~~overfitting~~

~~gradient optimisation~~

~~initialisation~~

paragraph code snippet

3rd problem & (3)

~~tikhonov vs lasso~~

what is it good for

~~problem~~ γ in (2)

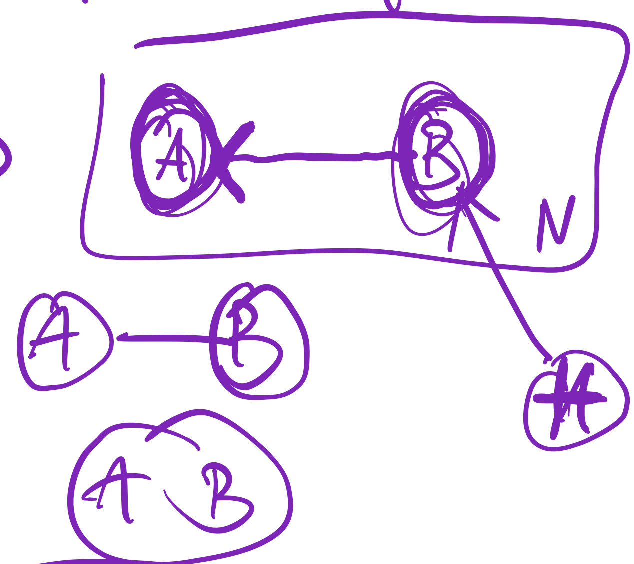
~~that data~~

'student constraints'
posts.
slides.

Graphical Models / plate diagrams

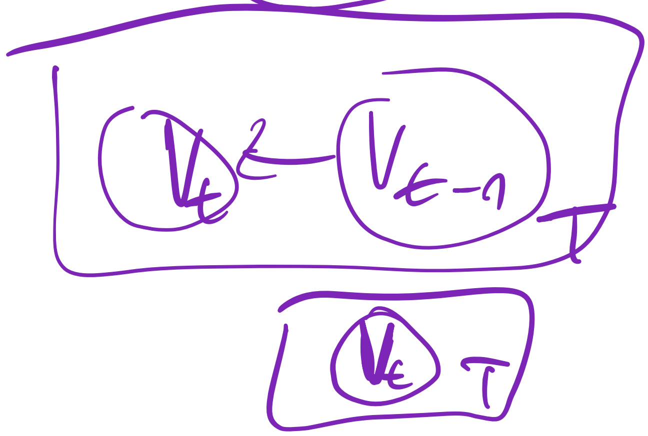
$$\prod^{N} P(A|B) \Longleftrightarrow$$

