Electronic Colloquium on Computational Complexity, Report No. 12 (2013)

A Simple Algorithm for Undirected Hamiltonicity

Hasan Abasi Nader H. Bshouty
Department of Computer Science
Technion,32000

January 16, 2013

Abstract

We develop a new algebraic technique that gives a simple randomized algorithm for the simple k-path problem with the same complexity $O^*(1.657^k)$ as in [1] and [3].

1 Introduction

Given an undirected graph G on n vertices, the *k-path problem* asks whether G contains a simple path of length k. For $k = n-1$ the problem is the Hamiltonian path problem in undirected graph.

Björklund proved in [1]

Theorem 1. *The undirected Hamiltonian path problem can be solved in time* $O^*(1.657^n)$ *by a randomized algorithm with constant, one sided error.*

Björklund et. al. proved in [3]

Theorem 2. *The undirected k-path problem can be solved in time* $O^*(1.657^k)$ *by a randomized algorithm with constant, one sided error.*

In this paper we give a simple algorithm for the above problems with same complexity as above.

2 Preliminary Results

In this section we give some preliminary results

Let $G(V,E)$ be an undirected graph with $n = |V|$ vertices. A *k-path* is $v_0, v_1, \ldots, v_k$ such that $\{v_i, v_{i+1}\} \in E$ for every $i = 0, \ldots, k-1$. A k-path $v_0, v_1, \ldots, v_k$ is called *simple* if the vertices in the path are distinct.

Let $V = V_1 \cup V_2$ be a partition of V. Let $E_1 = E(V_1)$ and $E_2 = E(V_2)$ be the set of edges with both ends in V_1 and V_2, respectively. Our goal is to find a simple k-path that starts from some fixed vertex. For a path $p = v_0, v_1, \ldots, v_k$ we define the multiset of vertices in p as $V(p) = \{v_0, \ldots, v_k\}$ and the (undirected) edges in p as the multiset

ISSN 1433-8092

$E(p) = \{\{v_0, v_1\}, \ldots, \{v_{k-1}, v_k\}\}$. When we write $V(p) \cap V_1$ (or $E(p) \cap E_2$) we mean the multiset that contains the elements in $V(p)$ that are also in V_1.

Define for every edge $e \in E$ a variable x_e, for every vertex $v \in V_1$ a variable y_v and for every edge $e \in E_2$ a variable z_e. Let $x = (x_e)_{e \in E}$ $y = (y_v)_{v \in V_1}$ and $z = (z_e)_{e \in E_2}$. For every k-path $p = v_0, v_1, \ldots, v_k$ we define a monomial over any field of characteristic 2 $M_p = X_p Y_p Z_p$ where

$$X_p = \left(\prod_{e \in E(p)} x_e \right), Y_p = \left(\prod_{v \in V(p) \cap V_1} y_v \right) \text{ and } Z_p = \left(\prod_{e \in E(p) \cap E_2} z_e \right).$$

Note here that if e appears twice in $E(p)$ then x_e appears twice in X_p.

A path $p = v_0, v_1, \ldots, v_k$ is called (r,s)-*legitimate* k-*path* with the partition $V = V_1 \cup V_2$ if $|V(p) \cap V_1| = r$, $|E(p) \cap E_2| = s$ and it contains no three consecutive vertices v_i, v_{i+1}, v_{i+2} where $v_{i+2} = v_i$, $v_i \in V_2$ and $v_{i+1} \in V_1$. Fix a vertex $v_0 \in V_1$. We denote by $\mathcal{L}_{k,r,s}(v_0, V_1, V_2)$ the set of all (r,s)-legitimate k-paths in G with the partition $V_1 \cup V_2 = V$ that starts from $v_0 \in V_1$. Define

$$F_{k,r,s}^{v_0,V_1,V_2}(x,y,z) = \sum_{p \in \mathcal{L}_{k,r,s}(v_0,V_1,V_2)} M_p.$$

We now prove the following results.

