ .$$

Furthermore these relations hold also if we bound for all pebblings and refutation with length/time bounded by the same function.

Therefore we can study lower bounds and trade-offs for pebbling formulas (to be defined later) using lower bounds and trade-offs for pebbling games.

Bounds and treade-offs for pebbling games

An interesting observation is that no graph requires linear pebbling number.[6]

Proposition 6 (Hopcrotf et al. 1977)**.** *Give a DAG with n vertice,s one sink and indegree at most* 2*, there is always a black pebbling of space* $O(n/\log n)$.

Exercise 7. Use the upper bounds for pebbling to show that any formula that has a refutation of length L can be refutated in space $O(L/\log L)$.

Exercise 8. Prove that any graph with a pebbling of constant space can be pebbled in polynomial time and constant space simultaneously.

The next two results clarify the relative power of black and black-white pebbling.[7]

Proposition 9 (Meyer auf der Heide, 1981)**.** *For any DAG G, it holds that* $\text{Black}(G) \leq \text{BW}(G))^2/2$.

Proposition 10 (Kalyanasundaram and Schnitger, 1991)**.** *There is a graph family* $\{G_s\}_{s=1}^{\infty}$ *of size* $\exp(s \log s)$ *where* $\text{Black}(G_n) \geq s^2$ *and* $\text{BW}(G_n \leq 3s - 1$.

It is still open to find a polynomial size constriction.

Nevertheless since we resolution space complexity is sandwiched between black-white and black-space, it makes sense to find graphs from which the upper bounds hold for black pebbling and the lower bounds hold for black-white pebbling.

The next lower bounds and trade-offs for the pebbling numbers of DAGs comes from a series of classic papers (and a newer one) on pebbling[8]. All results below refer (or can be easily adapted) to some family of DAGs G_n with $O(n)$ vertices one sink indegree at most two.

- (Gilbert, Tarjan, 1978) **Optimal lower bound**
  There is a family of graph with black-white pebbling lower bound $\Omega\left(\frac{n}{\log n}\right)$;
- (Lengauer, Tarjan, 1982) **Trade-offs for the polynomial range**
  There is a family of graph with
  - black pebbling of space 3;
  - for every $s \leq n$ any black pebbling of space s takes time $\Theta(n^2/s)$;
  - for every $s \leq \sqrt{n}$ any black-white pebbling of space s takes time $\Theta((n/s)^2)$.
- (Nordstrom, 2012; based on Carlson Savage, 1982) **Trade-offs for small non-constant space**
  For every slowly growing function $g(n) = O(n^{1/7})$ and $\epsilon > 0$, there is a graph family so that
  - has black pebblings of space $O(g(n))$;

[6] John Hopcroft, Wolfgang Paul, and Leslie Valiant. On time versus space. *Journal of the ACM (JACM)*, 24(2):332–337, 1977

[7] Friedhelm Meyer Auf Der Heide. A comparison of two variations of a pebble game on graphs. *Theoretical Computer Science*, 13(3):315–322, 1981; and Bala Kalyanasundaram and Georg Schnitger. On the power of white pebbles. *Combinatorica*, 11(2):157–171, 1991

[8] John R Gilbert and Robert E Tarjan. Variations of a pebble game on graphs. Technical report, DTIC Document, 1978; T. Lengauer and R.E. Tarjan. Asymptotically tight bounds on time-space trade-offs in a pebble game. *Journal of the ACM (JACM)*, 29(4):1087–1130, 1982; David A Carlson and John E Savage. Graph pebbling with many free pebbles can be difficult. In *Proceedings of the twelfth annual ACM symposium on Theory of computing*, pages 326–332. ACM, 1980; David A Carlson and John E Savage. Extreme time-space tradeoffs for graphs with small space requirements. *Information Processing Letters*, 14(5):223–227, 1982; and Jakob Nordström. On the relative strength of pebbling and resolution. *ACM Transactions on Computational Logic (TOCL)*, 13(2):16, 2012

  - has a black pebbling of linear time and space $O\big(\sqrt[3]{n/g^2(n)}\big)$;
  - any black-white pebbling in space $\big(n/g^2(n)\big)^{1/3-\epsilon}$ requires super-polynomial time;

