HUMBOLDT-UNIVERSITÄT ZU
BERLIN

An improved inclusion-exclusion algorithm for counting Hamiltonian paths

K. Dohmen

Informatik-Bericht Nr. 125

INFORMATIK-
BERICHTE

Herausgeber:	Professoren des Institutes für Informatik
Redaktion:	Publikationsstelle, Tel. (+49 30) 2093 3114
Druckerei:	Humboldt-Universität zu Berlin
ISSN:	0863 - 095

Die Reihe Informatik-Berichte erscheint aperiodisch.

Dr. Klaus Dohmen
Humboldt-Universität zu Berlin
Institut für Informatik
Sitz: Rudower Chaussee 25
Unter den Linden 6
10099 Berlin
e-mail: dohmen@informatik.hu-berlin.de

Mai 1999

An improved inclusion-exclusion algorithm for counting Hamiltonian paths

Klaus Dohmen
Institut für Informatik
Humboldt-Universität zu Berlin
Unter den Linden 6
D-10099 Berlin, Germany
e-mail: dohmen@informatik.hu-berlin.de

AMS Subject Classification (1991): 05A15, 05C45, 68R10

Abstract

We propose a modification of Karp's inclusion-exclusion algorithm for counting Hamiltonian paths in graphs. Experiments on random graphs suggest that this modification leads to a significant reduction of the average running time when the input graphs are sparse.

1 Introduction

Let G be a directed or undirected finite graph with two distinguished vertices s and t. The *Hamiltonian Path Counting Problem*, which is well-known to be #P-complete, asks for the number of directed resp. undirected paths in G starting at s and ending in t and containing each of the vertices in G exactly once. Such paths are usually referred to as *Hamiltonian paths.*

Karp [3] and later Bax [1] gave an exact algorithm for the Hamiltonian Path Counting Problem, based on the principle of inclusion-exclusion. In both cases, the average running time is dominated by the number of terms in the inclusion-exclusion formula. In fact, this number is equal to 2^n, where n is the number of vertices in G different from s and t.

In this paper, we propose a modification of the counting procedure which does not make use of all terms in the inclusion-exclusion formula. A similar approach has been proposed by Lozinski [4] for the problem of counting satisfying assignments of propositional formulae in conjunctive normal form.

Experiments on random graphs suggest that this modification leads to a reduction of the average running time when the input graphs are sparse.

2 The algorithm

The principle of inclusion-exclusion, which is also known as the sieve formula, states that for any subsets $A_1, \ldots, A_n$ of some finite universe U,

$$(1) \qquad \left| \bigcap_{i=1}^{n} \complement A_i \right| = \sum_{I \subseteq \{1,\ldots,n\}} (-1)^{|I|} \left| \bigcap_{i \in I} A_i \right| ,$$

where $\complement A_i = U \setminus A_i$, and $|\cdot|$ denotes cardinality.

In order to apply the principle of inclusion-exclusion to the Hamiltonian Path Counting Problem for some directed or undirected graph G having vertices $1, \ldots, n$ and two additional vertices s and t, we define U as the set of all directed resp. undirected walks of length $n+1$ from s to t, and A_i as the set of those walks in U that lack vertex i.

We thus have the following algorithm:

Algorithm 1 Standard algorithm for counting Hamiltonian paths

Require: G is a directed or undirected graph with vertices $1, \ldots, n; s, t$
Ensure: acc is the number of Hamiltonian s,t-paths in G

1: $acc \leftarrow 0$
2: **for all** subsets I of $\{1, \ldots, n\}$ **do**
3: $\quad h \leftarrow$ number of s,t-walks of length $n+1$ in G lacking each $i \in I$
4: $\quad acc \leftarrow acc + (-1)^{|I|}\, h$
5: **end for**

The quantity on the right-hand side of line 3 can be computed using a dynamic programming technique as proposed by Karp [3], or alternatively, by computing powers of the adjacency matrix of G as proposed by Bax [1]. Since this part of the computation can be carried out efficiently, we won't go into the details and instead refer to Bax [1] and Karp [3]. Note that the number of terms produced by Algorithm 1 (and hence its running time) is the same for all graphs on n vertices.

In combinatorial topology, with each family of sets $\mathcal{A} = (A_1, \ldots, A_n)$ one usually associates its *nerve*, that is, the abstract simplicial complex

$$\mathcal{N}(\mathcal{A}) := \left\{ I \subseteq \{1, \ldots, n\} \;\middle|\; \bigcap_{i \in I} A_i \neq \emptyset \right\} .$$

By the simple observation that the sum in (1) can be restricted to the nerve of $(A_1, \dots, A_n)$, we are led to the following modification of Algorithm 1:

Algorithm 2 Improved algorithm for counting Hamiltonian paths

Require: G is a directed or undirected graph with vertices $1, \dots, n; s, t$
Ensure: *acc* is the number of Hamiltonian s, t-paths in G

```
acc ← 0
Gen ← {∅}
while Gen ≠ ∅ do
    NewGen ← ∅
    for all I in Gen do
        h ← number of s,t-walks of length n+1 in G lacking each i ∈ I
        if h > 0 then
            acc ← acc + (−1)^|I| h
            if I ≠ ∅ then
                m ← max I
            else
                m ← 0
            end if
            for j = m+1 to n do
                NewGen ← NewGen ∪ {I ∪ {j}}
            end for
        end if
    end for
    Gen ← NewGen
end while
```

3 Experimental results

Consider the random graph model $\mathbf{G}_{n+2,p}$ obtained from the complete undirected graph having vertex-set $\{1, \dots, n\} \cup \{s, t\}$ by independently removing each edge with probability $1 - p$. For each pair $(n, p) \in \{1, \dots, 12\} \times \{0.1, 0.3, 0.5\}$ we generated a sample of size 1000 from $\mathbf{G}_{n+2,p}$ and calculated the mean $\overline{N}^{(2)}_{n,p}$ and empirical standard deviation $S^{(2)}_{n,p}$ of the number of terms produced by Algorithm 2. The obtained values of $\overline{N}^{(2)}_{n,p}$ and $S^{(2)}_{n,p}$ are shown in Table 1, along with the number $N^{(1)}_n$ of terms produced by Algorithm 1. Since this quantity is the same for all graphs on n vertices, $N^{(1)}_n$ does not depend on p. In fact, $N^{(1)}_n = 2^n$.

Table 1 Means and empirical standard deviations

n	$N_n^{(1)}$	$\overline{N}_{n,0.5}^{(2)}$	$S_{n,0.5}^{(2)}$	$\overline{N}_{n,0.3}^{(2)}$	$S_{n,0.3}^{(2)}$	$\overline{N}_{n,0.1}^{(2)}$	$S_{n,0.1}^{(2)}$
1	2	1.0	0.0	1.0	0.0	1.0	0.0
2	4	2.6	1.5	1.9	1.4	1.4	1.0
3	8	3.2	1.9	1.8	1.4	1.1	0.5
4	16	10.4	6.1	6.5	6.8	2.8	4.8
5	32	18.2	7.4	8.2	8.0	1.6	3.0
6	64	47.0	18.8	27.2	25.5	7.9	19.2
7	128	93.1	23.9	44.1	34.5	5.4	16.0
8	256	213.5	50.7	131.1	92.6	28.7	75.7
9	512	432.6	60.5	243.1	129.3	26.8	74.3
10	1024	923.3	129.8	621.4	324.6	133.6	320.0
11	2048	1864.9	155.1	1214.7	484.9	144.4	341.9
12	4096	3879.2	307.0	2779.6	1089.23	577.9	1259.9

The numerical values suggest that Algorithm 2 requires much fewer terms on the average than Algorithm 1, provided p is small. Therefore, Algorithm 2 is likely to have a less average-case complexity for sparse graphs.

4 Concluding remarks

We proposed a modification of the standard inclusion-exclusion algorithm for counting Hamiltonian paths in graphs. Empirically, this modification turned out to decrease the average-case complexity for sparse graphs. A theoretical analysis of the modified algorithm is left for future work.

The modification was achieved by restricting the sum to the nerve of the underlying sets. This raises the question, wether other abstract simplicial complexes would be appropriate as well. For instance, when counting proper vertex colorings, the broken circuit complex is appropriate (cf. [7, 8]).

A general framework for dealing with such questions is the emerging theory of discrete abstract tubes, founded by Naiman and Wynn [5, 6] in 1992. For some recent applications of this theory to combinatorics and reliability theory, the reader is referred to the present author [2].

References

[1] E.T. Bax, *Inclusion and exclusion algorithm for the Hamiltonian path problem*, Inform. Process. Lett. **47** (1993), 203–207.

[2] K. Dohmen, *Improved inclusion-exclusion identities and inequalities based on a particular class of abstract tubes*, Electron. J. Probab. **4** (1999), paper no. 5.

[3] R.M. Karp, *Dynamic programming meets the principle of inclusion-exclusion*, Oper. Res. Lett. **1** (1982), 49–51.

[4] E.L. Lozinski, *Counting propositional models*, Inform. Process. Lett. **41** (1992), 326–332.

[5] D.Q. Naiman and H.P. Wynn, *Inclusion-exclusion-Bonferroni identities and inequalities for discrete tube-like problems via Euler characteristics*, Ann. Statist. **20** (1992), 43–76.

[6] ———, *Abstract tubes, improved inclusion-exclusion identities and inequalities and importance sampling*, Ann. Statist. **25** (1997), 1954–1983.

[7] H. Whitney, *A logical expansion in mathematics*, Bull. Amer. Math. Soc. **38** (1932), 572–579.

[8] H.S. Wilf, *Which polynomials are chromatic?*, Proc. Colloq. Combinatorial Theory (Rome), 1973.