Lemma 1. *Given an undirected graph $G = (V,E)$, a partition $V = V_1 \cup V_2$, $v_0 \in V_1$ and two integers s and r. There is a deterministic polynomial time algorithm that construct a polynomial size circuit for the function $F_{k,r,s}^{v_0,V_1,V_2}(x,y,z)$.*

Lemma 2. *$M_p = X_p Y_p Z_p$ is a monomial in $F_{k,r,s}^{v_0,V_1,V_2}(x,y,z)$ and $Y_p Z_p$ is multilinear if and only if p is a (r,s)-legitimate simple k-path with the partition $V_1 \cup V_2 = V$ that starts from v_0.*

Lemma 3. *There is a randomized algorithm with constant, one sided error, that runs in time $O^*(2^{r+s})$ for the following decision problem: Given a black box for the multivariate polynomial $f(x,y,z) := F_{k,r,s}^{v_0,V_1,V_2}(x,y,z)$ over a field of characteristic 2. Decides whether f contains a monomial $M_p = X_p Y_p Z_p$ where $Y_p Z_p$ is multilinear.*

Proof of Lemma 1. For any two vertices $u_1, u_2 \in V$ we define $\mathcal{L}_{k,r,s}(v_0, V_1, V_2, u_1, u_2)$ the set of all (r,s)-legitimate k-paths that start with v_0 and end with u_1, u_2. Define

$$F_{k,r,s}^{v_0,V_1,V_2,u_1,u_2} = \sum_{p \in \mathcal{L}_{k,r,s}(v_0,V_1,V_2,u_1,u_2)} M_p.$$

Then

$$F_{k,r,s}^{v_0,V_1,V_2}(x,y,z) = \sum_{u_1,u_2 \in V} F_{k,r,s}^{v_0,V_1,V_2,u_1,u_2}.$$

We now show, using dynamic programming, that $F_{k,r,s}^{v_0,V_1,V_2,u_1,u_2}$ can be computed in polynomial time. For a vertex v let $N(v)$ be the neighbor vertices of v. For a predicate A we define $[A] = 1$ if A is true and 0 otherwise. Now it is easy to verify the following recurrence formula

1. If $k \geq 2$, $r \leq k+1, s \leq k$ and $\{u_1, u_2\} \in E$ then

$$\begin{aligned} F_{k,r,s}^{v_0,V_1,V_2,u_1,u_2} &= [u_2 \in V_1] \cdot x_{\{u_1,u_2\}} y_{u_2} \sum_{w \in N(u_1)} F_{k-1,r-1,s}^{v_0,V_1,V_2,w,u_1} \\ &\quad + [u_2 \in V_2 \wedge u_1 \in V_2] \cdot x_{\{u_1,u_2\}} z_{\{u_1,u_2\}} \sum_{w \in N(u_1)} F_{k-1,r,s-1}^{v_0,V_1,V_2,w,u_1} \\ &\quad + [u_2 \in V_2 \wedge u_1 \in V_1] \cdot x_{\{u_1,u_2\}} \sum_{w \in N(u_1) \backslash \{u_2\}} F_{k-1,r,s}^{v_0,V_1,V_2,w,u_1} \end{aligned}$$

2. If $k = 1$, $u_1 = v_0$, $u_2 \in V_2$, $r = 1$ and $s = 0$ then $F_{k,r,s}^{v_0,V_1,V_2,u_1,u_2} = x_{\{v_0,u_2\}} y_{v_0}$.

3. If $k = 1$, $u_1 = v_0$, $u_2 \in V_1$, $r = 2$ and $s = 0$ then $F_{k,r,s}^{v_0,V_1,V_2,u_1,u_2} = x_{\{v_0,u_2\}} y_{v_0} y_{u_2}$.

4. Otherwise $F_{k,r,s}^{v_0,V_1,V_2,u_1,u_2} = 0$.

Since u_1, u_2, k, r, s can take at most $k^2(k+1)n^2$ different values the above recurrence can be computed in polynomial time. □

Proof of Lemma 2. ($\Leftarrow$) Let $p = v_0, v_1, \ldots, v_k$ be any (r,s)-legitimate simple k-path with the partition $V_1 \cup V_2 = V$. Then $p \in \mathcal{L}_{k,r,s}(v_0, V_1, V_2)$. Since p is simple $Y_p Z_p$ is multilinear. We now need to show that no other path p' satisfies $M_{p'} = M_p$. If $M_p = M_{p'}$ then $X_p = X_{p'}$ and since p is simple and starts from v_0 by induction on the path $p \equiv p'$. Therefore M_p is a multilinear monomial in $F_{k,r,s}^{v_0,V_1,V_2}(x, y, z)$

($\Rightarrow$) We now show that all the monomials that correspond to (r,s)-legitimate non-simple k-path $p = v_0, v_1, \ldots, v_k$ with the partition $V_1 \cup V_2 = V$ either vanish (because the field is of characteristic 2) or are not multilinear.

