HIDDEN MARKOV MODELS

SEQUENTIAL DATA / TIME SERIES

we observe a process evolving in time

$x_1 \; x_2 \; x_3 \; x_4 \; \ldots \; x_t$ observations, x_t = STEP (TIME) ⟶ DISCRETE TIME

t or n = time coordinate / index

- FINANCIAL DATA
- WEATHER FORECAST
- SPEECH DATA.
- EPIDEMIOLOGICAL DATA

STATIONARY / TIME-HOMOGENEOUS

$$p(x_n \mid x_{n-1}) = p(x_2 \mid x_1)$$

MARKOV CHAINS

OBSERVED!

$$p(x_1 \ldots x_n) = p(x_1)\,p(x_2 \mid x_1) \cdots p(x_n \mid x_{n-1})$$

$$x_{n+1} \perp\!\!\!\perp x_{n-1} \mid x_n \qquad \text{SHORT MEMORY}$$

x_i DISCRETE ⟹ MARKOV CHAINS

x_i CONTINUOUS, $p(x_n \mid \ldots)$ GAUSSIAN

⟹ AUTOREGRESSIVE MODELS (of order k)

$$x_{n+2} \perp\!\!\!\perp x_{n-1} \mid x_n, x_{n+1}$$

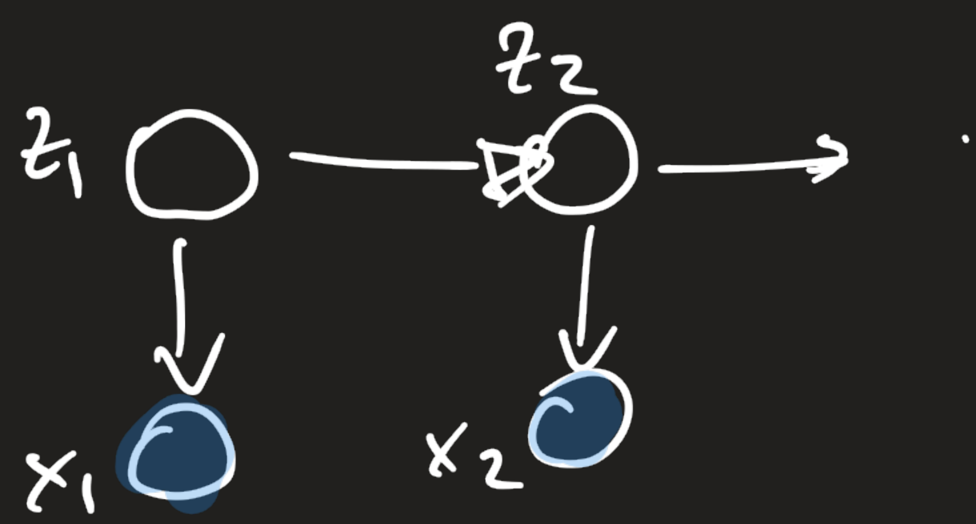
ORDER 2:

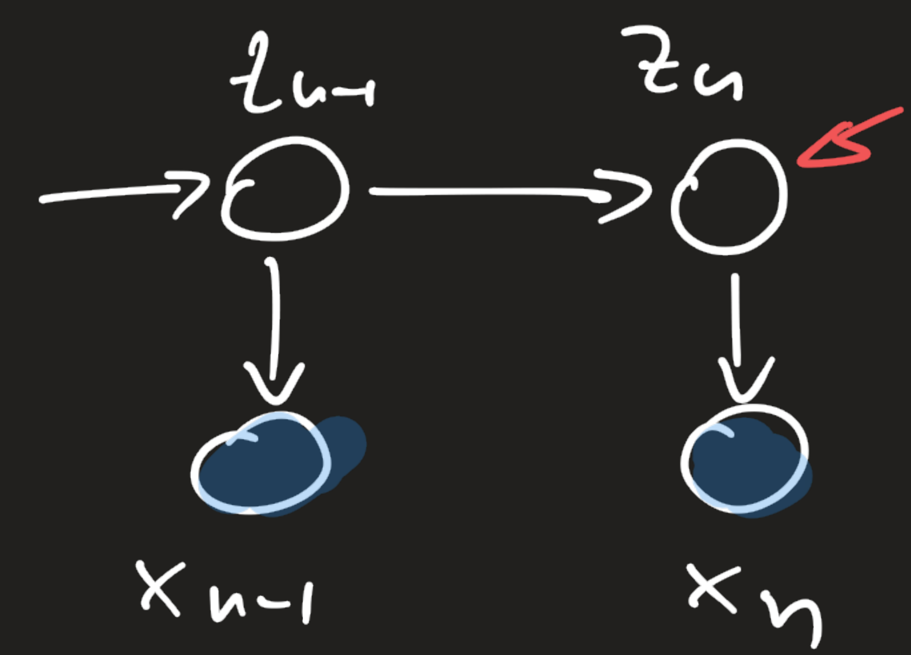
$$p(x_1, \ldots, x_n) = p(x_1)\,p(x_2 \mid x_1)\,p(x_3 \mid x_1, x_2) \cdots p(x_{n+1} \mid x_n, x_{n-1}) \cdots$$

STATE SPACE MODELS

$\underline{X} : x_1, \ldots x_n$ observed

- LATENT VARIABLES





$$x_i \not\perp\!\!\!\perp x_j \mid \underline{X}$$

transition matrix A

- $p(z_n \mid z_{n-1})$ TRANSITION PROBABILITY $\quad A_{ij} = p(z_n = j \mid z_{n-1} = i)$
- $p(z_1)$ INITIAL DISTRIBUTION $\quad \pi_i = P(z_1 = i)$
- $p(x_n \mid z_n)$ EMISSION PROBABILITY $\quad \psi$ parameters

DISCRETE X
CONTINUOUS X
- GAUSSIAN ..
- MIXTURE OF GAUSSIANS ..

FOR DISCRETE z_i $\Rightarrow$ <u>HIDDEN MARKOV MODELS</u>

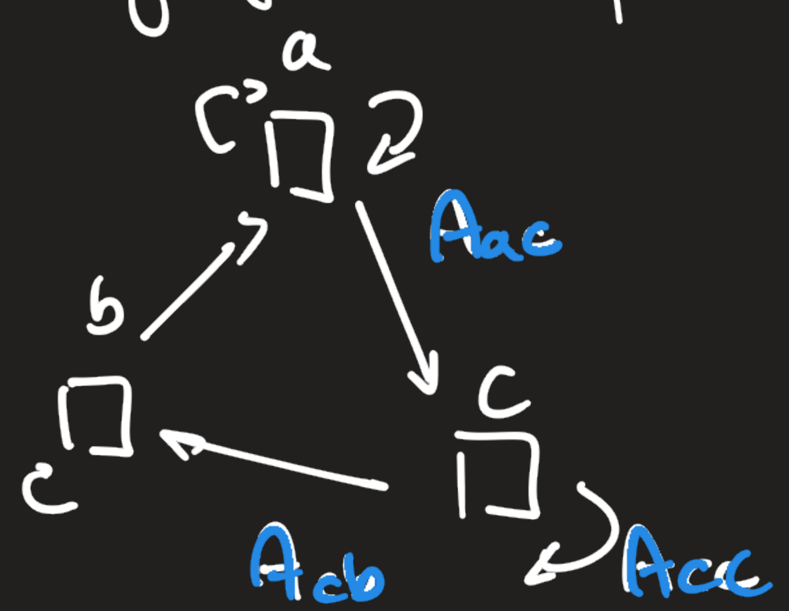
z_i CONTINUOUS, $p(z_i \mid z_{i-1})$ GAUSSIAN $\Rightarrow$ LINEAR DYNAMICAL SYSTEMS

$X, Z = (x_1, \ldots, x_N, z_1, \ldots, z_N)$

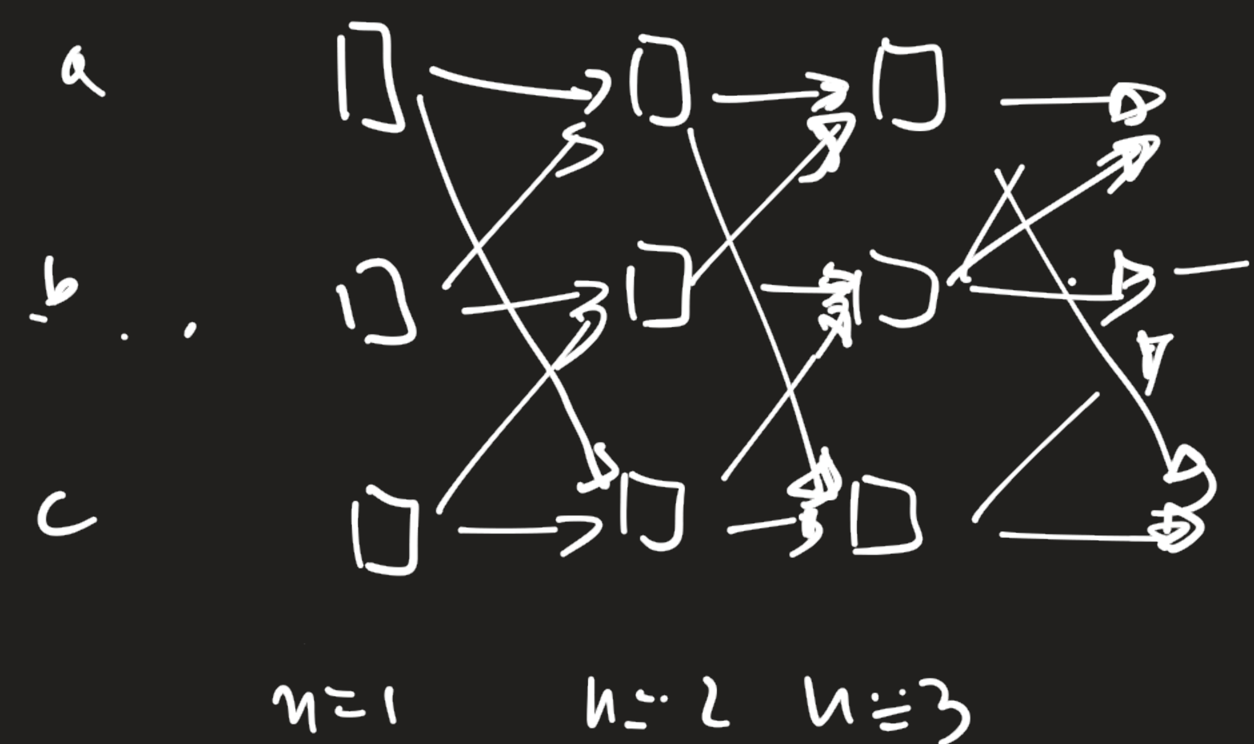
$\theta = (A, \pi, \psi)$ PARAMETERS

$$p(X,Z|\theta) = p(z_1|\pi)\left[\prod_{n=2}^{N} p(z_n|z_{n-1},A)\right]\left[\prod_{n=1}^{N} p(x_n|z_n,\psi)\right]$$

graphical representation for STATES of Z.

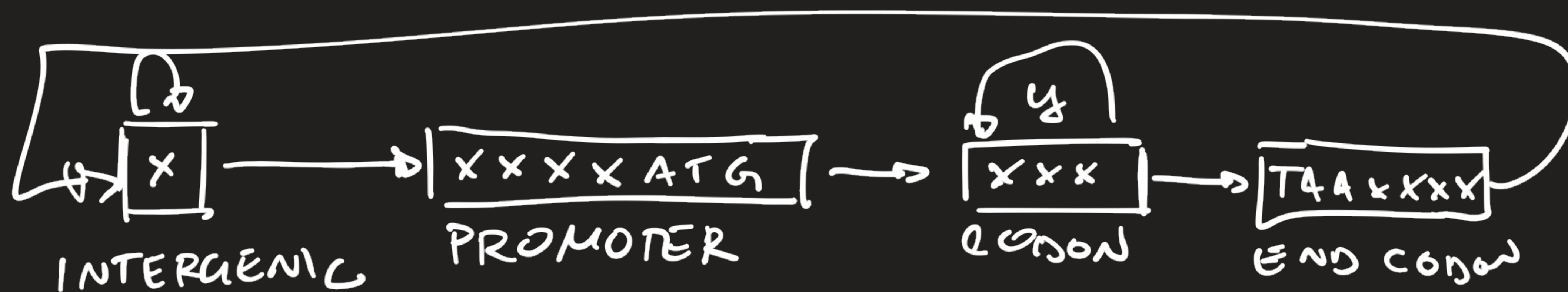


UNFOLDING ~>



HMM FOR GENE FINDING.

