Information & Coding theory.

Lec-9

$$a_1, \dots, a_K \quad \boxed{X} \longrightarrow \boxed{Y}$$

Typical set : $A_\varepsilon^{(L)}$

Source sequence $\longleftarrow \bar{x}_L \equiv (x_1, x_2, \dots, x_L)$

$x_i \in \mathcal{X} \equiv \mathcal{A} = \{a_1, \dots, a_K\}$

code alphabet : $\mathcal{C} = \{0, 1\}$

- ~~un~~ fixed length codewords
- variable length codewords.

$a_1 \longrightarrow n_1$ $\qquad n_i \rightarrow$ denotes the length of the

$a_2 \longrightarrow n_2$ $\qquad$ codewords

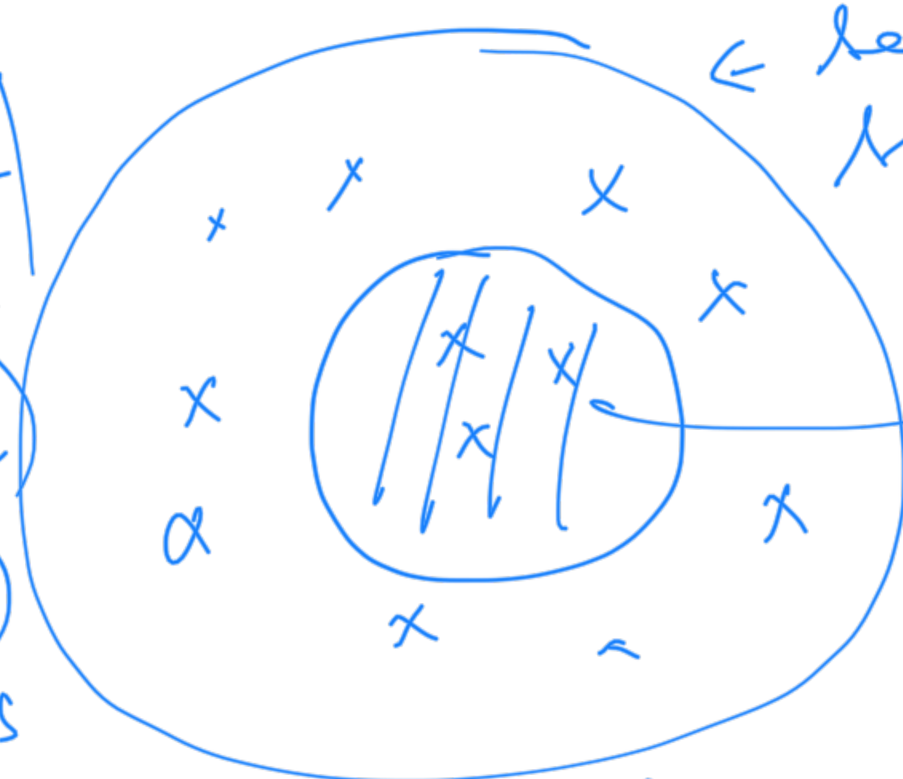
$a_2 \longrightarrow n_2$ codewords.

$\vdots$

$a_K \longrightarrow n_K$

We discussed about a bound: $n_1 = n_2 = \dots = n_K = N$.

$\left|\{(x_1, \dots, x_L) \mid x_i \in \mathcal{X}\}\right| = \boxed{K^L}$



← set of all source seq.

→ Typical set.

$\bar{X} \equiv (X_1, X_2, \dots, X_L)$
$\equiv (\underbrace{X, X, \dots, X}_{L \text{ times}})$

Memoryless source:

$\bar{x}_L \equiv (x_1, \dots, x_L)$

X_i

X_1

$X_i \overset{iid}{\sim} \boxed{\mathcal{X}}$ (source random variable) pmf $p(x) \equiv p(a_i)$, $i = 1, \dots, K$

$$\boxed{P_r(\bar{x}_L)} = \prod_{i=1}^{K} p(x_i)$$

$\varepsilon > 0$

Typical set:

$$2^{-L(H(X)+\varepsilon)} \leq p(\bar{x}_L) \leq 2^{-L(H(X)-\varepsilon)}$$

$$\boxed{p(\bar{x}_L) \approx 2^{-L H(X)}}$$

AEP

(1) $\boxed{P_r\{A_\varepsilon^{(L)}\} > 1-\varepsilon}$

$L(H(X)+\varepsilon)$

2. $|A_\varepsilon^{(L)}| \le 2^{L(H(X)+\varepsilon)}$

3. $|A_\varepsilon^{(L)}| \ge (1-\varepsilon)\, 2^{L(H(X)-\varepsilon)}$

when L is large

Understanding the importance of entropy

Goal: existence of a nonsingular binary code with the property of the codeword length.

$$\left| A_\varepsilon^{(L)} \right| \le 2^{L(H(X)+\varepsilon)}$$

We need to define a ~~bin~~ $\boxed{\text{unique}}$ binary string for each $\bar{x}_L \in A_\varepsilon^{(L)}$.