Consider a (r,s)-legitimate non-simple k-path with the partition $V_1 \cup V_2 = V$. Consider the first circuit C in this path. If $C = v_i, v_{i+1}, v_i$ then either $v_i \in V_1$ and then Y_p contains $y_{v_i}^2$ or $v_i, v_{i+1} \in V_2$ and then Z_p contains $z_{\{v_i,v_{i+1}\}}^2$. Notice that p is legitimate and therefore the case $v_i \in V_2$ and $v_{i+1} \in V_1$ cannot happen.

Now suppose $|C| > 2$, $C = v_i, v_{i+1}, \ldots, v_j, v_{j+1} (= v_i)$. Define $p_1 = v_1, \ldots, v_{i-1}$ and $p_2 = v_{j+2}, \ldots, v_k$. Then $p = p_1 C p_2$. If $v_i \in V_1$ then Y_p contains $y_{v_i} y_{v_{j+1}} = y_{v_i}^2$. Therefore we may assume that $v_i \in V_2$. Define the path

$$\rho(p) := p_1 C' p_2 = \underline{v_0, v_1, \ldots, v_{i-1}}, \underline{v_i, v_j, v_{j-1}, \ldots, v_{i+1}, v_i}, \underline{v_{j+2}, v_{j+3}, \ldots, v_k}.$$

We now show that

1. $\rho(\rho(p)) = p$.

2. $\rho(p)$ is (r,s)-legitimate non-simple k-path with the partition $V_1 \cup V_2 = V$ that starts with v_0.

3. $\rho(p) \neq p$ and $M_p = M_{\rho(p)}$.

This implies that M_p vanishes from $F^{v_0,V_1,V_2}_{k,r,s}(x,y,z)$ because the characteristic of the field is 2. Let $p'=\rho(p)$. Since C is the first circuit in p we have $v_0,v_1,\ldots,v_j$ are distinct and $v_{j+1}=v_i$. This implies that C' is the first circuit in p' and therefore $\rho(\rho(p))=\rho(p')=p$. This implies 1.

Obviously, $|V(p')\cap V_1|=|V(p)\cap V_1|=r$ and $|E(p')\cap E_2|=|E(p)\cap E_2|=s$. Suppose p' contains three consecutive vertices u,w,u such that $w\in V_1$ and $u\in V_2$. Then since p is (r,s)-legitimate path and $v_i\in V_2$ we have three cases

Case I. u,w,u is in C'. Then u,w,u is in C which contradict the fact that p is legitimate.

Case II. $u=v_{i-2}\in V_2, w=v_{i-1}\in V_1$ and $v_{i-2}=v_i$. In this case $C''=v_{i-2},v_{i-1},v_i$ is a circuit in p and then p is not legitimate. A contradiction.

Case III. $u=v_i\in V_2, w=v_{j+2}\in V_1$ and $v_{j+3}=v_i$. In this case, $C''=v_iv_{j+2}v_{j+3}$ is a circuit in p and then p is not legitimate. A contradiction.

This proves 2.

If $p=p'$ then $C=C'$ and since $|C|>2$ we get a contradiction. Therefore $p\neq p'$. Since $E(p)=E(p')$, $V(p')\cap V_1=V(p)\cap V_1$ and $E(p')\cap E_2=E(p)\cap E_2$ we also have $M_p=M_{p'}$. This proves 3. $\square$

Proof of Lemma 3. Consider the new indeterminates $y^{(i)}=(y_{i,v})_{v\in V_1}$, $i\in[r]$ and $z^{(j)}=(z_{j,e})_{e\in E_2}$, $j\in[s]$ where $[r]=\{1,2,\ldots,r\}$. Consider the operator

$$\Phi(f)=\sum_{S\subseteq[s]}\sum_{R\subseteq[r]} f\left(x,\sum_{i\in R}y^{(i)},\sum_{j\in S}z^{(j)}\right).$$

If $f(x,y,z)=\sum_{p\in P}X_pY_pZ_p$ then $\Phi(f)=\sum_{p\in P}X_p\Phi(Y_pZ_p)$. If $Y_p=y_{v_1}\cdots y_{v_r}$ and $Z_p=z_{e_1}\cdots z_{e_s}$ then, by Ryser formula for permanent and since permanent in field of characteristic 2 is equal to determinant we have

