An Algorithm for finding Hamilton Cycles in a Random Graph

by
B. Bollobás*, T.I. Fenner**, A.M. Frieze***

* Department of Pure Mathematics, University of Cambridge, Cambridge, England

** Department of Computer Science, Birkbeck College, University of London, England

*** Department of Computer Science and Statistics, Queen Mary College, University of London, London, England

Abstract

This paper describes a polynomial time algorithm HAM that searches for hamilton cycles in undirected graphs. On a random graph its asymptotic probability of success is that of the existence of such a cycle. If all graphs with n vertices are considered equally likely, then using dynamic programming on failure leads to an algorithm with polynomial expected time. Finally, it is used in an algorithm for solving the symmetric bottleneck travelling salesman problem with probability tending to 1, as n tends to ∞.

Introduction

This paper describes a polynomial time algorithm HAM that searches for hamiltonian cycles in undirected graphs. As one would expect this algorithm is not perfectly reliable, i.e. a graph G may have a hamiltonian cycle but our algorithm may fail to find one. However if G is chosen at random then our algorithm has an asymptotically small probability of failure.

To be precise: let Γ_0 denote the set of graphs with vertex set $V_n = \{1,2,\ldots,n\}$ and m edges.

We turn Γ_0 into a probability space by giving each $G \in \Gamma_0$ the probability $1/|\Gamma_0| = 1/\binom{N}{m}$ where $N = \binom{n}{2}$. Let $G_{n,m}$ denote a graph chosen randomly from Γ_0.

Now let

(1.1) $m = n\log n/2 + n\log\log n/2 + c_n n$ for some sequence c_n.

Komlós and Szemerédi [8] have shown that

$$\lim_{n\to\infty} \Pr(G_{n,m} \text{ is hamiltonian}) = \begin{cases} 0 & \text{if } c_n \to -\infty \\ e^{-e^{-2c}} & \text{if } c_n \to c \\ 1 & \text{if } c_n \to +\infty \end{cases}$$

$$= \lim_{n\to\infty} \Pr(G_{n,m} \text{ has minimum degree at least } 2).$$

Their proof is essentially non-constructive (see also Bollobás [1] or Frieze[6] for alternative non-constructive proofs). The best previous polynomial time algorithm is due to Shamir [10] whose algorithm, HAM1 say, satisfies

$$\lim_{n\to\infty} \Pr(\text{HAM1 finds a hamiltonian cycle in } G_{n,m}) = 1 \quad \text{if } c_n \geq (3/2 + \epsilon)\log\log n \text{ for } \epsilon > 0 \text{ fixed.}$$

We first improve this to obtain essentially the best possible result.

Theorem 1.1

(a) Let m be defined as in (1.1). Then

$$\lim_{n\to\infty} \Pr(\text{HAM finds a hamiltonian cycle in } G_{n,m}) = \begin{cases} 0 & \text{if } c_n \to -\infty \\ e^{-e^{-2c}} & \text{if } c_n \to c \\ 1 & \text{if } c_n \to \infty \end{cases}$$

(b) HAM runs in $O(n^{4+\epsilon})$ time.

(Note that result (a) cannot be improved, although (b) possibly could)

We next consider the case where each of the 2^N graphs with vertices V_n is equally likely to be chosen. Under this model the probability of failure is so small that, if we apply dynamic programming [7] when HAM fails, we obtain the following

Theorem 1.2

There is an algorithm for solving the hamiltonian cycle problem with polynomial expected running time.

Our algorithm also has application in solving the symmetric bottleneck travelling salesman problem(BTSP). An instance of BTSP is specified by the assignment of a numerical weight to the edges of a complete graph K_n on n vertices. The objective is to find a hamiltonian circuit for which the maximum edge-weight is minimised.

Let us assume that edge-weights are drawn independently from the uniform distribution over [0,1]. Karp and Steele [9] remark that Shamir's algorithm can be used to find a near optimal solution with probability tending to 1. (Throughout this paper all limits are taken as $n \to \infty$ and this is implied if it is not explicitly stated). A modification of our proof of Theorem 1.1 gives

Theorem 1.3

There is a polynomial time algorithm BOT satisfying

$$\lim_{n\to\infty} \Pr(\text{BOT solves BTSP exactly}) = 1$$

Algorithm HAM

The following idea has been used by many authors: given a path $P = (v_1,v_2,\ldots v_k)$ plus an edge $e = \{v_k,v_i\}$ where $1 \le i \le k-2$, we can create another path of length k-1 by deleting edge $\{v_i,v_{i+1}\}$ and adding e. Thus let

$\text{ROTATE}(P,e) = (v_1,v_2,\ldots,v_i,v_k,v_{k-1},\ldots,v_{i+1})$.

