Crypto

Preliminaries and notation

- $S \subseteq \{0,1\}^*$ defines set S as a finite subset of $\{0,1\}$ all finite-length strings.
- $x \in_R S$, R indicates x is chosen randomly / uniformly from S
- U_n x is chosen from the set of all n-bit strings
- $\mu(\cdot)$ means the negligible function can take any input. Negligible functions decrease faster than inverse polynomial as n increases
- must use positive polynomial, if not, maybe it won't be negligible. We want $\mu(n) < \frac{1}{p(n)}$
- λ is an empty string

Polynomial time, Security parameter

- Polynomial time refers to the computational complexity of an algorithm with respect to the security parameter
- it means the protocol is efficient and practical because algorithms are feasible in polynomial time
- But also secure because breaking the algorithm is infeasible
- security parameter determines level of security e.g. length of keys in bits
- an algorithm running in polynomial time in the security parameter can be expressed as a polynomial function of the security parameter. That is, there exists $p(\lambda)$ such that $\mathcal{O}(p(\lambda))$
- E.g. if $\lambda = 128$ bits runs in polynomial time, running time will be a function of λ like $\mathcal{O}\lambda^2$
- In contrast to exponential time which are $\mathcal{O}2^\lambda$

Theory of Computation

- a turing machine is a theoretical device that "manipulates symbols on a strip of tape according to a table of rules" simulating algorithm logic
- Includes an infinitely long tape divided into blocks, a head can read and write symbols on the tape, a state register storing the machine state, a finite table of instructions
- We use unary 1^n, a string of 1's on the security paramter tape for reasons:
  - Unary: 1^n provides unary representation of n meaning input length corresponds with n, the longest possible representation (Worst case and Lower bound)
  - in binary, n is represented in $log_2(n)$ bits. e.g. 1000 is 11111101000 = 10 bits long. In unary, 1^{1000} is a string of 1000 ones.
  - using binary, an algorithm taking time proportional to input length would run in $\mathcal{O}(log(n))$ which runs in $\mathcal{O}(n)$
- Security parameter tape is used to model how a system scales with security parameter, 1^λ is written on it e.g. string of 1's
- This means, the function is bounded by the length of the input on the security parameter tape

Computational Indistinguishability

$$X = \{X(a,n)\}_{a \in \{0,1\}^*; n \in \mathbb{N}}$$

"A probability ensemble" is a collection or family of random variables, denoted by $X = \{X(a,n)\}$ i.e. X represents the set of $\{X(a,n)\}$.

- Set $X = \{X(a,n)\}$ defines the set of random variables X
- Set Subscript $a \in \{0,1\}^*; n \in \mathbb{N}$ defines the indexing, that is, exactly what values of a, n can take for infinitely any element in set X. e.g. $X('01', 3)$ is the index of a random variable X where a is a binary string of any finite length "0, 1, 01, 000", n is a natural number 1,2
- $\{0,1\}^*$ is the Kleene star operation, means all finite strings, an infinite set because there's no limit to the length

- $n \in \mathbb{N}$ means n can be any natural number, also an infinite set
- indexing gives us an address or a way to talk about 1 specific random variable rather than the collection
- "probability ensemble" is a term in cryptography / probability theory referring to a collection or family of probability distributions or random variables. Used to describe systems where behaviours depend on on input length, security parameter, etc
- "ensemble" means we're dealing with a collection of probabilistic objects rather than a single fixed distribution

Two probability ensembles, X, Y are computationally indistinguishabile: $X \stackrel{c}{\equiv} Y$ if

$$|\Pr[D(X(a,n)) = 1] - \Pr[D(Y(a,n)) = 1]| \leq \mu(n)$$

The (absolute difference) of the probability that the distinguisher D outputs 1 when given a sample from $X(a,n)$ and the probability D outputs 1 when given a sample from $Y(a,n)$ is $\leq$ the negl. function $\mu(\cdot)$

Notes:

- D is a PPT algorithm trying to distinguish between samples from X and Y, not guessing the bit
- $D's$ output is binary $\{0,1\}$. Output 1 means "The sample came from X" or "This is a real sample" i.e. distinguish
- We look at the difference in probability of D outputting 1 for X vs Y
- a negl. difference means D can't distinguish
- absolute value —— means we only want the magnitude of the difference, doesn't matter which is larger
- $\mu(n)$ being a negl. function means the difference becomes smaller as the security parameter grows larger
- All parties run in polynomial time in the security parameter means

Non-uniformity "Above notion of computational indistinguishability is inherently non uniform: "We allow D to be non-uniform":

Order of quantifiers for computational indistinguishability Questions - What are the tapes that are referred to? - How does the definition claim uniformity? Something to do with D Ding the random variabnles based on the indexed input