$$\begin{aligned}
\Phi(Y_pZ_p) &= \sum_{S\subseteq[s]}\sum_{R\subseteq[r]}\left(\prod_{i_1=1}^{r}\sum_{i_2\in R}y_{i_2,v_{i_1}}\prod_{j_1=1}^{s}\sum_{j_2\in S}z_{j_2,e_{j_1}}\right)\\
&= \left(\sum_{R\subseteq[r]}\prod_{i_1=1}^{r}\sum_{i_2\in R}y_{i_2,v_{i_1}}\right)\left(\sum_{S\subseteq[s]}\prod_{j_1=1}^{s}\sum_{j_2\in S}z_{j_2,e_{j_1}}\right)\\
&= \left(\sum_{R\subseteq[r]}(-1)^{r-|R|}\prod_{i_1=1}^{r}\sum_{i_2\in R}y_{i_2,v_{i_1}}\right)\left(\sum_{S\subseteq[s]}(-1)^{s-|S|}\prod_{j_1=1}^{s}\sum_{j_2\in S}z_{j_2,e_{j_1}}\right)\\
&= \mathrm{Per}\left(y_{i_2,v_{i_1}}\right)_{i_1,i_2\in[r]}\mathrm{Per}\left(z_{j_2,v_{j_1}}\right)_{j_1,j_2\in[s]}=\det\left(y_{i_2,v_{i_1}}\right)_{i_1,i_2\in[r]}\det\left(z_{j_2,v_{j_1}}\right)_{j_1,j_2\in[s]}.
\end{aligned}$$

Now if in Y_p (or Z_p) we have $y_{v_a}=y_{v_b}$, or equivalently $v_a=v_b$, for some $a\neq b$ then $\det\left(y_{i_2,v_{i_1}}\right)_{i_1,i_2\in[r]}=0$ and $\Phi(Y_pZ_p)\equiv 0$. If Y_pZ_p is multilinear then $\Phi(Y_pZ_p)\not\equiv 0$. Therefore $\Phi(f)\not\equiv 0$ if and only if f contains a monomial $M_p=X_pY_pZ_p$ where Y_pZ_p is multilinear.

Now since substitution in $\Phi(f)$ can be simulated by 2^{r+s} substitutions in f and the degree of $\Phi(f)$ is $k+r+s\leq 3k+1$ we can randomly zero test f in time $O(poly(k,n)2^{r+s})$ [2, 4, 5]. $\square$

3 Main Results

In this section we prove the Theorems

The following lemma is proved in [3] we give its proof for completeness

Lemma 4. *Let $p = v_0, v_1, \ldots, v_k$ be a simple path. For a partition V_1, V_2 selected uniformly at random where $v_0 \in V_1$,*

$$\Pr_{V_1,V_2}\left(|V(p)\cap V_1| = r, |E(p)\cap E_2| = s\right) = 2^{-k}\binom{r}{k-r-s+1}\binom{k-r}{s}.$$

Proof. We will count the number of partitions V_1, V_2 that satisfies $|V(p)\cap V_1| = r$, $|E(p)\cap E_2| = s$ and $v_0 \in V_1$. Obviously, the probability in the lemma is 2^{-k} times the number of such partitions.

Let V_1, V_2 be a partition such that $|V(p)\cap V_1| = r$, $|E(p)\cap E_2| = s$ and $v_0 \in V_1$. Let $v_{i_1} = v_0, v_{i_2}, \ldots, v_{i_r}$ be the nodes in V_1. Let $\bar{s}_j \geq 0$, $j = 1, \ldots, r-1$ be the number of nodes in V_2 that are between v_{i_j} and $v_{i_{j+1}}$. Let $\bar{s}_r$ be the number of nodes in V_2 that are after v_{i_r}. Let t be the number of $\bar{s}_i$ that are not zero. For $j < r$ the number of edges in E_2 that are between v_{i_j} and $v_{i_{j+1}}$ is $s_i := \max(\bar{s}_i - 1, 0)$. The number of edges in E_2 that are after v_{i_r} is $s_r := \max(\bar{s}_r - 1, 0)$. Therefore

$$\sum_{i=1}^{r} \bar{s}_i = \sum_{i=1}^{r} s_i + t = s + t. \tag{1}$$

Since the number of nodes in the path is

$$k + 1 = r + \sum_{i=1}^{r} \bar{s}_i = r + s + t \tag{2}$$

we must have $t = (k+1) - (r+s)$.