The algorithm we describe is based on ideas in the proof used in [6]. It proceeds by a sequence of stages. At the beginning of the k^{th} stage we have a path P_k of length k, with endpoints w_0 and w_1. We try to extend P_k from either w_0 or w_1. If we fail but $\{w_0,w_1\} \in E$ then connectivity tells us that we can find a longer path. Failing this, we do a sequence of rotations which creates new paths that we can try to extend or close. We apply the same construction to all these paths and so on until either we have succeeded in obtaining a path of length k+1 or we have exceeded a certain length of rotation sequence. We now give a formal description:

Algorithm HAM

```
Input: a connected graph G = (V_n,E) of minimum degree at least 2.
begin
    let  P_1 be the path (1,w) where w = min{v: {1,v} ∈ E};
    k:=1
L1  begin {stage k begins here}
      Q_1:=P_k; s:=1; t:=1; δ(Q_1):=0;
{Remark: δ(Q_s) is the number of rotations in the sequence constructing Q_s from Q_1}
      repeat
        let path Q_s have endpoints w_0, w_1 where w_0 < w_1;
        for i = 0,1  do
        begin
          Suppose that the edges incident with w_i and not contained in
          Q_s are {w_i,x_1},...{w_i,x_p} where x_1<x_2<...<x_p;
          for j = 1 to p do

          if x_j is not on Q_s then
          begin
            P_k+1:= Q_s + {w_i,x_j};    {extension}
            k:=k+1;goto L1
```

```
          end
          else if x_j = w_{1-i} then
          begin
            let C be the cycle Q_s+{w_0,w_1};
            if C is a hamiltonian cycle then terminate successfully
            else
            begin
              starting from w_0, let u be the first vertex along Q_s which is
              adjacent to some vertex not in C; let v be the lowest
              numbered neighbour of u not in C and let u_1 and u_2 be the
              neighbours of u on C where u_1 < u_2, then
              P_{k+1}:= C+{u,v}-{u,u_1};k:=k + 1;
              goto L1 {cycle extension}
            end
          end else
          begin
            t:=t+1; Q_t:=ROTATE(Q_s,{w_i,x_j}); δ(Q_t):=δ(Q_s)+1
          end
          {next j}
        end;{next i}
        s:=s+1
      until δ(Q_s) > 2T+1;
          {where T = ⌈logn/(logd-loglogd⌉+1 and d=2m/n}
    terminate unsuccessfully
  end
end
```

We now introduce some notation used in the analysis of HAM. Suppose that HAM terminates unsuccessfully in stage k on input G. Let

$END(G) = \{ v:$ there exists, in stage k, a path Q_s with v as an endpoint and $\delta(Q_s) = t, 1 \leq t \leq T\}$

For $x \in END(G)$ let

$END(G,x) = \{ v:$ there exists, in stage k, a path Q_s with x,v as endpoints and $\delta(Q_s) = t, 1 \leq t \leq 2T\}$

We note that

(2.1) G cannot contain an edge $\{x,y\}$ where $x \in END(G)$ and $y \in END(G,x)$.

Consider P_k, the initial path in stage k. It is the final path in a sequence $P^{(0)}=P_1$, $P^{(1)}$, $P^{(2)}, \ldots P^{(M)}=P_k$ where $P^{(i+1)}$ is obtained from $P^{(i)}$ by a single extension, cycle extension, or rotation.

Let $W(G) = \{$ edges in $P^{(1)}, P^{(2)}, \ldots P^{(M)} \} \cup \{\{w_0,w_1\}:$ HAM executes a cycle extension on a path with endpoints w_0 and $w_1\}$

For $X \subseteq E$ let $G_X = (V_n, E-X)$, we can then deduce

Lemma 2.1

Suppose that HAM terminates unsuccessfully in stage k on input G. If $X \subseteq E-W(G)$ then HAM will also terminate unsuccessfully in stage k on G_X.

[On input G_X HAM will actually generate P_k at the start of stage k via the same sequence $P^{(0)}, P^{(1)}, \ldots P^{(M)}$]

The following inequality is straightforward:

(2.2) $\qquad |W(G)| \leq n(2T+2)$

Proof of Theorems 1.1 and 1.2

We say that an event A_n, depending on n, occurs almost surely (a.s.) if $\lim_{n\to\infty} Pr(A_n) = 1$.

We now prove a structural lemma concerning $G=G_{n,m}$. Let $d = 2m/n$ as in HAM.
A vertex is small if $deg(v) \leq d/20$ and large otherwise.
For $S \subseteq V_n$ let $N(S,G)=\{w \in V_n-S:$ there exists $v \in S$ such that $\{v,w\} \in E\}$.

