Prof. Friedrich Eisenbrand
Nicolai Hähnle

Discussions from: March 8, 2011

Computer Algebra

Spring 2011

Assignment Sheet 2

Warning: These are just notes and not necessarily full solutions for each exercise. Full solutions may require some additional details to be fleshed out. Please report any mistakes you may find.

Exercise 1

Let $a, b \in \mathbb{N}$ be odd numbers with $a - b = 2^k$ for some $k \in \mathbb{N}$. Assume that a and b are not coprime, that is, there is a prime p that divides both a and b. We can write $a = px$ and $b = py$ so that

$$2^k = a - b = p(x - y)$$

Comparing left and right hand sides implies that $p = 2$, but this contradicts the fact that a and b are odd. Therefore, a and b cannot have been coprime.

Exercise 2

Let $f : \mathbb{N} \to \mathbb{R}_+$ be a monotone increasing function with $f(a) + f(b) \leq f(a+b)$ and let $n = 2^k$. Then

$$\begin{aligned}
& f(1) + f(2) + f(4) + f(8) + \ldots + f(n) \\
&= \sum_{j=0}^{k-1} f(2^j) + f(2^k) \\
&\leq f(\sum_{j=1}^{k-1} 2^j) + f(2^k) \\
&= f(2^k - 1) + f(2^k) \\
&\leq f(2^k) + f(2^k) = 2f(n)
\end{aligned}$$

For the first inequality, we used the property $f(a) + f(b) \leq f(a+b)$, for the second inequality the fact that f is monotone increasing. Though note that the first property actually implies that the function is increasing, because $f(a) \leq f(a) + f(1) \leq f(a+1)$.

Exercise 3

1. Let $x = a \cdot 2^e$ and $y = b \cdot 2^f$. Then $xy = ab \cdot 2^{e+f}$, so the numbers can be multiplied using one multiplication and one addition.

Suppose $e \geq f$, then $y = (b \cdot 2^{f-e}) \cdot 2^e$ and so $x + y = (a + b \cdot 2^{f-e}) \cdot 2^e$, that is, addition can be done by adding shifted mantissas. It is usually reasonable to truncate the mantissa afterwards, to reflect the relative precisions of x and y.

2. Let $x_n = a \cdot 2^e$ be a t-bit approximation of $1/b$, that is

$$|x_n - \frac{1}{b}| \leq 2^{-t+1} \cdot \frac{1}{b} \tag{1}$$

So in fact we have an approximation of a root of $f(x) = \frac{1}{x} - b$. For Newton approximation, we approximate f by its tangent g in x_n and find the root of the tangent.

$$f'(x) = -\frac{1}{x^2}$$
$$g(x) = f(x_n) + f'(x_n) \cdot (x - x_n) = \frac{1}{x_n} - b - \frac{x - x_n}{x_n^2}$$

Setting $g(x) = 0$ gives us:

$$x = x_n + x_n^2 \cdot \left(\frac{1}{x_n} - b\right) = 2x_n - bx_n^2$$

We can efficiently compute this x; let us check how well it approximates $1/b$:

$$\begin{aligned} |x - \frac{1}{b}| = |2x_n - bx_n^2 - \frac{1}{b}| &= |\sqrt{b}x_n - \frac{1}{\sqrt{b}}|^2 \\ &\leq \left(2^{-t+1} \cdot \frac{1}{\sqrt{b}}\right)^2 \\ &= 2^{-2t+2}\frac{1}{b} \end{aligned}$$

The inequality comes from (1), multiplied by $\sqrt{b}$.

3. For a constant t_0, we can compute a t_0-bit approximation to $\frac{1}{b}$ by setting up the exponent based on the size of b and using the schoolbook method on the most significant bits of b to obtain an initial approximation of the mantissa. If the constant t_0 is large enough, then the iterated Newton approximation almost double the precision in every iteration, so that a variant of the argument in exercise 2 shows the bound on the running time.

Exercise 4

We have $a \leq 2^n$, and we can assume $b \geq 2$ (else the problem is trivial) so that $\frac{a}{b} \leq 2^{n-1}$. Let $z = a \cdot 2^e$ be an n-bit approximation of $\frac{1}{b}$, that is

$$|z - \frac{1}{b}| \leq 2^{-n+1} \cdot \frac{1}{b}$$

Then

$$|az - \frac{a}{b}| \le a|z - \frac{1}{b}| \le 2^{-n+1}\frac{a}{b} \le 1$$

In other words, we can use the previous exercise to compute a floating point approximation z of $\frac{1}{b}$, then perform floating point multiplication az, then truncate the resulting floating point number to an integer q using bit operations. The resulting value will be very close to the desired result; using trial computation of the remainder $r = a - qb$ we can easily check whether the result q needs to be adjusted and by how much.

Exercise 6

1. Show that adding an integer multiple of one basis vector to another basis vector does not change the lattice generated by the basis:

   This is essentially the same argument as for elementary column operations of a matrix in linear algebra, except that a little care needs to be taken to make sure all data is integral.

2. Let $P = \{\sum_{j=1}^{n} \lambda_j b_j \mid 0 \le \lambda_j < 1 \text{ for all } j = 1 \dots n\}$ be the *fundamental parallelepiped*. Show that $P \cap L = \{0\}$.

   Use the fact that since $b_1, \dots, b_n$ are linearly independent, every lattice point $v \in L$ has a *unique* representation of the form $v = \sum_{j=1}^{n} \lambda_j b_j$. Since in that representation $\lambda_j \in \mathbb{Z}$, the rest of the argument is elementary.

Exercise 7

1. We need to show that the μ_n are well-defined. Clearly, the set of possible choices for μ_n is non-empty because it contains zero. Furthermore, as long as the sequence does not stop, b_{2j} is always strictly below and b_{2j+1} strictly above the line L_α. It follows that adding large enough multiples of b_{n+1} to b_n eventually crosses that line, and so the possible choices for μ_{n+2} are bounded from above.

   Furthermore, if the sequence is unbounded, at least every second μ_n is non-zero, because the vector $b_n + b_{n+1}$ shows that at least one of $\mu_{n+2} \ge 1$ or $\mu_{n+3} \ge 1$ must be true. The coordinates of the b_n are non-negative and their sum is increasing — strictly increasing when μ_n is non-zero — and integer and therefore not bounded.

2. This follows directly from the previous exercise.

3. Consider the (half-open) fundamental parallelepiped spanned by b_{2j} and b_{2j+1}. A better approximation would imply that this parallelepiped contains a non-zero integer point (why?), which does not exist by the previous exercise.