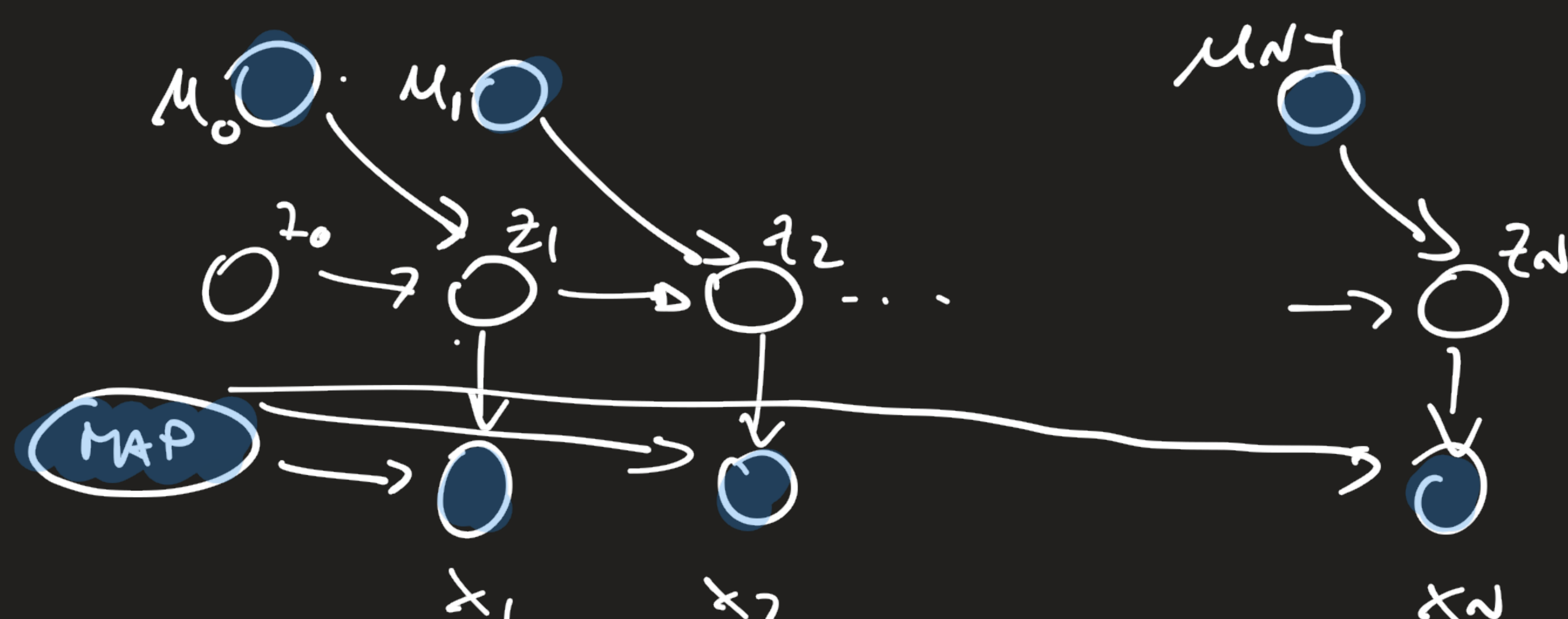
LATENT VARIABLE STATE SPACE



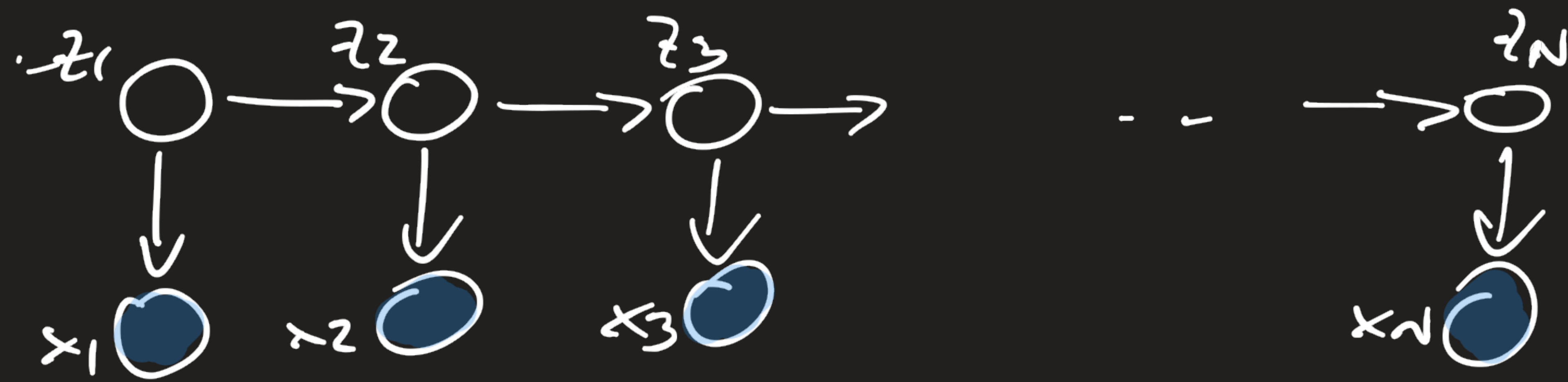
output $x \in \{A, C, G, T\}$

$y \in \{[A, C, G, T]\}^3$

ROBOT LOCALIZATION



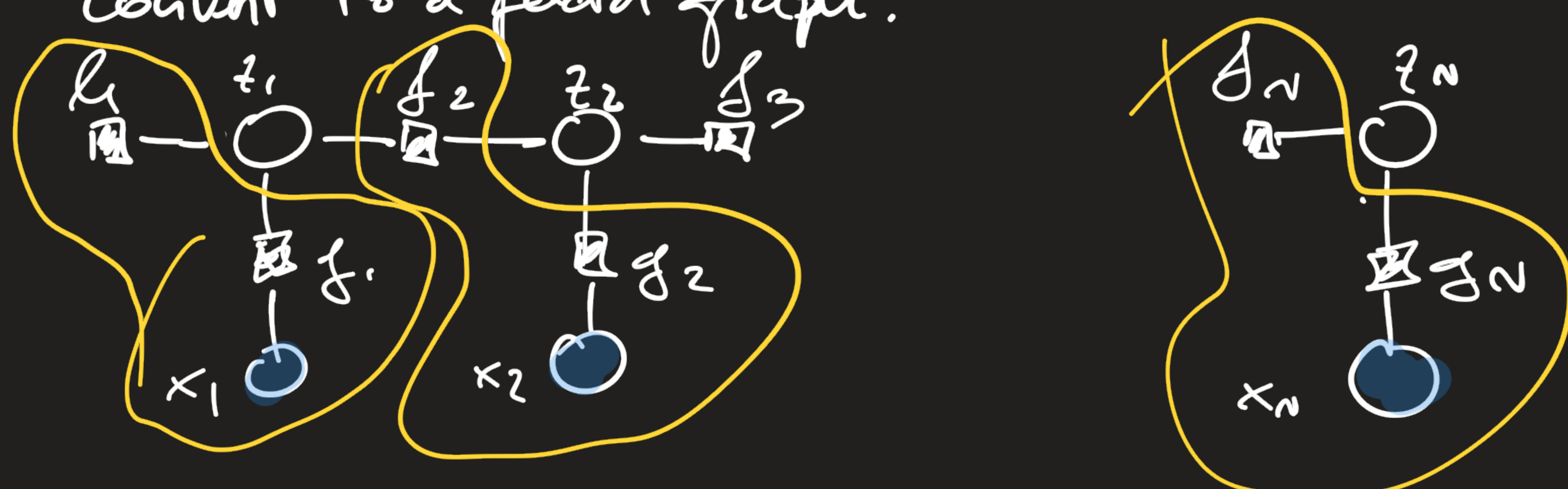
INFERENCE IN HMM



$\underline{x}: x_1, \ldots, x_N$

- **FILTERING**: compute $p(z_N \mid \underline{x})$ (last observed instant N)
- **SMOOTHING**: compute $p(z_k \mid \underline{x})$, for $K<N$
- **MOST LIKELY EXPLANATION**: compute $z^* = \arg\max_{z} p(\underline{z} \mid \underline{x})$, $\underline{z}=z_1, \ldots, z_N$

Convert to a factor graph:



OBSERVED, hence CLAMPED

$$h(z_1) = p(z_1)\underbrace{p(x_1|z_1)}$$

CLAMPED

$$f_2(z_1, z_2) = p(z_2|z_1)\underbrace{p(x_2|z_2)}_{g_2:}$$

$$\vdots$$

$$f_N(z_{N-1}, z_N) = p(z_N|z_{N-1})\,p(x_N|z_N)$$

- FILTERING + SMOOTHING on sum-product algorithm

$$\mu_{z_{n-1}\to f_n}(z_{n-1}) = \mu_{f_{n-1}\to z_{n-1}}(z_{n-1})$$