Lemma 3.1

The following statements hold a.s., provided $c_n \rightarrow -\infty$:

(a) $G_{n,m}$ contains no more than $n^{1/3}$ small vertices.

(b) $G_{n,m}$ does not contain 2 small vertices at a distance of 4 or less apart.

(c) $G_{n,m}$ contains no vertex of degree exceeding 5d.

(d) There does not exist a set of large vertices S with $|S| \leq n/d$ and $|N(S,G)| \leq d|S|/300$.

Proof

It is much easier to work with the independent model $G_{n,p}$ which is a random graph with vertices V_n, in which each possible edge is independently included with probability p and excluded with probability 1-p. It is well known that if $p = m/N$ then $G_{n,m}$ and $G_{n,p}$ have similar properties. We shall calculate with $G_{n,p}$, $p = m/N$ and translate our results to $G_{n,m}$.

Let $E_{n,p}$ denote the (random) set of edges in $G_{n,p}$. We note first that $|E_{n,p}|$ is distributed as a binomial random variable with parameters N,p and that conditional on $|E_{n,p}| = m$, $G_{n,p}$ is distributed as $G_{n,m}$.

It follows from Stirling's inequalities for factorials that

$$(3.1)\ \Pr(\ |E_{n,p}| = m\ ) = (1-o(1))(N/(2\pi m(N-m)))^{1/2} \geq (1-o(1))(2/\pi N)^{1/2}.$$

Also for a graph property A

$$(3.2)\ \Pr(G_{n,p} \text{ has } A) = \sum_{m'} \Pr(G_{n,p} \text{ has } A \mid |E_{n,p}|=m')\ \Pr(|E_{n,p}|=m')$$

$$= \sum_{m'} \Pr(G_{n,m} \text{ has } A)\ \Pr(|E_{n,p}|=m')$$

Thus from (3.1) and (3.2) we have

$$(3.3)\ \Pr(G_{n,m} \text{ has } A) \leq (1+o(1))(\pi N/2)^{1/2}\Pr(G_{n,p} \text{ has } A).$$

In our proof we will often use non-integral quantities where we should really round up or down. It will be clear that such aberrations do not affect the validity of our arguments.

(a) If $G_{n,p}$ has $\geq n^{1/3}$ small vertices then there exists a set S, $|S| = n^{1/3}$ such that each vertex of S is adjacent to no more than d/20 vertices of V_n-S. Thus

$$(3.4)\ \Pr(G_{n,p} \text{ has } \geq s{=}n^{1/3} \text{ small vertices}) \leq \binom{n}{s} \Big(\sum_{k=0}^{d/20} \binom{n-s}{k} p^k(1-p)^{(n-s-k)}\Big)^s$$

$$\leq (ne/s)^s\ (c((n-s)20ep/d)^{d/20}\ \exp(-19d/20 + sd/n))^s$$

$$(3.5)\quad \leq \exp(\ -n^{1/3}d/12)$$

using $d \geq \log n$.

(Note that the summation in (3.4) is dominated by its last term)

Thus, using (3.3)

$$\Pr(G_{n,m} \text{ has } \geq n^{1/3} \text{ small vertices}) = O(n\ \exp(-n^{1/3}d/12)) = o(1).$$

(b) Let A_1 denote '2 small vertices at a distance $\leq$ 4 apart'. Then

$$\Pr(G_{n,p} \text{ has } A_1) \leq \binom{n}{2} \Big(\sum_{k=0}^{d/20} \binom{n-2}{k} p^k(1-p)^{n-k-2}\Big)^2(n^3p^4+n^2p^3+np^2+p)$$

$$(3.6)\quad \leq n^5p^4e^{-1.5d}.$$

For $p \geq 2\log n/n$ we can use (3.3) and (3.6). For smaller p in the range $\log n/n \leq p \leq 2\log n/n$ we need a bit more work. It follows from (3.2) that there exists m', $m-(n\log^3 n)^{1/2} \leq m' \leq m$ such that

$$\Pr(G_{n,m'} \text{ has } A_1) \leq 3n^5p^4e^{-1.5d}.$$

Now $G_{n,m}$ is obtained from $G_{n,m'}$ by adding m-m' random edges. Thus

$$\Pr(G_{n,m} \text{ has } A_1) \leq \Pr(G_{n,m'} \text{ has } A_1) + \pi_1\Pr(G_{n,m'} \text{ does not have } A_1)$$

where $\pi_1 = \Pr$(one of the m-m' added edges meets a vertex that is within distance 1 of a small vertex)

That π_1 is o(1) follows from (a) and $m-m' \leq (n\log^3 n)^{1/2}$.

(c) $$\Pr(G_{n,p} \text{ has a vertex of degree } \geq 5d) \leq n \binom{n}{5d} p^{5d} \leq n(e/5)^{5d}.$$

Now use (3.3).

