

Algorithms, Operators and Transfer Matrices

Richard P. Brent

MSI & RSISE
ANU

11 October 2007

Outline

Average-case analysis of algorithms uses techniques that may have applications in other areas, such as statistical mechanics. To illustrate this, we consider the analysis of one nontrivial algorithm, the binary Euclidean algorithm.

- The binary Euclidean algorithm – results of Brent & Maze
- Ruelle operators – results of Vallée
- Possible connection with Transfer Matrices

Notation

$\lg(x)$ denotes $\log_2(x)$.

N, n, u, v are positive integers.

$\text{Val}_2(u)$ denotes the dyadic valuation of the positive integer u , i.e. the greatest integer j such that $2^j \mid u$.

The Binary Euclidean Algorithm

The idea of the *binary* Euclidean algorithm is to avoid the “division” operation $r \leftarrow m \bmod n$ of the classical algorithm, but retain $O(\log N)$ worst (and average) case.

We assume that the algorithm is implemented on a binary computer so division by a power of two is easy. In particular, we assume that the “shift right until odd” operation

$$u \leftarrow u/2^{\text{Val}_2(u)}$$

or equivalently

$$\text{while even}(u) \text{ do } u \leftarrow u/2$$

can be performed in constant time, although time $O(\text{Val}_2(u))$ would be sufficient.

Definition

There are several almost equivalent ways to define the algorithm. It is easy to take account of the largest power of two dividing the inputs, so for simplicity we assume that u and v are *odd* positive integers.

Following is a simplified version of the algorithm given in Knuth, §4.5.2.

Algorithm B

B1. $t \leftarrow |u - v|$;
if $t = 0$ terminate with result u

B2. $t \leftarrow t/2^{\text{Val}_2(t)}$

B3. if $u \geq v$ then $u \leftarrow t$ else $v \leftarrow t$;
go to B1.

A Heuristic Continuous Model

To analyse the expected behaviour of Algorithm B, we can follow what Gauss did for the classical algorithm. This was first attempted in my 1976 paper¹ and there is a summary in Knuth (Vol. 2, *third* edition, §4.5.2).

Assume that the initial inputs u_0, v_0 to Algorithm B are uniformly and independently distributed in $(0, N)$, apart from the restriction that they are odd. Let (u_n, v_n) be the value of (u, v) after n iterations of step B3.

Let

$$x_n = \frac{\min(u_n, v_n)}{\max(u_n, v_n)}$$

and let $F_n(x)$ be the probability distribution function of x_n (in the limit as $N \rightarrow \infty$). Thus $F_0(x) = x$ for $x \in [0, 1]$.

¹R. P. Brent, Analysis of the Binary Euclidean Algorithm, *New Directions and Recent Results in Algorithms and Complexity*, (J. F. Traub, editor), Academic Press, New York, 1976, 321–355.

Plausible Assumption

We make the plausible assumption² that $\text{Val}_2(t)$ takes the value k with probability 2^{-k} at step B2.

It is plausible because $\text{Val}_2(t)$ at step B2 depends on the least significant bits of u and v , whereas the comparison at step B3 depends on the most significant bits, so one would expect the steps to be (almost) independent.

²Not proved by Brent (1976), but later justified by Vallée's results (*c.* 1998).

The Recurrence for F_n

Consider the effect of steps B2 and B3. We can assume that $u > v$ so $t = u - v$. If $\text{Val}_2(t) = k$ then $X = v/u$ is transformed to

$$\begin{aligned} X' &= \min \left(\frac{u - v}{2^k v}, \frac{2^k v}{u - v} \right) \\ &= \min \left(\frac{1 - X}{2^k X}, \frac{2^k X}{1 - X} \right). \end{aligned}$$

It follows that $X' < x$ iff

$$X < \frac{1}{1 + 2^k/x} \quad \text{or} \quad X > \frac{1}{1 + 2^k x}.$$

Thus, the recurrence for $\tilde{F}_n(x) = 1 - F_n(x)$ is

$$\begin{aligned} \tilde{F}_{n+1}(x) &= \\ &\sum_{k \geq 1} 2^{-k} \left(\tilde{F}_n \left(\frac{1}{1 + 2^k/x} \right) - \tilde{F}_n \left(\frac{1}{1 + 2^k x} \right) \right) \end{aligned}$$

and $\tilde{F}_0(x) = 1 - x$ for $x \in [0, 1]$.

The Recurrence for f_n

Differentiating the recurrence for \tilde{F}_n we obtain (formally) a recurrence for the probability density $f_n(x) = F'_n(x) = -\tilde{F}'_n(x)$:

$$\begin{aligned} f_{n+1}(x) &= \sum_{k \geq 1} \left(\frac{1}{x + 2^k} \right)^2 f_n \left(\frac{x}{x + 2^k} \right) \\ &+ \sum_{k \geq 1} \left(\frac{1}{1 + 2^k x} \right)^2 f_n \left(\frac{1}{1 + 2^k x} \right). \end{aligned}$$

Operator Notation

The recurrence for f_n may be written as

$$f_{n+1} = \mathcal{B}_2 f_n,$$

where the operator \mathcal{B}_2 is the case $s = 2$ of a more general operator \mathcal{B}_s which will be defined later.

Convergence Results

In my 1976 paper I gave numerical and analytic evidence that $F_n(x)$ converges to a limiting distribution $F(x)$ as $n \rightarrow \infty$, and that $f_n(x)$ converges to the corresponding probability density $f(x) = F'(x)$ (note that $f = \mathcal{B}_2 f$ so f is a “fixed point” of the operator \mathcal{B}_2).

Proofs of these results were eventually given by Gérard Maze³. However, today I will discuss a different (and slightly earlier) approach due to Brigitte Vallée.

³G. Maze, Existence of a limiting distribution for the binary GCD algorithm, *J. of Discrete Algorithms* 5 (2007), 176–186.

Expected number of iterations

The expected number of iterations of Algorithm B is $\sim K \lg N$ as $N \rightarrow \infty$, where $K = 0.705\dots$ is a constant defined by

$$K = \ln 2 / E_\infty ,$$

and

$$E_\infty = \ln 2 + \int_0^1 \left(\sum_{k=2}^{\infty} \left(\frac{1 - 2^{-k}}{1 + (2^k - 1)x} \right) - \frac{1}{2(1+x)} \right) F(x) dx .$$

We can simplify the expression for K to obtain

$$K = 2/b ,$$

where

$$b = 2 - \int_0^1 \lg(1-x) f(x) dx \approx 2.833 .$$

Another Formulation – Algorithm V

It will be useful to rewrite Algorithm B in the following equivalent form (using pseudo-Pascal):

Algorithm V { Assume $u \leq v$ }

```
while  $u \neq v$  do  
  begin  
    while  $u < v$  do  
      begin  
         $j \leftarrow \text{Val}_2(v - u)$ ;  
         $v \leftarrow (v - u)/2^j$ ;  
      end;  
     $u \leftrightarrow v$ ;  
  end;  
return  $u$ .
```

Continued Fractions

Vallée (*Algorithmica*, 1998) shows a connection between Algorithm V and continued fractions of a certain form:

$$\frac{u}{v} = 1/a_1 + 2^{k_1}/a_2 + 2^{k_2}/\dots/a_r + 2^{k_r},$$

where a_j is odd, $k_j > 0$, and $0 < a_j < 2^{k_j}$.

Some Useful Operators

Operators $\mathcal{B}_s, \mathcal{U}_s, \tilde{\mathcal{U}}_s, \mathcal{V}_s$, useful in the analysis of the binary Euclidean algorithm, are defined on suitable function spaces by

$$\mathcal{U}_s[f](x) = \sum_{k \geq 1} \left(\frac{1}{1 + 2^k x} \right)^s f \left(\frac{1}{1 + 2^k x} \right), \quad (1)$$

$$\tilde{\mathcal{U}}_s[f](x) = \left(\frac{1}{x} \right)^s \mathcal{U}_s[f] \left(\frac{1}{x} \right), \quad (2)$$

$$\mathcal{B}_s = \mathcal{U}_s + \tilde{\mathcal{U}}_s,$$

$$\mathcal{V}_s[f](x) = \sum_{k \geq 1} \sum_{\substack{a \text{ odd,} \\ 0 < a < 2^k}} \left(\frac{1}{a + 2^k x} \right)^s f \left(\frac{1}{a + 2^k x} \right). \quad (3)$$

In these definitions s is a complex variable, and the operators act linearly on certain function spaces (in fact Hardy spaces $\mathcal{H}^2(\mathcal{D})$ where \mathcal{D} is a suitable open disk).