$$P(A,B) \Rightarrow$$



$$P(V_t | V_{t-1})$$

↳



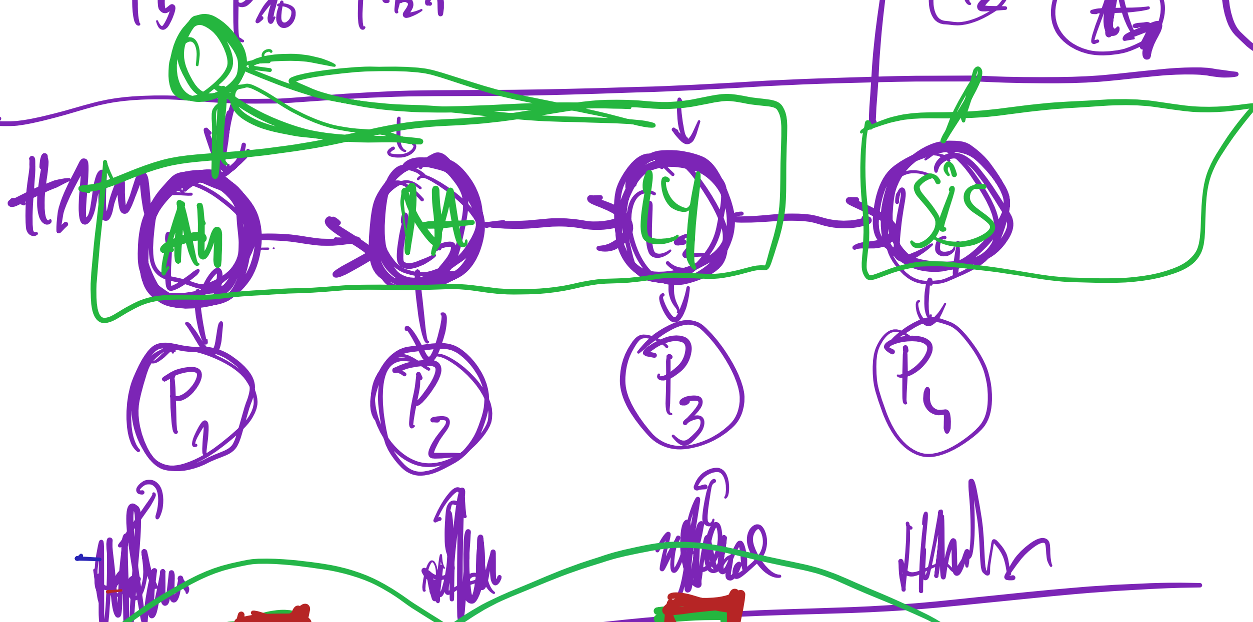
$L = \prod^{N} \ldots\ldots$

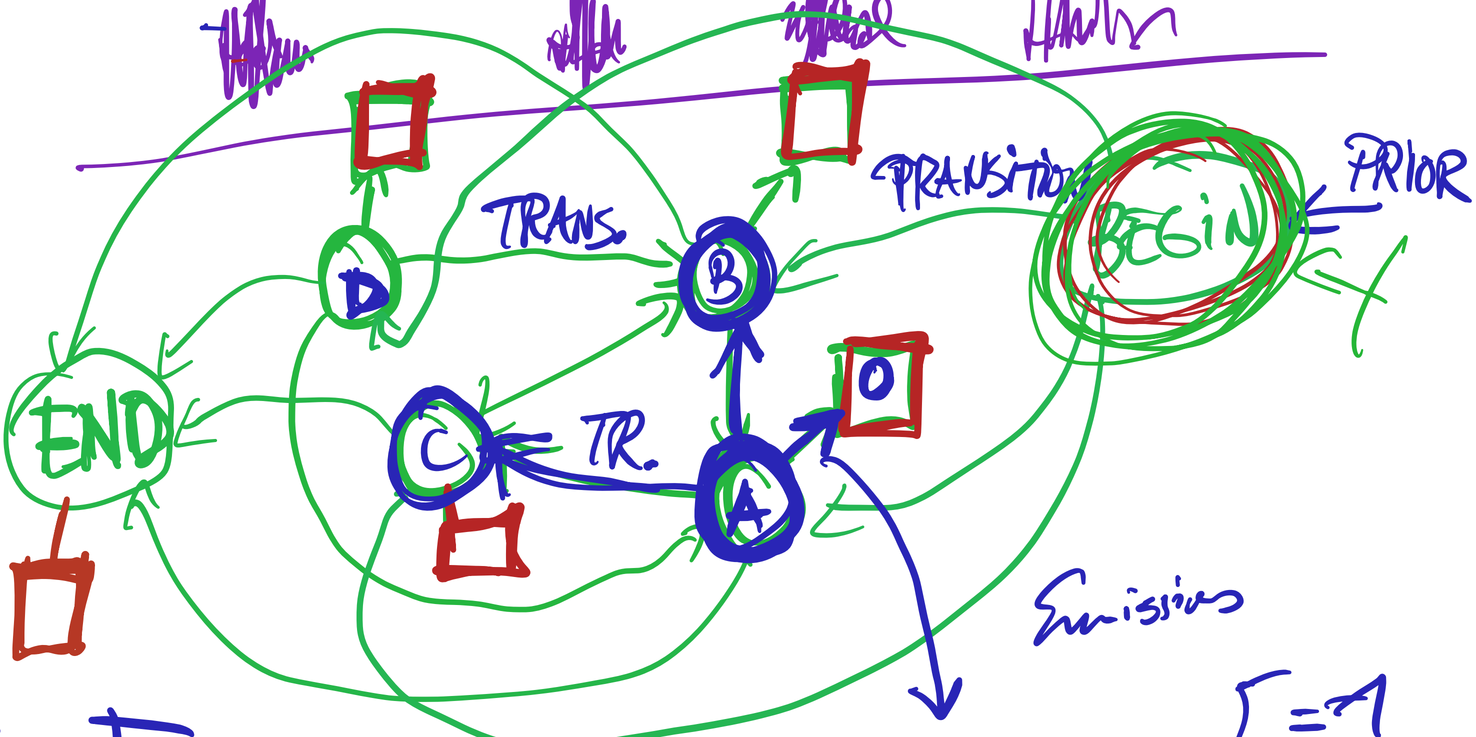
MC

$P_1 = p_1$ $P_2 = p_2$ $P_3 = p_3$ ....

P_1 P_4 P_6 ...

P_5 P_{10} P_{121} ...