(d) We prove the result for $G=G_{n,p}$, the result for $G_{n,m}$ follows from (3.3). For a set $K \subseteq V_n$ we define the event

A_K = '$|N(K,G)| \leq \alpha|K|d$' where $\alpha = 1/300$.

We first consider |K| large. Suppose first that $n^{.1} \leq k=|K| \leq n/d$. We prove a stronger result than needed. Now

Pr(there exists K, |K| = k and A_K) $\leq$

$$\Delta = \binom{n}{k} \sum_{t=0}^{\alpha kd} \binom{n-k}{t} p_1{}^t\ (1-p_1)^{n-k-t}$$

Where $p_1 = (1-(1-p)^k) \lessapprox kp \lessapprox 1$ is the probability that a vertex not in K is adjacent to at least one vertex of K, if |K| = k.

For large n, $\alpha kd \leq (n-k)p_1/2$ and so for some constant c>0

$\Delta \le c\binom{n}{k}\binom{n-k}{\alpha kd} p_1{}^{\alpha kd}(1-p)^{k(n-k-\alpha kd)}$

$\le c(ne/k)^k\ (ne/\alpha kd)^{\alpha kd}\ (kp)^{\alpha kd}\ \exp(-kd+k+\alpha kd)$

$\le c((ne^2/k)\ (e^2/\alpha)^{\alpha d}\ e^{-d})^k,$

$= O(n^{-\gamma})$ for any constant $\gamma > 0$,

provided $n^{.1} \le k \le n/d$.

For $1 \le k \le \min(n^{.1}, n/d)$ we use two methods of proof which cover the range of possibilities. We first assume that $p = O(n^{-.75})$.

If there exists a set K of large vertices such that A_K occurs then, by considering $T = K \cup N(K,G)$, there exists a set T, $|T| = t$, $d/20 \le t \le n^{.1}(1+\alpha d)$ which contains at least $2t$ edges. Then for some constant $c>0$

$$\Pr(\text{there exists such a } T) \le c\sum_t \binom{n}{t}\binom{t^2/2}{2t} p^{2t}\ (1-p)^{t^2/2-2t}$$

$$\le \sum_t (ne^4tp^2/16)^t$$

$$= O(n^{-\gamma}) \qquad \text{for any constant } \gamma > 0,$$

We finally consider $p \ge n^{-.8}$. We independently orient the edges of G randomly to obtain a digraph G', i.e. if $\{u,v\} \in E_{n,p}$ then we direct from u to v with probability 1/2 and from v to u with probability 1/2.

Let B be the event 'there exists $v \in V_n$ such that v is large in G but has outdegree $\le d/50$ in G'. Since $d \ge n^{.2}$ we find

$\Pr(B) = O(n^{-\gamma})$ for any constant $\gamma > 0$.

Suppose then that B does not occur and A_K occurs for some small set K of large vertices. Then there exists a set K all of whose vertices have outdegree $\ge d/50$ in G' for which the outdegree of the set K is no more than $\alpha|K|d$. Then if $k = |K|$

$$\Pr(\text{there exists such a } K) \le \binom{n}{k}\binom{n}{\alpha kd}\left(\binom{\alpha kd}{d/50}(p/2)^{d/50}\right)^k$$

$$= O(n^{-\gamma}) \qquad \text{for any constant } \gamma > 0.$$

Let $\Gamma_1 = \{G \in \Gamma_0$: G is connected, has minimum degree at least 2 and satisfies all the conditions listed in Lemma 3.1$\}$

Suppose that HAM terminates unsuccessfully in stage k on $G_{n,m}$. Now let $X \subseteq E$ be <u>deletable</u> if:

(i) no edge of X meets a small vertex;

(ii) no large vertex meets more than $d/1000$ edges of X;

(iii) $X \cap W(G) = \Phi$.

<u>Lemma 3.2</u>

Suppose HAM terminates unsuccessfully in stage k on $G = G_{n,m} \in \Gamma_1$. Suppose $X \subseteq E$ is deletable. Then for, n large,

(3.6) $|END(G_X)| \ge n/1000$;

(3.7) $|END(G_X,x)| \ge n/1000$ for $x \in END(G_X)$.

<u>Proof</u>

Consider the execution of HAM on G_X. From Lemma 2.1 we know that HAM will start stage k with the same P_k as for G and terminate unsuccessfully in this stage.

Suppose P_k has endpoints w_0 and w_1. Let

$S_t = \{v$: v is large (in G) and there exists a path Q_s with endpoints w_0, v such that $\delta(Q_s) = t\}$.