Q. How many bits are required to define a codeword / what is the minimum codeword length?

H.W.

Ans. For each sequence in $A_\varepsilon^{(L)}$ we need no more than $\boxed{L(H(X)+\varepsilon)} + 1$ bits.

$L\,(H(X)+\varepsilon)$

In order to keep the uniqueness property preserved, we can add one extra bit for a prefix but $\boxed{\text{'0'}}$ for each of such strings.

Then the total # of bits

$$= \boxed{L\,(H(X)+\varepsilon) + 2}$$

Note: For the atypical set, note that $K^L = |A|^L$

$= |X|^L$

Then for any sequence in the atypical set it is enough to consider

$$\log |X|^L \cong L \log |X| + 1$$

bits for the binary string.

Then we can prefix any such string by one more bit '1' — in order to preserve uniqueness property.

∴ total # of bits for such a sequence is ... bits.

$$\boxed{2 \log |x| + 2}.$$

H. W. Give an example existential criteria where solution is shown to exist but need not be easy to find.

$$f(x,y) = x^2y^2 + xy + y^2x + \cdots$$

$\boxed{1899}$, Hilbert 17th problem.

$$f(x,y) = (x+y)^2 + 1 \geq 0$$

$\boxed{1937}$ Artin

Goal: Give a bound for expected length of the codeword

$$\mathbb{E}\left(\ell(X^L)\right) = \sum_{\bar{x}_L \in \mathcal{X}^L} p(\bar{x}_L)\,\ell(\bar{x}_L)$$

$$\mathbb{E}\left(\frac{1}{L}\,\ell(X^L)\right) \le \boxed{H(X)+\varepsilon}$$

length

Pf. $\mathbb{E}\left(\ell(X^L)\right)$

$= \sum p(\bar{x}_L)\,\ell(\bar{x}_L)$

$$= \sum_{\bar{x}_L \in \mathcal{X}^L}$$

$$= \sum_{\bar{x}_L \in A_\varepsilon^{(L)}} p(\bar{x}_L)\, l(\bar{x}_L) + \sum_{\bar{x}_L \in A_\varepsilon^{(L)c}} p(\bar{x}_L)\, l(\bar{x}_L)$$

$$\leq \sum_{\bar{x}_L \in A_\varepsilon^{(L)}} p(\bar{x}_L) \boxed{\left(L(H(X)+\varepsilon)+2\right)} + \sum_{\bar{x}_L \in A_\varepsilon^{(L)c}} p(\bar{x}_L) \left(L \log|\mathcal{X}| + 2\right)$$

$$= \underbrace{\Pr\left(A_\varepsilon^{(L)}\right)}\left(L\left(H(X)+\varepsilon\right)+2\right)$$

$$+ \underbrace{\Pr\left(A_\varepsilon^{(L)^c}\right)}\underbrace{\left(L \log|X|+2\right)}$$

$$\overset{??}{\leq} L\left(H(X)+\varepsilon\right)+2$$

$$+ \varepsilon L\left(\log|X|\right)+2$$

$$\left[\text{Since } \underbrace{\Pr\left(A_\varepsilon^{(L)}\right) > 1-\varepsilon}\right].$$

$$\left(H(X)+\varepsilon'\right)$$

$\approx \boxed{L(H(\cdot)}$

$\varepsilon + \varepsilon \log |\mathcal{X}| + \frac{2}{2}$

$E(\ell(x)) = ?$

$$E\left(\frac{1}{L}\ell(x^L)\right) \le \boxed{H(X) + \delta}$$

Practical importance of entropy.

Thm: Can we say more if the ... fixed

Think: code is of length.

Procedure for construction of a code with the above property

Variable length code:

$p(a_i)$, $i = 1, \ldots, K$.

$$\boxed{E(C) = \sum_{i=1}^{K} p(a_i)\, n_i}$$

desired code

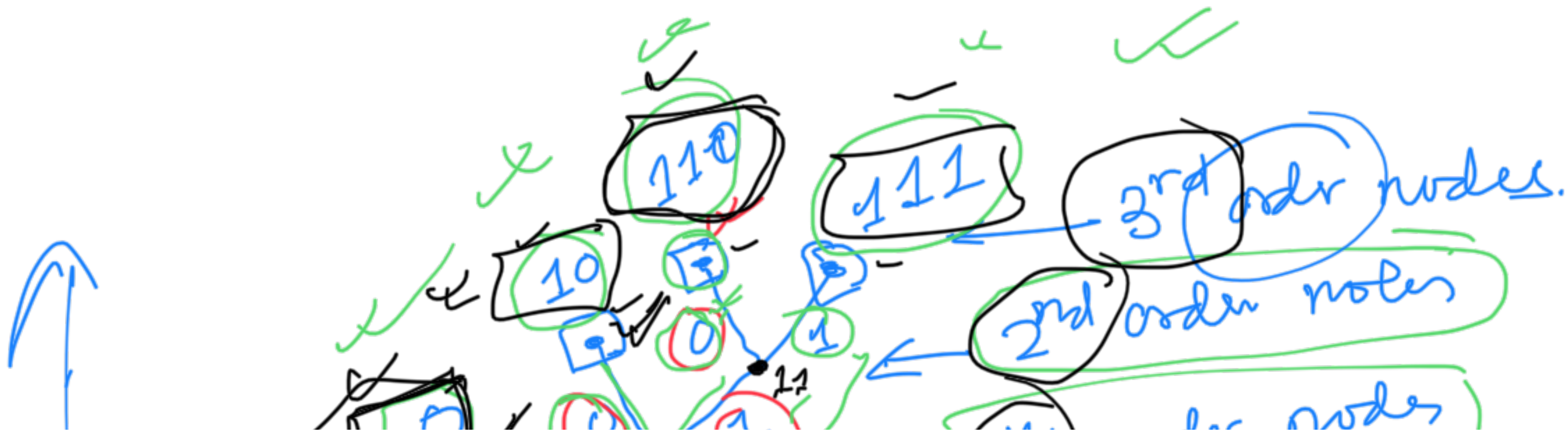
expected length of the code.

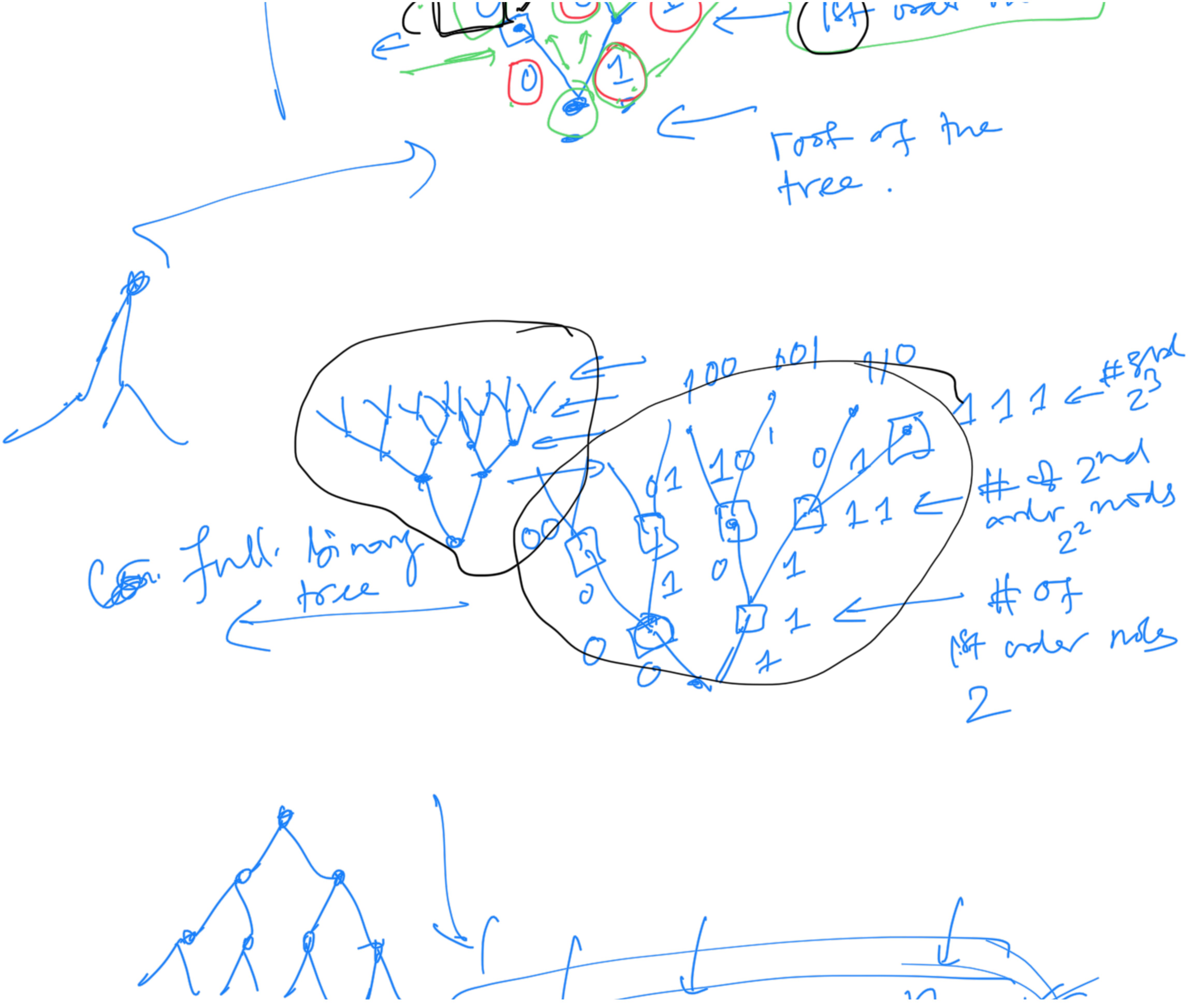
is the length of codeword

ext leng... the codeword... for a_i.

	$p(a_i)$	Code	
a_1	$1/2 = 0.5$	0	1
a_2	$1/4 = 0.25$	10	01
a_3	$1/8 = 0.125$	110	001
a_4	$1/8 = 0.125$	111	000

Tree — a connected graph without a cycle.





root of the tree .
Con. full binary tree
100 001 110
111 ← # 3rd
2^3
# of 2nd order nodes
2^2
11 ←
# of 1st order nodes
2
1 ←
00
01
10

Goal: $n_1, n_2, \ldots, n_k$

D-ary

Kraft Inequality (1949)

If $n_1, n_2, \ldots, n_k$ satisfies

$$\sum_{k=1}^{K} D^{-n_k} \leq 1 \quad \text{——} (*)$$

then there exist a prefix code with D code letters and n_k's are the codeword lengths. Converse, the codeword lengths of any prefix code satisfy $(*)$.

H. W. Can we ~~pro~~duce a code with $n_1, n_2, \ldots, n_K$ as length of codewords that satisfy (*) and it is not a prefix code?

~~Tutorial class~~

Not attending Tutorial ~~cut~~ class meaning no doubts / questions

Lec-10

No classes on "Feb 21 & 22"

10 marks.
~~or~~ viva exam. First class Test.

March 1st → 50 / 100 (20 marks)

(Tuesday)

Design of a prefix code

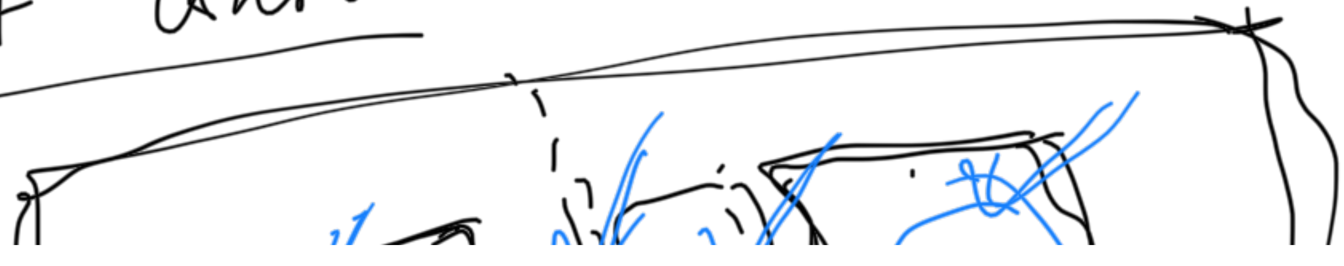
$a_1, \dots, a_K$ → source letters

$(2 \le D)$ → code letters. $n_1, \dots, n_K$

$$\sum_{k=1}^{K} \boxed{D^{-n_k}} \le 1 \leftarrow (*)$$

~~Import~~. Given: $n_1, n_2, \dots, n_K$

First consider the full tree



Observations:

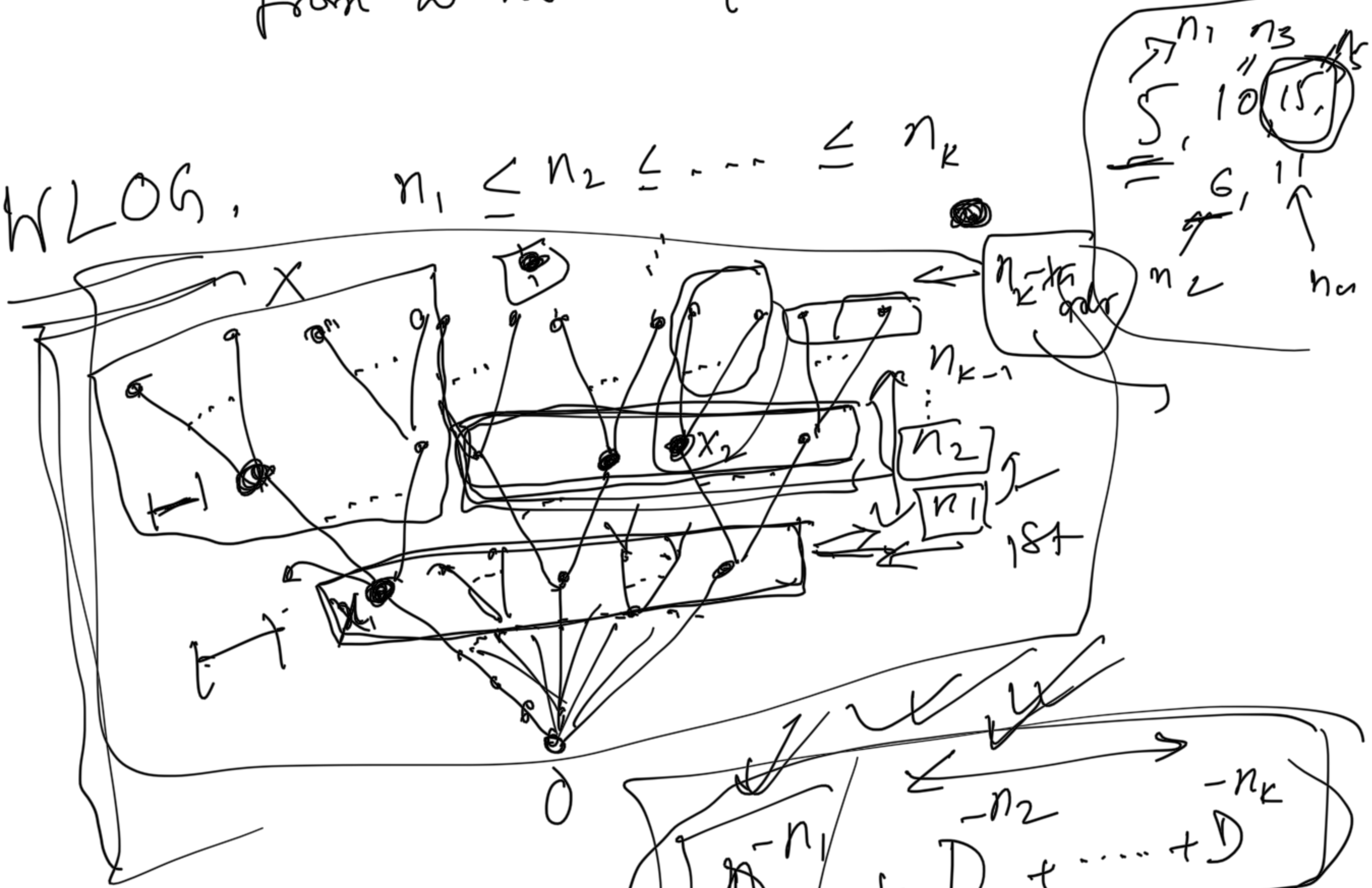
① there are D # of nodes stemming from each node at any order, say "$i-1$"

② there will be D^i # of nodes at ith order of the tree

③ the fraction of nodes originating from a node of order one is $\frac{1}{D} = D^{-1}$.
the fraction of nodes originating [illegible] order 'i' is $\frac{1}{D^i} = D^{-i}$

from a node of

WLOG, $n_1 \le n_2 \le \cdots \le n_k$



$$D^{-n_1} + D^{-n_2} + \cdots + D^{-n_k} \le 1$$

Conversely, it follows similar arguments.

$n_1 = 1$

$n_2 = 3$

$n_3 = 5$