$$\underbrace{\mu_{f_n\to z_n}(z_n)}_{\alpha(z_n)} = \sum_{z_{n-1}} f_n(z_{n-1}, z_n)\cdot\mu_{z_{n-1}\to f_n}(z_{n-1})$$

$$\alpha(z_n) = \sum_{z_{n-1}} f_n(z_{n-1}, z_n)\,\alpha(z_{n-1})$$

$$\underbrace{\mu_{f_{n+1}\to z_n}(z_n)}_{\beta(z_n)} = \sum_{z_{n+1}} f_{n+1}(z_n, z_{n+1})\overbrace{\mu_{f_{n+2}\to z_{n+1}}(z_{n+1})}^{\beta(z_{n+1})}$$

SMOOTHING. $p(z_n, \underline{x}) = \alpha(z_n)\beta(z_n)$

FILTERING $p(z_N, \underline{x}) = \alpha(z_N)$

$p(z_n, z_{n+1}, \underline{x}) = \alpha(z_n) f(z_n, z_{n+1}) \beta(z_{n+1})$

$$p(z_n|\underline{x}) = \frac{p(z_n, \underline{x})}{\sum_{z_n} p(z_n, \underline{x}) = p(\underline{x})}$$

$$\boxed{p(\underline{x}) = \sum_{z_N} \alpha(z_N)}$$

MOST LIKELY SEQUENCE → max plus algorithm

$$\hat{\mu}_{f_n \to z_n} = \max_{z_{n-1}} \{ \underbrace{\log f_n(z_n, z_{n-1}) + \hat{\mu}_{f_{n-1} \to z_{n-1}}(z_{n-1})} \}$$

$\to \phi(z_n) = \arg\max_{z_{n-1}}$ ← VITERBI ALGORITHM.