We prove (3.6) by showing that

$$(3.8)\quad \left|\bigcup_{i=1}^{T} S_t\right| \ge n/1000.$$

We show first that, for n large, $S_1 \ne \Phi$. Let $x_0, x_1, \ldots, x_p$ be the neighbours of w_1 in G_X where $\{x_0, w_1\}$ is an edge of $P_k = Q_1$. Let y_i be the endpoint, other than w_0, of $\mathrm{ROTATE}(P_k,\{w_1,x_i\})$ for $i=1,2,\ldots p$.

Case 1: w_1 is small.

Then $p \ge 1$ as $G \in \Gamma_1$ and X is deletable. Also y_1 is large (Lemma 3.1(b)) and so $y_1 \in S_1$.

Case 2: w_1 is large.

$p \ge d/20 - d/1000 - 1$ for n large. Also at most one of $y_1, y_2, \ldots, y_p$ can be small (Lemma 3.1(b)) and so $|S_1| \ge p-1 > 0$ for n large.

We show next that, for n large,

(3.9) $|S_t| \le n/d$ implies $|S_{t+1}| \ge d|S_t|/1000$.

For each vertex $v \in S_t$ choose one path $Q_{s(v)}$ with endpoints w_0 and v such that $\delta(Q_{s(v)}) = t$. Consider now pairs (v,w) where $v \in S_t$ and $w \in W(v) = N(\{v\},G_X)$. If $\{v,w\}$ is not an edge of $Q_{s(v)}$ let $x(v,w)$ be the endpoint of ROTATE$(Q_{s(v)},\{v,w\})$ other than w_0. If $x(v,w)$ is large then $x \in S_{t+1}$. Let

$$\alpha(v,w) = 1 \text{ if } \begin{cases} \text{(a) } \{v,w\} \text{ is an edge of } Q_{s(v)} \\ \text{OR} \\ \text{(b) } x = x(v,w) \text{ is small} \\ \text{OR} \\ \text{(c) } \{x,w\} \text{ is not an edge of } P_k. \end{cases}$$

$$= 0 \quad \text{otherwise.}$$

Now for each $v \in S_t$ there are at most t+2 w's such that $\alpha(v,w) = 1$ (1 for each of (a) and (b) and t for (c) as $Q_{s(v)}$ is obtained from P_k by t rotations and hence contains at most t edges not in P_k). On the other hand, for each $w \in N(S_t,G_X)$ there can be at most 2 $x \in S_{t+1}$ such that for some $v \in S_t$, $x = x(v,w)$ and $\alpha(v,w) = 0$, since x will be a neighbour of w on P_k. Thus

$$\begin{aligned}
|S_{t+1}| &= |\{ x(v,w): v \in S_t,\ w \in W(v) \text{ and } x(v,w) \text{ is large}\}| \\
&\ge |\{ x(v,w): v \in S_t,\ w \in W(v) \text{ and } \alpha(v,w) = 0\}| \\
&\ge |\{ w \in N(S_t,G_X): \text{there exists } v \in S_t \text{ with } \alpha(v,w) = 0\}|/2 \\
&\ge (|N(S_t,G_X)| - (t+2)|S_t|)/2 \\
&\ge (|N(S_t,G)| - (d/1000 + t + 2)|S_t|)/2 \\
&\ge (d/300 - (d/1000 + T + 2))|S_t|/2 \\
&\ge d|S_t|/999 \qquad \text{for n large.}
\end{aligned}$$

Since $S_1 \ne \Phi$ and (3.9) holds, we know that for some $\tau \le T-1$ that $|S_\tau| \ge n/d$. Let $S' \subseteq S_\tau$ be of size $\lfloor n/d \rfloor$. Applying the same argument as used to prove (3.9), using S' in place of S_τ we have

$$|S_{\tau+1}| \ge d|S'|/999 \ge n/1000 \qquad \text{for n large.}$$

This verifies (3.6). To prove (3.7) consider $x \in END(G_X)$, choose a path $Q = Q_\sigma$ having x as one of its endpoints and $\delta(Q_\sigma) \le T$ and then redefine

$S_t = \{v$: v is large (in G) and there exists a path Q_s with endpoints x,v such that $\delta(Q_s) \ge \delta(Q)\}$.

Now apply the argument used to prove (3.9), using Q in place of P_k, to prove (3.7).

We can now prove Theorem 1.1.