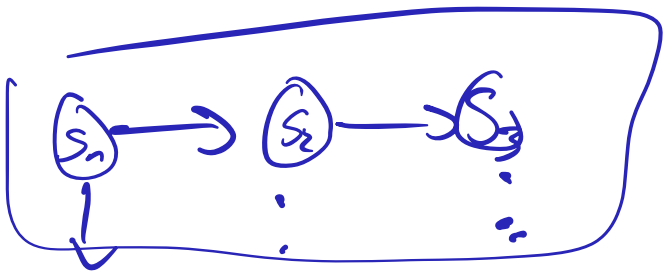


Transitions

$P(C|A) = \dots$ $\Sigma = 1$
$P(B|A) = \dots$ Multinomial
$P(D|A) = 0$

$P(O_1|A) = \dots$
$P(O_2|A) = \dots$ Multin.
$\vdots$
$P(O_N|A) = \dots$

$P(S_1, S_2, S_3)$

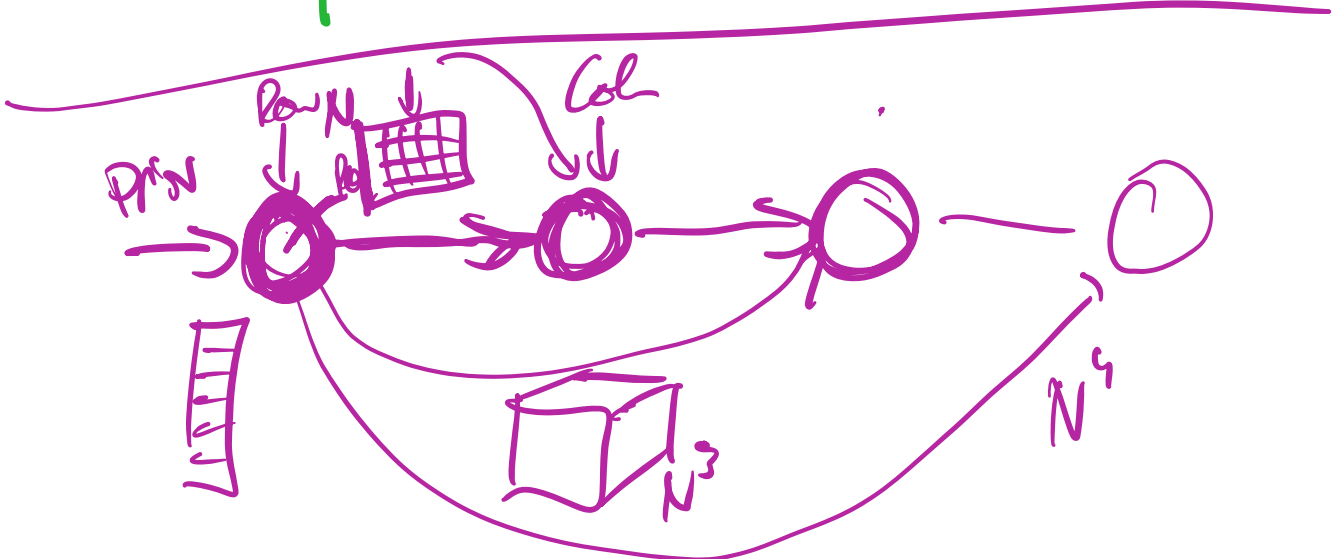


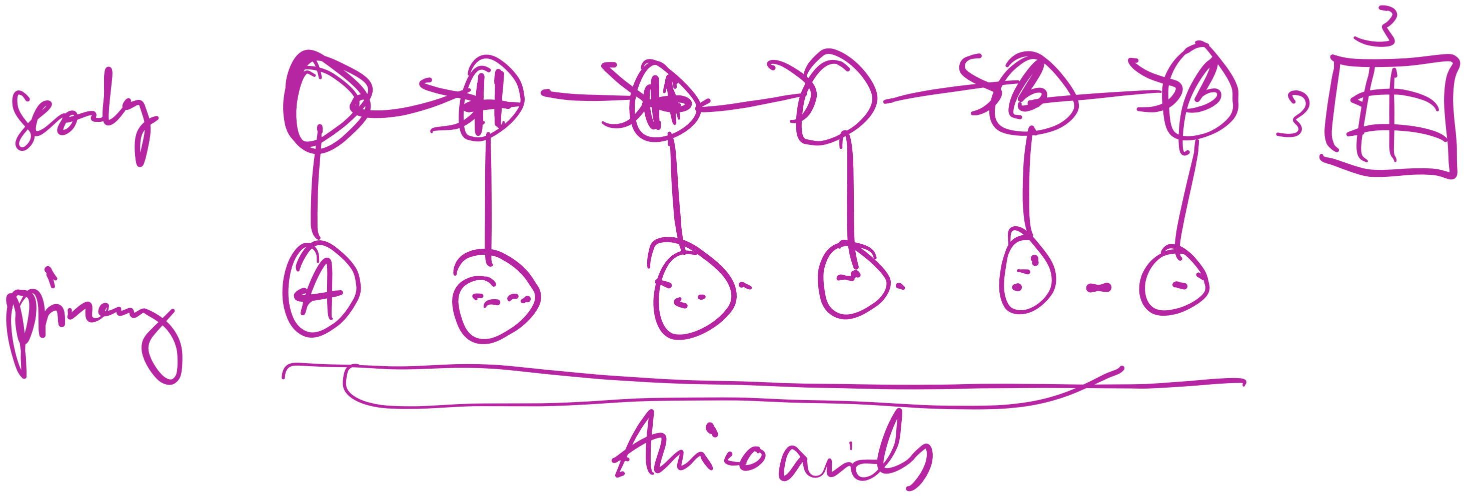
$P(S_1) \cdot P(S_2|S_1) \cdot P(S_3|S_1, S_2)$ ~~$P(S_4|S_2, S_3)$~~

$P(S_1)$ $P(S_2|S_1)$ $P(S_3|S_2)$ $P(S_4|S_3)$ ...

$L = P(S_1|\text{BEGIN}) \prod P(S_t|S_{t-1})$

$P(\text{BEGIN}) \cdot P(S_1|\text{BEGIN})$





---

① ② ③

$P(O|I)$ → ① Forward Backward...

→ ② Viterbi

$P(S_t | X_1 \ldots X_t)$

$P(S_1 \ldots S_t | \vec{X})$

---

③ Maximum likelihood

$D = \{(\vec{I}, \vec{O})\}$ → $P(\text{Rasen...} | S) = 0$

$P(\vec{O} | \vec{I}) = \prod \ldots = 0$

Laplace Smoothing

dictionary D

$$P(O_t | S_t) = \frac{\#(O_t, S_t) + 1}{N + |D|}$$

freq

Baum-Welch / EM