Strings and Encodings

Recall the general principle driving science and theoretical work:

Objects/concepts we want to study and understand

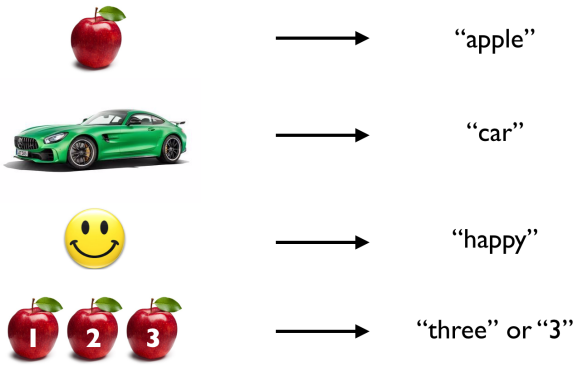
↓

Mathematical model (formal, precise definitions)

↓

Mathematically/rigorously prove facts/theorems

How do we mathematically/formally represent data?



English alphabet: $\Sigma = \{a,b,c,d,e,f,g,h,i,j,k,l,m,n,o,p,q,r,s,t,u,v,w,x,y,z\}$

Turkish alphabet: $\Sigma = \{a,b,c,ç,d,e,f,g,ğ,h,ı,i,j,k,l,m,n,o,ö,p,r,s,ş,t,u,ü,v,y,z\}$

Does the alphabet size matter?

Binary alphabet: $\Sigma = \{0, 1\}$

alphabet:

symbol/character:

string/word:

length of a string:

Σ^*:

Questions

- Does every object have a corresponding encoding?
- Can two objects have the same encoding?
- Does every string correspond to a valid encoding?

encoding:

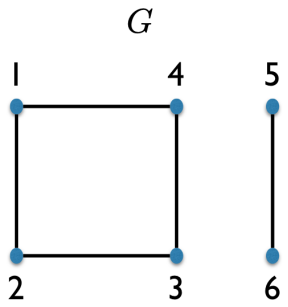
Example: $A = \mathbb{N}$

Does Σ affect encodability?

Example: $A = \mathbb{Z}$

Example: $A = \mathbb{N} \times \mathbb{N}$

Example: $A =$ all undirected graphs



Example: $A =$ all Python functions

```
def isPrime(N):
    if (N < 2):
        return False
    for factor in range(2, N):
        if (N % factor == 0):
            return False
    return True
```

Does $|\Sigma|$ matter?

Going from $|\Sigma| = k$ to $|\Sigma'| = 2$:

$A = \mathbb{N}$	**Binary**	**vs**	**Unary**
0	0		ϵ
1	1		1
2	10		11
3	11		111
4	100		1111
5	101		11111
6	110		111111
7	111		1111111
8	1000		11111111
9	1001		111111111
10	1010		1111111111
11	1011		11111111111
12	1100		111111111111

Which sets are encodable?

Function/Decision Problems

Reasonable assumptions about computation:

-
-
-
-

How can we characterize the input/output behavior of a computer?

Function Problem:

Examples of function problems:

-
-
-

How do we typically describe the last example?

More general description: A function problem is

How do we convert this into our original definition of a function problem?

Important Note:

Decision Problem:

Why?

1.

   Language:

   Important Note:

2.

   Important Note:

   Integer Factorization Problem: Given as input a natural number N, output its prime factorization.

   Decision version: