1

Stat methods of AI/ML.

CS 6301.005

Revanth Segu

2) Sampling Algorithm

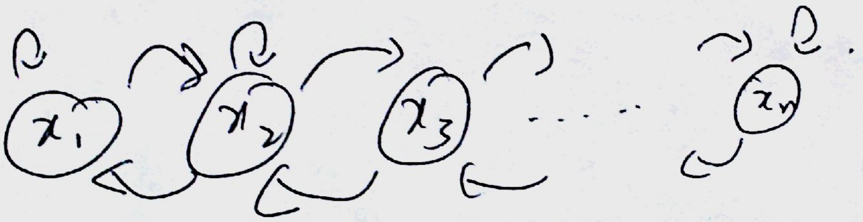
(i) Consider Metropolis Hasting Algorithm.

$$P(X_i = x_i / E=e) = \frac{1}{T}\sum_{t=1}^{T} \delta_{x_i}(z^{(t)})$$

let us assume ~~st~~ $X_i = x_i$ takes n values

$$X = \{x_1, x_2, x_3 \ldots x_n\} \rightarrow (1)$$

the markov chain looks like



Consider Gibbs Sampling.

here we randomly assign variables to all non evidence variables

then we compute

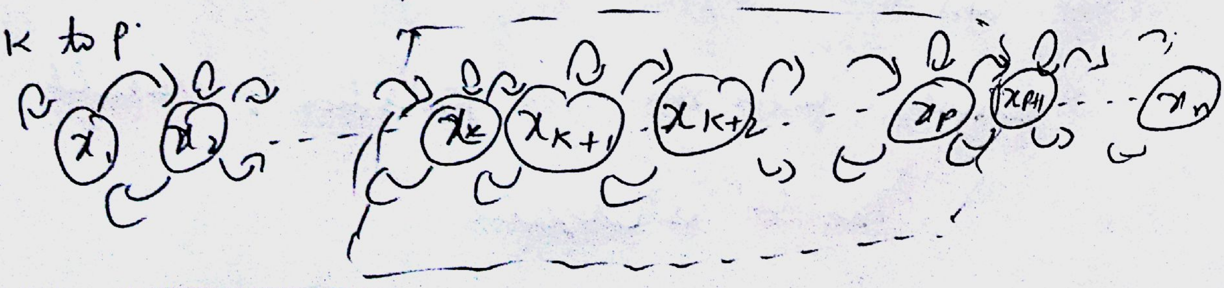
$$P(X_i / E=e, X_{-i}).$$ X_{-i} is currently sampled from $X / E \cup X_i$

Let => $E, X_{-i} \subset X$. $\{ (X) \text{ from } (1) \}$

=> ∴ we can assume $E=e, X_{-i}$ as subset of X ranging from k to p.

T

so p.



so gibbs Sampling will sample from subset of the set from which metropolis hasting algorithm will sample.

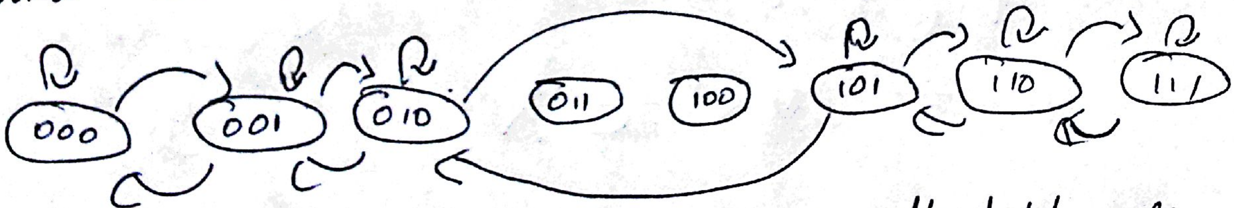
$\therefore$ Gibbs Sampling is special Case of Metropolis sampling algorithm

---

(2)- Consider a network of three variables (binary) x_1, x_2, x_3
Let the some of them has probability 0 (011, 100) in our Case)

x_1	x_2	x_3	θ
0	0	0	$P_0 > 0$
0	0	1	$P_1 > 0$
0	1	0	$P_2 > 0$
0	1	1	0
1	0	0	0
1	0	1	$P_5 > 0$
1	1	0	$P_6 > 0$
1	1	1	$P_7 > 0$

The markov network looks like shown below.



We Can still construct the markov network with left over examples with non zero transition probability state

$\therefore$ The It is Ergodic

Q3 MPE

10.7

without evidence of C

$MPE = O(k^2)$ k is a constant.

$MAP = O(k^{n+1})$

with evidence on C

its just one calculation on each variable

$\therefore$ $MPE = O(n)$

$MAP = O(n)$

10.8

① when instantiate $m_x = MAP(M, e_x)$ we can instantiate and sumout ~~the~~ on the given evidence variables. And hence we can compute

$\therefore$ It is True

② we can not compute MAP instantiation as. the given $X \in M$

$\therefore$ It is False

(10.9) Given all MAP variables are present in same cluster. We can order MAP variables to the last and hence we can calculate MAP in less than the expected time & space i.e $O(n \exp(w))$