Now it is known (see for example Erdös and Rényi [3]) that if $c_n \to -\infty$ then $G_{n,m}$ is a.s. connected and in general

$$\Pr(G_{n,m} \text{ has a vertex of degree } 1) \approx 1-e^{-e^{-2c_n}}$$

Thus if $c_n \to -\infty$, $G_{n,m}$ a.s. has a vertex of degree 1 and so there is nothing to prove. If $c_n \to -\infty$ then, using Lemma 3.1, we have

$$(3.10) \quad |\Gamma_1| = (1-o(1))e^{-e^{-2c_n}} |\Gamma_0|$$

Now let $\Gamma_2 = \{ G: G \in \Gamma_1$ and HAM terminates unsuccessfully on G$\}$. It follows from (3.10) that to prove Theorem 1.1 we need only show that

$$(3.11) \quad \lim_{n\to\infty} |\Gamma_2|/|\Gamma_0| = 0$$

To prove (3.11) we use a colouring argument developed in Fenner and Frieze [5]. Let now $w = \lceil \lambda d \rceil$ for some constant $\lambda > 0$. For each $G \in \Gamma_0$ let

(G,j), $j = 1,2,\ldots, J = \binom{m}{w}$ enumerate all the possible ways of colouring w edges of G green and the remaining m-w edges blue. Let $X=X(G,j)$ denote the set of green edges. Let

$$a(G,j) = 1 \quad \text{if} \quad \begin{cases} \text{(3.12a) HAM fails on G and } G_X; \\ \text{(3.12b) there does not exist } e = \{x,y\} \in X \text{ such that } x \in END(G_X) \text{ and } y \in END(x); \\ \text{(3.12c) } |END(G_X)| \ge n/1000 \text{ and } |END(G_X,x)| \ge n/1000 \text{ for all } x \in END(G_X). \end{cases}$$

$$= 0 \quad \text{otherwise}$$

We show first that for $G \in \Gamma_2$

$$(3.13)\qquad \sum_{j=1}^{J} a(G,j) \geq (1-o(1))\binom{m_1}{w} \quad \text{where } m_1 = m-(2T+2)n$$
$$= (1-o(1))m.$$

To see this let $G\in\Gamma_2$ and let HAM terminate unsuccessfully in stage k on G. As $G\in\Gamma_1$, it follows from (2.1), Lemma 2.1 and Lemma 3.2 that if $X=X(G,j)$ is deletable then $a(G,j)=1$. Let $G'=(V',E')$ be the subgraph of G induced by the large vertices and those edges not in $W(G)$. Then $|V'| \geq n-n^{1/3}$ and $|E'|=m'\geq m-n(2T+2)$. The number of deletable sets is the number of ways of choosing w edges from E' subject to the condition that no vertex in V' has more than d/1000 of its incident vertices chosen. Using Lemma 3.1(c) it is not difficult to show that this is $(1-o(1))\binom{m'}{w}$, which implies (3.13).

(Choose edges of E' independently with probability $4\lambda/n$. One almost surely chooses more than w edges. Furthermore the number of edges chosen incident with a given vertex is dominated stochastically by a binomial random variable with parameters $\lceil 5d\rceil$ and $4\lambda/n$).

On the other hand, let H be a fixed graph with vertices V_n and $m-w$ edges. Let
$b(H) = |\{(G,j)\colon H=G_X,\ G\in\Gamma_0 \text{ and } a(G,j) = 1\}|$

We see that

$$(3.14)\quad b(H) \leq \binom{N'-m+w}{w} \qquad \text{where } N' = \binom{n}{2} - \binom{\lceil n/1000\rceil}{2}$$

If (3.12a) or (3.12b) do not hold for H (replace G_X by H in these statements) then $b_H=0$. Given (3.12a), (3.12b) there are at most $N'-m+w$ edges to choose from in order to ensure (3.12c).

Now

$$\sum_{G\in\Gamma_0}\sum_{j=1}^{J} a(G,j) = \sum_H b(H)$$

Thus

$$\begin{aligned}(1-o(1))|\Gamma_2| &\leq \sum_{G\in\Gamma_2}\sum_{j=1}^{J} a(G,j) && \text{by (3.13)}\\
&\leq \sum_{G\in\Gamma_0}\sum_{j=1}^{J} a(G,j)\\
&= \sum_H b(H)\\
&\leq \binom{N'-m+w}{w}\binom{N}{m-w} && \text{by (3.14)}\end{aligned}$$

$$\begin{aligned}\text{Thus } |\Gamma_2|/|\Gamma_0| &\leq (1+o(1))\binom{N'-m+w}{w}\binom{N}{m-w}\Big/\left(\binom{m_1}{w}\binom{N}{m}\right)\\
&\leq (1+o(1))((N'-m+w)(m-w)/((N-m+w)(m_1-w)))^w\\
&\leq (1+o(1))e^{-w(N-N')/N}\,(1+o(1))^w\\
(3.15)\qquad &\leq e^{-\lambda d/1000001} && \text{for n large}\end{aligned}$$

We can take any constant value $\lambda>0$ here and this will complete the proof of Theorem 1.1(a). To prove part (b) we note that on $G\in\Gamma_1$ HAM executes $O(n(5d)^{2T}) = O(n^{3+\epsilon})$ rotations. Thus, as given, HAM runs in time $O(n^{4+\epsilon})$ with probability $1-o(1)$. We can easily make it run in time $O(n^{4+\epsilon})$ by imposing a suitable time limit.