The case $s = 2$ is of particular interest. \mathcal{B}_2 encodes the effect of one iteration of the inner “while” loop of Algorithm V, and \mathcal{V}_2 encodes the effect of one iteration of the outer “while” loop.

History and Notation

\mathcal{B}_2 (denoted T) was introduced in my 1976 paper and was generalised to \mathcal{B}_s by Vallée. \mathcal{V}_s was introduced by Vallée. We shall call

- \mathcal{B}_2 the *binary Euclidean operator* and
- \mathcal{V}_s *Vallée's operator*.

In the context of dynamical systems⁴, \mathcal{V}_s is called the *Ruelle operator* relative to the system. The generating functions of interest involve the quasi-inverse operator

$$\Lambda_s = (I - \mathcal{V}_s)^{-1} .$$

Vallée studied partial sums of coefficients of these Dirichlet series, and her results come from the application of Tauberian theorems due to Delange.

⁴David Ruelle, *Thermodynamic Formalism*, Addison-Wesley, 1978.

Relation Between the Operators

The operators are closely related, as the following results show.

Lemma 1

$$\mathcal{V}_s = \mathcal{V}_s \tilde{\mathcal{U}}_s + \mathcal{U}_s.$$

The Lemma can be proved algebraically, and there is also a nice algorithmic interpretation in the case $s = 2$.

The following Theorem gives a simple relationship between \mathcal{B}_s , \mathcal{V}_s and \mathcal{U}_s . The proof is immediate from Lemma 1 and the definitions of the operators.

Theorem 1

$$(\mathcal{V}_s - \mathcal{I})\mathcal{U}_s = \mathcal{V}_s(\mathcal{B}_s - \mathcal{I}).$$

Fixed Points

It follows immediately from Theorem 1 that, if

$$g = \mathcal{U}_2 f,$$

then

$$(\mathcal{V}_2 - \mathcal{I})g = \mathcal{V}_2(\mathcal{B}_2 - \mathcal{I})f.$$

Thus, if f is a fixed point of the operator \mathcal{B}_2 , then g is a fixed point of the operator \mathcal{V}_2 . From a result of Vallée we know that \mathcal{V}_2 , acting on a certain Hardy space $\mathcal{H}^2(\mathcal{D})$, has a unique positive dominant simple eigenvalue 1, so g must be (a constant multiple of) the corresponding eigenfunction (provided $g \in \mathcal{H}^2(\mathcal{D})$). Also, from the definitions of \mathcal{B}_2 and \mathcal{U}_2 , we have

$$\lambda = f(1) = 2g(1)$$

which is useful for proving the consistency of two of the expressions for K given below.

Some Results of Vallée

Using her operator \mathcal{V}_s , Vallée proved that

$$K = \frac{2 \ln 2}{\pi^2 g(1)} \sum_{\substack{a \text{ odd,} \\ a > 0}} 2^{-\lfloor \lg a \rfloor} G\left(\frac{1}{a}\right)$$

where g is a nonzero fixed point of \mathcal{V}_2 (i.e. $g = \mathcal{V}_2 g \neq 0$) and $G(x) = \int_0^x g(t) dt$. This is the only expression for K which has been rigorously proved.

Because \mathcal{V}_s can be proved to have nice spectral properties, the existence and uniqueness (up to scaling) of g can be proved rigorously.

A Conjecture of Vallée

Let $\lambda = f(1)$, where f is the limiting probability density (conjectured to exist) as above. Vallée (see Knuth, third edition, §4.5.2(61)) conjectured that

$$\frac{\lambda}{b} = \frac{2 \ln 2}{\pi^2},$$

or equivalently that

$$K = \frac{4 \ln 2}{\pi^2 \lambda}. \quad (4)$$

Vallée proved the conjecture under the assumption that the operator \mathcal{B}_s satisfies a certain spectral condition.

