# Exercise 2, Discrete Mathematics for Bioinformatics

Sascha Meiers, Martin Seeger

Winter term 2011/2012

## 2.1 Modulo Arithmetic

a) We show that $\langle a \rangle \subset \langle d \rangle$.
Since $d = \gcd(a, n)$, there is a $k \in \mathbb{N}$ such that $a = kd$. Hence, if $v \in \langle a \rangle$, i.e. $v = ai \mod n$, then $v = dki \mod n$ which implies that $v \in \langle d \rangle$.

b) We show that $\langle a \rangle \supset \langle d \rangle$.
Any element $v$ of $\langle d \rangle$ can be written as $v = di \mod n$ (*). On the other hand, $v \in \langle a \rangle$ iff $v = aj \mod n$.
We now use Bezout's lemma to find $x$, $y$, such that $ax + ny = d$. This is inserted into (*) to yield

$$v = di \mod n = (ax + ny)i \mod n = axi \mod n.$$

In other words, $v \in \langle a \rangle$. $\square$

## 2.2 Hashing

Let $x, y$ be character strings both of length $n$. Now we can interprete their characters as numbers in radix $2^p$, leading to a hash function

$$h(x) = \sum_{i=0}^{n} x_i 2^{p \cdot i} \bmod 2^p - 1$$

If $y$ is nothing else than a permutation of the characters in $x$, then especially their sum of the digits is equal, i.e.

$$\sum_{i=0}^{n} x_i = \sum_{i=0}^{n} y_i$$

Proof: $h(x) = h(y)$

$$\begin{aligned}
h(x) &= \sum_{i=0}^{n} x_i 2^{p \cdot i} \bmod 2^p - 1 && (1)\\
&= \sum_{i=0}^{n} \left(x_i 2^{p \cdot i} \bmod 2^p - 1\right) \bmod 2^p - 1 && (2)\\
&= \sum_{i=0}^{n} (x_i \bmod 2^p - 1) \Big(\underbrace{2^p \bmod 2^p - 1}_{1}\Big)^i \bmod 2^p - 1 && (3)\\
&= \sum_{i=0}^{n} x_i \bmod 2^p - 1 && (4)\\
&= \sum_{i=0}^{n} y_i \bmod 2^p - 1 && (5)\\
&= h(y) && (6)
\end{aligned}$$

## 2.3 Hashing

x

## 2.4 Expected value

Let $K$ = Number of probes accessing occupied slots. We want to show that

$$E(K) = \sum_{i=1}^{\infty} P(K \geq i).$$

Proof: We note that $K \geq 0$ and by definition $E(K) < \infty$ so that it is permitted to rearrange the sums. It follows that

$$\begin{aligned}
E(K) = \sum_{i=0}^{\infty} iP(K = i) &= && (7)\\
&= 1 \cdot P(K = 1) + 2 \cdot P(K = 2) + 3 \cdot P(K = 3) + \ldots && (8)\\
&= P(K = 1) + && (9)\\
&+ P(K = 2) + P(K = 2) + && (10)\\
&+ P(K = 3) + P(K = 3) + P(K = 3) \ldots && (11)
\end{aligned}$$

We rearrange this term by columns and obtain due to absolute convergence

$$\begin{aligned}
E(K) = P(K = 1) + P(K = 2) \quad &+ \quad P(K = 3) + \ldots + \\
+P(K = 2) \quad &+ \quad P(K = 3) + \ldots + \\
&+ \quad P(K = 3) + \ldots \\
= q_1 + q_2 + q_3 + \ldots \quad &= \quad \sum_{i=1}^{\infty} P(K \geq i). \square && (12)
\end{aligned}$$