- (Lengauer, Tarjan, 1982) **Robust trade-offs**
  There is a graph family so that
  - has black pebbling of space $O(\log^2 n)$;
  - has a black pebbling of linear time and space $O(n/\log n)$;
  - there is a constant K such that any black-white pebbling in space $<$ $Kn/\log n$ requires time $n^{\Omega(\log\log n)}$;

- (Lengauer, Tarjan, 1982) **Exponential trade-offs**
  Let $K > 0$ a large enough constant. There is a graph family and a constant $K' \ll K$ such that
  - has black pebbling of space $K'n/\log n$;
  - has a black pebbling of linear time and linear space;
  - any black-white pebbling in space $< Kn/\log n$ requires time $\exp(n^{\Omega(1)})$.

Definitions of the Pebbling formulas

We consider a formula, defined on single sink directed acyclic graphs G. The formula says that (1) sources can be pebbled; (2) if all predecessors of a vertex can be pebbled, then the vertex can be pebbled; (3) the sink cannot be pebbled. The formula is essentially the negation of the claim that G has some pebbling strategy in the pebbling game (i.e. the sink is reachable from all sources).

Definition 11 (Pebbling formula $\mathrm{Peb}(G)$)**.** *Consider a DAG G with sink z, the pebbling formula* $\mathrm{Peb}(G)$ *has a variable x_v for every vertex $v \in V(G)$. The constraints are*

$$\neg x_z \qquad\qquad z \text{ is the unique sink of } V(G); \tag{1}$$

$$\Big(\bigwedge_{u\in\mathrm{pred}(v)} x_u\Big) \longrightarrow x_v \qquad\qquad \text{for every } v \in V(G). \tag{2}$$

Assuming G has in-degree 2 this is a 3-CNF in Horn form. The harder variant of the pebbling formula is the one in which the reachability of a vertex is represented by the XOR or two distinct variables.

Definition 12 (Pebbling formula $\mathrm{Peb}^{\oplus_2}(G)$)**.** *Consider a DAG G with sink z, the pebbling formula* $\mathrm{Peb}^{\oplus_2}(G)$ *has variables x_v and y_v for every vertex $v \in V(G)$.*

$$\neg x_z \oplus y_z \qquad\qquad z \text{ is the unique sink of } V(G); \tag{3}$$

$$\Big(\bigwedge_{u\in\mathrm{pred}(v)} x_u \oplus y_u\Big) \rightarrow x_v \oplus y_v \qquad\qquad \text{for every } v \in V(G). \tag{4}$$

While the constraints are not given in CNF form, each one can be expressed with at most 8 clauses on 6 variables.

Exercise 13. Show that for every G there is a refutation of $\mathrm{Peb}(G)$ of simultaneously

- length $O(n)$ and width $O(1)$;

and one of simultaneously

- length $O(n)$ and clause space $O(1)$.

Exercise 14. Show that for every G there is a refutation of $\mathrm{Peb}^{\oplus_2}(G)$ of simultaneously length $O(n)$ and width $O(1)$.

Corollary 15 (From Proposition 6)**.** *Pebbling formulas* $\mathrm{Peb}(G)$ *and* $\mathrm{Peb}^{\oplus_2}(G)$ *can always be refuted in sublinear space.*

Connection between pebbling games and pebbling formulas

Now we have all tools to develop a theory of pebbling formula. The main tenets of this theory is the connection between pebbling cost of a graph G and the space complexity of the corresponding pebbling formula over G. We will leave the simple statements as exercises and we will discuss or prove (time permitting) the more difficult part.

1. black pebblings upper bounds give resolutions refutations upper bounds;
2. black-white pebbling lower bounds give resolution refutation lower bounds;
3. the bounds apply for variable space of $\mathrm{Peb}(G)$ and for clause space of $\mathrm{Peb}^{\oplus_2}(G)$.

Black pebbling strategies give resolution refutations. If we have a strategy for a space efficient black pebbling of a graph we can find a space efficient refutation of pebbling formula.[9]

[9] Eli Ben-Sasson, Russell Impagliazzo, and Avi Wigderson. Near optimal separation of tree-like and general resolution. *Combinatorica*, 24(4):585–603, 2004; and Eli Ben-Sasson. Size-space tradeoffs for resolution. *SIAM Journal on Computing*, 38(6):2511–2525, 2009

Exercise 16. Show how to obtain, from a black pebbling strategy of graph G in time t and space s, a refutation of $\mathrm{Peb}(G)$ of length $O(t)$, variable space $O(s)$ and width $O(1)$

Exercise 17. Show how to obtain, from a black pebbling strategy of graph G in time t and space s, a refutation of $\mathrm{Peb}^{\oplus_2}(G)$ of length $O(t)$ and clause space $O(s)$ and width $O(1)$.

It is natural to ask whether we can simulate black-white pebbling in resolution. Here the question is still open even though there are restricted versions of black-white pebbling that (1) can be simulated in resolution (2) it is strictly stronger than black pebbling on some graphs.[10] This means that black pebbling strategies are not optimal in refuting pebbling formulas.

[10] Jakob Nordström. On the relative strength of pebbling and resolution. *ACM Transactions on Computational Logic (TOCL)*, 13(2):16, 2012

Resolution refutations give black-white pebblings. It seems that power of resolution refutations on pebbling formulas in sandwiched between the power of black-white pebbling and black pebbling. It is quite an interesting exercise to see that this holds for variable space of $\mathrm{Peb}(G)$ formulas.

Exercise 18. Consider, without loss of generality, a resolution refutation

$$M_0, M_1, \dots M_T$$

for $\mathrm{Peb}(G)$, where

- there is no weakening;
- if we let say that the empty clause in M_T is essential, and than each clause used to derive an essential clause is essential, then the proof does only contain essential clauses.

Show that we can extract from any such proof of length t and variable space s a black-white pebbling strategy of length $O(t)$ and space at most s.
(Hint: turn each memory configuration into a pebbling configuration, so that negative literals correpsond to white pebbles and positive literals corresponds to black pebbles.)
(Hint: argue that the sink node is mentioned at some points and that this implies that the conversion produces a pebbling.)

For clause space and XOR-ified pebbling formulas $\mathrm{Peb}^{\oplus_2}(G)$ the proof is much longer and involved, but in the end Ben-Sasson, Nordström[11] showed that it is possible to turn a refutation of $\mathrm{Peb}^{\oplus_2}(G)$ into a refutation of $\mathrm{Peb}(G)$ so that clause space turns into variables space, and length is preserved.[12]

[11] Eli Ben-Sasson and Jakob Nordström. Short proofs may be spacious: An optimal separation of space and length in resolution. In *FOCS*, pages 709–718, 2008; and Eli Ben-Sasson and Jakob Nordström. Understanding space in proof complexity: Separations and trade-offs via substitutions. In *ICS*, pages 401–416, 2011

[12] Their theorem is much more general but we only need this simpler statement.

Theorem 19 (Ben-Sasson, Nordström (2011))**.** *Any refutation of* $\mathrm{Peb}^{\oplus_2}(G)$ *in clause space s and length l can be transformed into a refutation of* $\mathrm{Peb}(G)$ *of length $O(l)$ and variable space $O(s)$.*

Corollary 20. *Any proof of length t and clause space s for $XPebF(G)$ can be transformed into black-white pebbling strategy of length at most $O(t)$ and space at most $O(s)$.*

We won't prove this theorem, but we will prove a simpler theorem which will still give some weaker space trade-offs and lower bounds.

The simpler version of the clause space lower bound in term of black-white pebbling only applies for short pebbling refutations. While the result is per se limited, has a much easier proof.[13]

[13] Eli Ben-Sasson. Size-space tradeoffs for resolution. *SIAM Journal on Computing*, 38(6):2511–2525, 2009

Theorem 21 (Ben Sasson, 2009)**.** *Consider a n unsatisfiable k-CNF formula F that requires refutations of variable V. Let $F[\oplus_2]$ the $2k$-CNF obtained by substituting each variable in F by the $\oplus_2$ of two new variables. For any resolution proof of $F[\oplus_2]$ with clause space C and length L it holds that*

$$C \cdot \log_{4/3} L \geq V \ . \tag{5}$$

Proof sketch. We start with $F[\oplus_2]$ formula, and we assume it has a refutation Π that uses clause space C and it has length L. Consider a random restriction that picks at random one variable in each pair of the xorification and fix it with to a random bit. Observe that after the restriction, the formula is equal to $\mathrm{Peb}(G)$ up to renaming of variables and up to polarity inversions.