Numerical Results

Using an improvement of the “discretization method” of my 1976 paper, and the MP package with the equivalent of more than 50 decimal places (50D) working precision, I computed the limiting probability density f , then K , $\lambda = f(1)$, and $K\lambda$. The results were

$$\begin{aligned} K &= 0.7059712461\ 0191639152\ 9314135852\ 8817666677 \\ \lambda &= 0.3979226811\ 8831664407\ 6707161142\ 6549823098 \\ K\lambda &= 0.2809219710\ 9073150563\ 5754397987\ 9880385315 \end{aligned}$$

These are believed to be correctly rounded values.

Vallée’s conjecture (4) is that

$$K\lambda = 4 \ln 2 / \pi^2 .$$

The computed value of $K\lambda$ agrees with $4 \ln 2 / \pi^2$ to 40 decimals!

Open Problems

Since the work of Vallée and Maze, analysis of the average behaviour of the binary Euclidean algorithm has a rigorous foundation. However, some interesting open questions remain.

For example, does the binary Euclidean operator \mathcal{B}_2 have a unique positive dominant simple eigenvalue 1? Vallée has proved the corresponding result for her operator \mathcal{V}_2 .

In order to estimate the speed of convergence of f_n to f (assuming f exists), we need more information on the spectrum of \mathcal{B}_2 . What can be proved? Preliminary numerical results indicate that the sub-dominant eigenvalue(s) are a complex conjugate pair:

$$\lambda_2 = \bar{\lambda}_3 = 0.1735 \pm 0.0884i ,$$

with $|\lambda_2| = |\lambda_3| = 0.1948$ to 4D.

Vallée has proved related results for some other algorithms (variants of the Euclidean algorithm, algorithms for computing the Jacobi symbol), but many analogous questions remain open.

Transfer Matrices

Transfer matrices are used in statistical mechanics to generate series expansions for certain models. The argument z of the series is a parameter of the model (e.g. temperature). (Of course, there may be several parameters.)

Usually we need to find or approximate the largest one or two eigenvalues of an $N \times N$ matrix, and we are interested in the thermodynamic limit as $N \rightarrow \infty$. In the limit, we are not strictly dealing with matrices, but with linear operators.

Compare our discussion of the binary Euclidean algorithm: the operators \mathcal{B}_s and \mathcal{V}_s are linear operators on certain function spaces, and to obtain numerical results (as in the numerical verification of Vallée's conjecture) we approximate these operators by (large) matrices. In some cases we know that the operators have an dominant eigenvalue 1, but we are interested in whether there is a “spectral gap”, i.e. whether the other eigenvalues λ satisfy $|\lambda| \leq c < 1$.

Research Topic

Is the analogy between transfer matrices and Vallée's operator useful ? Can we transfer some of the techniques used in analysis of algorithms to statistical mechanics (or *vice versa*) ?

References

- [1] Richard P. Brent, Analysis of the Binary Euclidean Algorithm, *New Directions and Recent Results in Algorithms and Complexity*, (J. F. Traub, editor), Academic Press, New York, 1976, 321–355.
- [2] Richard P. Brent, *Further analysis of the Binary Euclidean algorithm*, Report PRG-TR-7-99, Oxford University Computing Laboratory, Nov. 1999.
- [3] Yao-ban Chan, *Selected Problems in Lattice Statistical Mechanics*, Ph.D. thesis, Mathematics and Statistics, Univ. of Melbourne, Sept. 2005.
- [4] Donald E. Knuth, *The Art of Computer Programming, Volume 2: Seminumerical Algorithms* (third edition). Addison-Wesley, Menlo Park, 1997.

- [5] Gérard Maze, Existence of a limiting distribution for the binary GCD algorithm, *J. of Discrete Algorithms* 5, 1 (March 2007), 176–186.
- [6] David Ruelle, *Thermodynamic Formalism*, Addison Wesley, 1978.
- [7] Brigitte Vallée, The complete analysis of the Binary Euclidean Algorithm, *Proc. ANTS'98, Lecture Notes in Computer Science* **1423**, Springer-Verlag, 1998, 77–94.
- [8] Brigitte Vallée, Dynamics of the binary Euclidean algorithm: functional analysis and operators, *Algorithmica* **22** (1998), 660–685.