Now any partition that satisfies $\sum_{i=1}^{r} \bar{s}_i = s + t$ and $t = (k+1) - (r+s)$ must also satisfy $|V(p)\cap V_1| = r$ and $|E(p)\cap E_2| = s$. Therefore the number of such partitions is equal to the number ways of writing $s + t$ as $\bar{s}_1 + \bar{s}_2 + \cdots + \bar{s}_r$ where exactly t of them are not zero. We first select those $\bar{s}_{j_1}, \ldots, \bar{s}_{j_t}$ that are not zero. This can be done in $\binom{r}{t}$ ways. Then the number of ways of writing $s + t$ as $\bar{s}_{j_1} + \cdots + \bar{s}_{j_t}$ where $\bar{s}_{j_i} \geq 1$ is equal to the number of ways of writing s as $x_1 + \cdots + x_t$ where $x_i \geq 0$. The later is equal to $\binom{t+s-1}{t-1}$. Therefore the number of such partitions is

$$\binom{t+s-1}{t-1}\binom{r}{t} = \binom{k-r}{s}\binom{r}{k-r-s+1}.$$

$\square$

We now give the algorithm.

The algorithm is in Figure 1. In the algorithm we randomly uniformly choose a partition $V = V_1 \cup V_2$ where $v_0 \in V_1$. This is done T times for each vertex $v_0 \in V$. If $p = v_0, v_1, \ldots, v_k$ is simple path then by Lemma 4, the probability that no partition satisfies $|V(p)\cap V_1| = r$ and $|E(p)\cap E_2| = s$ is at most

$$\left(1 - 2^{-k}\binom{r}{k-r-s+1}\binom{k-r}{s}\right)^T \leq \frac{1}{4}.$$

Algorithm Hamiltonian($G(V,E)$,k,r,s).

For every $v_0 \in V$
  For $i = 1$ to $T := 2^{k+1} / \left(\binom{r}{k-r-s+1} \binom{k-r}{s} \right)$
    Choose a random uniform partition $V = V_1 \cup V_2$ where $v_0 \in V_1$
    Build the circuit $f = F_{k,r,s}^{v_0,V_1,V_2}(x,y,z)$ using Lemma 1
    Test if $\Phi(f) = \sum_{S\subseteq[s]} \sum_{R\subseteq[r]} f\left(x, \sum_{i\in R} y^{(i)}, \sum_{j\in S} z^{(j)}\right) \equiv 0$.
    If $\Phi(f) \not\equiv 0$ answer "YES" and halt.
Answer "NO"

Figure 1: *An algorithm for generating .*

Then by Lemma 1, $f = F_{k,r,s}^{v_0,V_1,V_2}(x,y,z)$ can be constructed in $poly(n)$ time. By Lemma 2 f has a multilinear monomial. Then, by Lemma 3, this can be tested with probability at least 3/4. Therefore, if there is a simple path then the algorithm answer "YES" with probability at least 1/2. If there is no simple path then, by Lemma 2, for every $v_0 \in V$ and every partition $V_1 \cup V_2$, $f = F_{k,r,s}^{v_0,V_1,V_2}(x,y,z)$ has no multilinear monomial. By Lemma 3, $\Phi(f) \equiv 0$ and the answer is "NO" with probability 1.

This proves the following

Lemma 5. *Let G be undirected graph. Algorithm* **Hamiltonian** $(G(V,E)$,k,r,$s)$ *runs in time*

$$O\left(\frac{2^{r+s+k} \cdot poly(n)}{\binom{r}{k-r-s+1}\binom{k-r}{s}} \right) \tag{3}$$

and satisfies the following. If G contains a simple path of length k then **Hamiltonian** $(G(V,E)$,k,r,$s)$ *answer "YES" with constant probability. If G contains no simple path of length k then* **Hamiltonian**$(G(V,E)$,k,r,$s)$ *answer "NO" with probability* 1.

Now to minimize (3) we choose $r = \lfloor 0.5 \cdot k \rfloor$ and $s = \lfloor 0.208 \cdot k \rfloor$ and get the result.

References

[1] A. Björklund. Determinant Sums for Undirected Hamiltonicity. FOCS 2010, pp. 173–182, (2010).

[2] R. A. DeMillo and R. J. Lipton. A probabilistic remark on algebraic program testing, *Inform. Process Lett.* 7, pp. 193-195, (1978).

[3] A. Björklund, T. Husfeldt, P. Kaski, M. Koivisto. Narrow sieves for parameterized paths and packings. *arXiv*:1007.1161v1, (2010).

[4] J. T. Schwartz, Fast probabilistic algorithms for verification of polynomial identities, *J. Assoc. Comput. Mach.* 27, pp. 701-717, (1980).

[5] R. Zippel. An Explicit Separation of Relativised Random Polynomial Time and Relativised Deterministic Polynomial Time. *Inf. Process. Lett.* 33(4). pp. 207–212 (1989).

ECCC ISSN 1433-8092

http://eccc.hpi-web.de