We now turn to the proof of Theorem 1.2. If all graphs are equally likely to be chosen then this is the same model as $G_{n,p}$, $p=1/2$. We use (3.2) where property A will mean that the given graph is connected, has minimum degree at least 2 and yet HAM terminates <u>unsuccessfully</u>. We will show that

(3.16) $\Pr(G_{n,1/2} \text{ has } A) = o(1/2^n)$.

Since Dynamic Programming requires time $O(n^2 2^n)$, this will prove the theorem.

Using the Chernoff bound [4] for the tails of the Binomial Distribution we see that

(3.17) $\Pr(|\,|E_{n,1/2}| - n^2/4\,| \geq (n^3\log n)^{1/2}) = o(1/2^n)$.

Thus, using (3.2), we need only prove

(3.18) $\Pr(G_{n,m'} \text{ has } A) = o(1/2^n)$ for $|m'-n^2/4| < (n^3\log n)^{1/2}$.

Letting Γ_i, $i=0,1,2$ refer to $G_{n,m'}$ we define

$\Gamma_1' = \{G \in \Gamma_0 : G \text{ does not satisfy all the conditions of Lemma 3.1}\}$

and

$\Gamma_A = \{ G\in\Gamma_0\colon$ G is connected, has minimum degree at least 2 and yet HAM terminates unsuccessfully on G$\}$.

Then $\Pr(G_{n,m'} \text{ has } A) = |\Gamma_A|/|\Gamma_0|$
$\leq |\Gamma_1'|/|\Gamma_0| + |\Gamma_A-\Gamma_1'|/|\Gamma_0|$

$= |\Gamma_1'|/|\Gamma_0| + |\Gamma_2|/|\Gamma_0|$.

Now let $d' = 2m'/n = (1-o(1))n/2$ in our range of interest. We see that, for large n, conditions (b) and (d) of Lemma 3.1 are always true for $G_{n,m'}$. It follows from (3.3),(3.5) and (3.6) that

$$|\Gamma_1'|/|\Gamma_0| = \Pr(G_{n,m'}\in\Gamma_1') = O(n\exp(-n^{4/3}/25) + n^6\exp(-.74n))$$
$$= o(1/2^n).$$

Putting $\lambda = 2000002$ in (3.15) shows that $|\Gamma_2|/|\Gamma_0| = o(e^{-n})$ and this completes the proof of Theorem 1.2.

Algorithm BOT

We turn now to BTSP. Given an instance of this problem, let the edges of K_n be ordered $e_1,e_2,\dots e_n$ where $c(e_1) \le c(e_2) \le \dots \le c(p_N)$ -- $c(e)$ is the 'cost' of edge e. Let $E_t = \{e_1,e_2,\dots,e_t\}$ and let $G_t = (V_n,E_t)$. Note that G_t has the same distribution as $G_{n,t}$.

Algorithm BOT

begin

let $\mu = \min\{t\colon G_t$ has mimimum degree at least 2$\}$;

apply HAM to G_μ

end

It is clear that if HAM terminates successfully on G_μ then BOT solves BTSP exactly. We cannot apply Theorem 1.1 directly as G_μ as defined in BOT above has a slightly different distribution to $G_{n,\mu}$ conditional on minimum degree 2.

Let $m' = \lfloor n\log n/2 + n\log\log n/2 - n\log\log\log n/2\rfloor$ and $m'' = m'+\lceil n\log\log\log n\rceil$. It is known that $G_{m'}$ a.s. is connected and has minimum degree 1 and that $G_{m''}$ a.s. has minimum degree 2. Thus $m' < \mu \le m''$ a.s..

Now for $m' < m \le m''$ define the events:

A_m = 'G_m is connected and satisfies the conditions of Lemma 3.1, where $d=\log n$',

B_m = 'A_m and G_m has minimum degree at least 2',

C_m = 'G_m is connected, has minimum degree at least 2 and HAM terminates unsuccessfully on G_m'.

Then, where, $M = \{m\colon m' < m \le m''\}$,

$$\Pr(\text{ BOT fails}) = \Pr\Big(\bigcup_{m\in M} C_m\Big) + o(1)$$

$$\le \Pr\Big(\Big(\bigcup_{m\in M} C_m\Big) \cap \Big(\bigcap_{m\in M} A_m\Big)\Big) + \Pr\Big(\bigcup_{m\in M}\bar{A}_m\Big) + o(1)$$

$$\le \sum_{m\in M}\Pr(C_m \cap A_m) + \Pr\Big(\bigcup_{m\in M}\bar{A}_m\Big) + o(1)$$

$$= \sum_{m\in M}\Pr(C_m \cap B_m) + \Pr\Big(\bigcup_{m\in M}\bar{A}_m\Big) + o(1)$$

$$(4.1)\qquad \le \sum_{m\in M}\Pr(C_m \mid B_m) + \Pr\Big(\bigcup_{m\in M}\bar{A}_m\Big) + o(1)$$

For $x\in\{a,b,c,d\}$ let

$A_m(x)$ = 'G_m satisfies condition x of Lemma 3.1'

and let

D_m = 'G_m is connected'.

