Math464/564 O. Knill, **February 1997**

Homework 3

Due: Tuesday, February 11, 1997 in class

Topics: Combinatorics and finite probability spaces.

Remark. This HW contains 5 questions. Each question can give 10 points, so that you can get a maximal credit of 50 points for this HW.

1) (10 points) Topic: **Permutations and maps**.
A permutation of a set $X = \{1, 2, \dots, n\}$ is a map on this set which is bijective. A general map from X to X is not a permutation. What is the probability that a random map on X is a permutation?

2) (10 points) Topic: **Combinations. Random walk**.
A random walker on the line moves in each step with probability 1/2 to the right and with probability 1/2 to the left. Assume the walker starts at 0.
a) What is the probability that the walker is returning back to the origin after n steps.
b) Each step has length 1. What is the probability that the walker is in distance $k \in \mathbb{N}$ from the origin after n steps.

3) (10 points) Topic: **Scrabble**.
While playing scrabble, you get the letters $F, E, R, I, T, R, C, F, I$. How many words can you write with these letters? An example is $TERRIFFIC$.

4) (10 points) Topic: **Amino acids**.
The genetic code of all living cells is built by nucleotides. There are four kind of those T(hymine),A(denine),C(ytosine),G(uanine) which are used to build up the DNA (DeoxyriboNucleicAcid). An amino acid is specified by a sequence of three nucleotides. How many amino acids can be coded in this manner?

5) (10 points) Topic: **RSA and factoring integers by a probabilistic method**. Let n be a number and p a factor of n. Given m random numbers out of the set $X = \{1, \dots, n\}$. What is the probability that two of them are congruent modulo p? Hint: the number n is irrelevant, you can as well take random numbers modulo p.

MOTIVATING STORY FOR PROBLEM 5) The RSA scheme is a widely used encryption scheme for example used on the internet. Its security is based on the fact that it is hard to factor integers with large prime factors and that this problem is considered by so many people that nobody could resist publishing a better method. The best current factoring algorithms can handle numbers with say up 150 integers (currently nobody is able to factor 200 digit numbers which are products of two 100 digit primes). There are current challenges on the internet where prizes are given out for successes (this is also politically motivated, the last week a prize was won, see the article below).
One method for factoring integers is the Pollard ρ method. This method was used in 1980 to factor the first time the eight's Fermat number $n = F_8 = 2^{(2^8)} + 1$. The idea to find a prime factor p like here $p = 1238926361552897$ of $n = F_8$ is to produce randomly numbers $1 \leq x_k \leq n$ for example with a pseudo random number generator $x_{k+1} = x_k^2 + 3 \bmod n$ and to check whether any pair $x_k - x_l$ has a common prime factor with n by forming the product $Q_m = \prod_{j=1}^m (x_{2j} - x_j) \bmod n$ finding occasionally the greatest common divisor $GCD(Q_m, n)$. If two numbers x_k, x_l have a common prime factor p among the m elements $x_1, \dots, x_m$, then it shows up in $GCD(Q_m, n)$.

(*) If you have time and if you are interested in the extremely fascinating story of computational number theory (which has a lot of probabilistic stuff inside), try this:
How big does m have to be, in order to factor $n = 2^{(2^8)} + 1$ with probability 1/2 with the

Pollard rho method?

Here is my Mathematica implementation of the simplest Pollard -ρ method (you can find the source code on the web site of the course). The six'th Fermat number has already 20 digits and a usual "Baby method"="trying out all possible factors" up to $\sqrt{n}$ would be hopeless already. So, this primitive algorithm is already quite effective. One can improve the method.

```
FactorWithPollard[n_] := Module[{a=3, x=17, y=17, q=1},
   While[q<2,
     Do[x=Mod[x*x-a,n];
        y=Mod[y*y-a,n];
        y=Mod[y*y-a,n];
        q=Mod[q*(x-y),n],{i,20}];
     q=Mod[GCD[n,x-y],n] ];q];

FermatNumber[n_]:=2^(2^n)+1;
Example=FactorWithPollard[FermatNumber[6]]
```

Bonus question: Why is the method called ρ? Hint: look at the form of the letter ρ and relate it with what the algorithm does.

See the cite:

```
http://www.cnn.com/TECH/9701/30/encryption.reut/index.html
```

about the story of the "California student unscrambling the internet code" (January 30, 1997 in CNN) from last week!

BERKELEY, California (Reuters) – As the White House and the Internet community battle over U.S. encryption laws, **a University of California graduate student said he broke a code** said to have the strongest encryption that U.S. law allows to be exported without restrictions.

It took him a mere three and a half hours, he said.

"It shows how silly the export restrictions are because 40-bit key length is ridiculously weak," Ian Goldberg, a graduate student of computer science at the University of California at Berkeley, told Reuters.

The 40-bit encrypted message was published Tuesday morning by RSA Data Security Inc., a software firm in Redwood City, California, which developed encryption widely used on the Internet, as a challenge to code breakers.

RSA, owned by Security Dynamics Technologies Inc., is one of dozens of companies trying to get the U.S. government to loosen its restrictions on the export of encryption, which currently prohibit U.S. firms or citizens from putting encrypted code of more than 40-bits of length on the Internet unless the government is supplied a code key.

U.S. law allows encryptions of up to 56-bits if the government is given a key to the code, which it will hold in escrow in case a national security need arises.

The government has argued that distribution of encryption codes outside of the United States would impede its ability to fight drug trafficking and political terrorism. Congress is considering bills to loosen these restrictions.

But Internet users and Internet technology companies argue that the restrictions impede electronic commerce and widespread use of the Internet for many private business transactions.

Because the Internet has no national borders, anything posted on it by a U.S. based company would be considered exporting.

Goldberg used about 250 computer workstations networked together to test various computations to break the code. The university said those resources would be pretty commonly available to people in university settings.

At a data security and encryption conference being held here this week by RSA Data Security, people said Goldberg's break of the code is proof that U.S. laws need changing.

"Nobody in that room's going to trust 40-bit (cryptography) any more," said Peter Trei, senior software engineer at Process Software Corp., of Framingham, Massachusetts., as he nodded toward the San Francisco auditorium where 2,500 people were attending the cryptography conference.

The gathering included some of the world's leading experts on cryptography, and a number of panelists in presentations were openly critical of the White House policy of prohibiting export of strong cryptography.

Cryptography experts said the government policy must enable businesses to stay ahead of the capabilities of computer hackers, but that current standards do not allow this to be exported, which also can limit Internet distribution.