Now we use the simple observation for for any refutation

$$\text{clause space} \times \text{width} \geq \text{variable space} \tag{6}$$

and since the restriction does not increase neither width nor clause space, we have that the width needed to refute the restricted formula is $\frac{V}{C}$.

The probability that one clause of width $\frac{V}{C}$ (a parameter to be fixed later) is killed by the restriction is

$$\left(\frac{3}{4}\right)^{\frac{V}{C}} \tag{7}$$

therefore

$$L \geq \left(\frac{4}{3}\right)^{\frac{V}{C}} \tag{8}$$

otherwise we would get a contradiction. □

Theorem 21, together with the pebbling results in the previous sections, implies clause space lower bounds for XOR-ified pebbling formulas, at least when we consider only short refutation. To avoid this last condition and get unconditional lower bounds and strong trade-offs we need to use Theorem 19 instead.

References

[BS09] Eli Ben-Sasson. Size-space tradeoffs for resolution. *SIAM Journal on Computing*, 38(6):2511–2525, 2009.

[BSIW04] Eli Ben-Sasson, Russell Impagliazzo, and Avi Wigderson. Near optimal separation of tree-like and general resolution. *Combinatorica*, 24(4):585–603, 2004.

[BSN08] Eli Ben-Sasson and Jakob Nordström. Short proofs may be spacious: An optimal separation of space and length in resolution. In *FOCS*, pages 709–718, 2008.

[BSN11] Eli Ben-Sasson and Jakob Nordström. Understanding space in proof complexity: Separations and trade-offs via substitutions. In *ICS*, pages 401–416, 2011.

[CS74] Stephen Cook and Ravi Sethi. Storage requirements for deterministic/polynomial time recognizable languages. In *Proceedings of the sixth annual ACM symposium on Theory of computing*, pages 33–39. ACM, 1974.

[CS80] David A Carlson and John E Savage. Graph pebbling with many free pebbles can be difficult. In *Proceedings of the twelfth annual ACM symposium on Theory of computing*, pages 326–332. ACM, 1980.

[CS82] David A Carlson and John E Savage. Extreme time-space tradeoffs for graphs with small space requirements. *Information Processing Letters*, 14(5):223–227, 1982.

[GT78] John R Gilbert and Robert E Tarjan. Variations of a pebble game on graphs. Technical report, DTIC Document, 1978.

[HPV77] John Hopcroft, Wolfgang Paul, and Leslie Valiant. On time versus space. *Journal of the ACM (JACM)*, 24(2):332–337, 1977.

[KS91] Bala Kalyanasundaram and Georg Schnitger. On the power of white pebbles. *Combinatorica*, 11(2):157–171, 1991.

[LT82] T. Lengauer and R.E. Tarjan. Asymptotically tight bounds on time-space trade-offs in a pebble game. *Journal of the ACM (JACM)*, 29(4):1087–1130, 1982.

[MADH81] Friedhelm Meyer Auf Der Heide. A comparison of two variations of a pebble game on graphs. *Theoretical Computer Science*, 13(3):315–322, 1981.

[Nor12] Jakob Nordström. On the relative strength of pebbling and resolution. *ACM Transactions on Computational Logic (TOCL)*, 13(2):16, 2012.

[Nor13] Jakob Nordström. Pebble games, proof complexity and time-space trade-offs. *Logical Methods in Computer Science*, 9:15:1–15:63, September 2013.

[Nor15] Jakob Nordström. New wine into old wineskins: A survey of some pebbling classics with supplemental results. Manuscript in preparation., 2015.

[PH70] Michael S Paterson and Carl E Hewitt. Comparative schematology. In *Record of the Project MAC conference on concurrent systems and parallel computation*, pages 119–127. ACM, 1970.

[Set75] Ravi Sethi. Complete register allocation problems. *SIAM journal on Computing*, 4(3):226–248, 1975.