Now

$$(4.2)\quad \Pr\Big(\bigcup_{m\in M}\bar{D}_m\Big) = \Pr(\bar{D}_{m'+1}) = o(1)$$

The calculations in Lemma 3.1 show that

$$(4.3)\qquad \Pr\Big(\bigcup_{m\in M}(\bar{A}_m(a) \cup \bar{A}_m(d))\Big) = O(n^{-\alpha}) \qquad \text{for any constant } \alpha > 0.$$

(Although Lemma 3.1 specifically excludes $c_n \to -\infty$, the calculations are still valid for $c_n \geq -\log\log\log n$.)

(4.4) $Pr(\bigcup_{m \in M} \bar{A}_m(c)) = Pr(\bar{A}_{m''}(c)) = o(1)$

By considering the addition of the m+1'st edge we obtain

$$Pr(\bar{A}_{m+1}(b) \cap A_m(b) \cap A_m(a) \cap A_m(c)) = O((\log n)^4 n^{-4/3})$$

Thus

$$Pr(\bigcup_{m \in M} (\bar{A}_{m+1}(b) \cap A_m(b) \cap A_m(a) \cap A_m(c)) = o(1)$$

It then follows from (4.3) that

$$Pr(\bigcup_{m \in M} (\bar{A}_{m+1}(b) \cap A_m(b))) = o(1)$$

and hence

(4.5) $Pr(\bigcup_{m \in M} \bar{A}_m(b)) \leq Pr(\bar{A}_{m'+1}(b)) + Pr(\bigcup_{m \in M} (\bar{A}_{m+1}(b) \cap A_m(b))$

(4.2) - (4.5) yield

(4.6) $Pr(\bigcup_{m \in M} \bar{A}_m) = o(1)$

Now in the notation used to prove Theorem 1.1, we have

$$\begin{aligned} Pr(C_m | B_m) &= |\Gamma_2|/|\Gamma_1| \\ &= (|\Gamma_2|/|\Gamma_0|)(|\Gamma_0|/|\Gamma_1|) \\ &\leq n^{-\lambda/1000001} \log n \end{aligned}$$

for any constant $\lambda > 0$, using (3.15) and (3.10). By choosing $\lambda = 2000002$ we obtain

(4.7) $\sum_{m \in M} Pr(C_m | B_m) = o(1)$.

Theorem 1.3 now follows from (4.1), (4.6) and (4.7).

Conclusions

The results of this paper show that the hamiltonian cycle problem can be considered to be well-solved in a probabilistic sense. They can be extended to cover the problem of finding a number of disjoint hamiltonian cycles by following the approach described in Bollobás and Frieze [2].

References

[1] B. Bollobás, 'Almost all regular graphs are hamiltonian', European Journal of Combinatorics 4 (1983) 311-316.

[2] B.Bollobás and A.M.Frieze, 'On matchings and hamilton cycles in random graphs', Annals of Discrete Mathematics, to appear.

[3] P.Erdős and A.Rényi, 'On the strength of connectedness of a random graph', Acta. Math. Acad. Sci. Hungar. 12 (1961) 261-267.

[4] P.Erdős and J.Spencer, 'Probabilistic methods in combinatorics', Academic Press, New York and London.

[5] T.I. Fenner and A.M. Frieze, 'On the Existence of Hamiltonian Cycles in a Class of Random Graphs', Discrete Mathematics 45(1983) 301-305.

[6] A. M. Frieze, 'Limit distribution for the existence of hamiltonian cycles in a random bipartite graph' to appear in the European Journal of Combinatorics.

[7] M. Held and R. M. Karp, ' A dynamic programming approach to sequencing problems', SIAM Journal on Applied Mathematics 10 (1962) 196-210.

[8] M. Komlós and E. Szemerédi, 'Limit distribution for the existence of a hamiltonian cycle in a random graph', Discrete Mathematics 43 (1983) 55-63.
[9] R. M. Karp and J. M. Steele, 'Probabilistic analysis of travelling salesman problem, to appear in a forthcoming book on the travelling salesman problem.
[10] E. Shamir, 'How many edges are needed to make a random graph hamiltonian?', Combinatorica (1983).