Plate models

Probabilistic Graphical Models

Jerónimo Hernández-González

Plate models

Example: Latent Dirichlet Allocation (LDA) for topic modeling

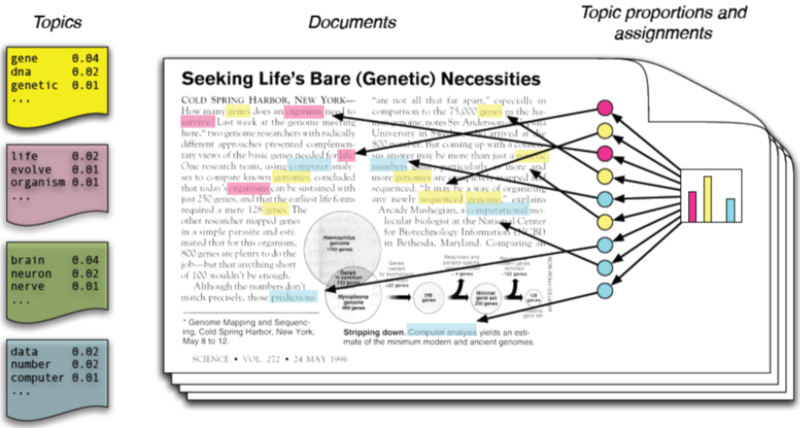
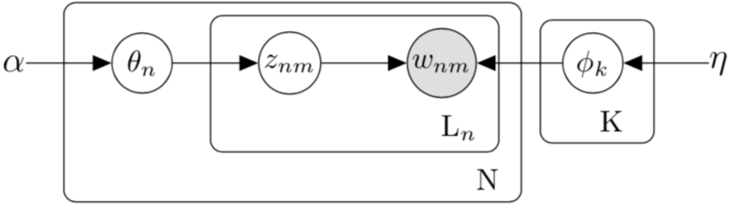


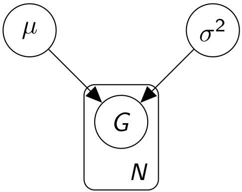
Plate models

Example: Latent Dirichlet Allocation (LDA) for topic modeling



BN with repeated structure

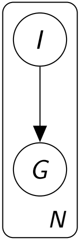
Student-grade



N students;
for each student n, her grade (G_n) is a sample from a Normal distribution $\mathcal{N}(\mu, \sigma^2)$

BN with repeated structure

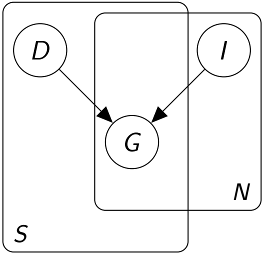
Student-grade II



N students;
for each student n, her intelligence (I_n) determines her grade (G_n)

BN with repeated structure

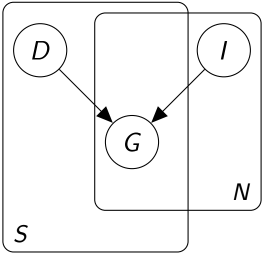
Intersection: Difficulty-Intelligence-Grade



S subjects and N students
for each subject s with a specific difficulty D_s,
and for each student n with a specific intelligence I_n,
both the difficulty of the subject (D_s) and her intelligence (I_n)
determine her grade (G_{sn})

BN with repeated structure

Intersection: Difficulty-Intelligence-Grade

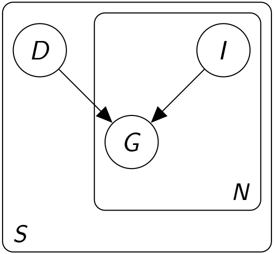


S subjects and N students
for each subject s with a specific difficulty D_s,
and for each student n with a specific intelligence I_n,
both the difficulty of the subject (D_s) and her intelligence (I_n)
determine her grade (G_{sn})

General intelligence

BN with Repeated Structure

Nesting: Difficulty-Intelligence-Grade

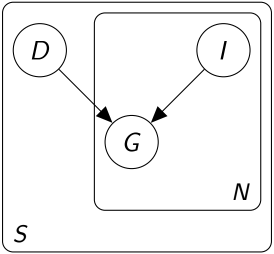


S subjects with N students

for each subject s with a specific difficulty D_s, there are N students;
for each student n her intelligence (I_{sn}) and the difficulty of the subject (D_s) determine her grade (G_{sn})

BN with Repeated Structure

Nesting: Difficulty-Intelligence-Grade

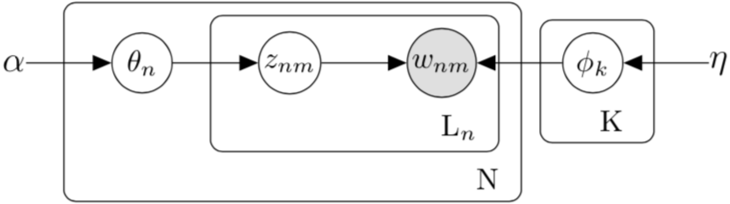


S subjects with N students
for each subject s with a specific difficulty D_s, there are N students;
for each student n her intelligence (I_{sn}) and the difficulty of the subject (D_s) determine her grade (G_{sn})

Subject-specific intelligence

Plate models

Example: Latent Dirichlet Allocation (LDA) for topic modeling



For each topic $k = 1 \ldots K$

$\phi_k \sim \mathrm{Dir}(\eta)$

For each document $n = 1 \ldots N$

$\theta_n \sim \mathrm{Dir}(\alpha)$

For each word $m = 1 \ldots L_n$

$z_{nm} \sim \mathrm{Cat}(\theta_n)$

$w_{nm} \sim \mathrm{Cat}(\phi_{z_{nm}})$

Exercise

Plate Semantics

Let A and B be random variables inside a common plate indexed by i. Which statement is true?

[You may select 1 or more options]

a) For each i, $A(i)$ and $B(i)$ have different CPDs.

b) For each i, $A(i)$ and $B(i)$ have edges connecting them to the same variables outside of the plate.

c) For each i, $A(i)$ and $B(i)$ have the same CPDs.

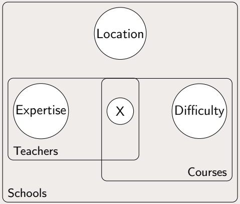
d) There is an instance of A and an instance of B for every i.

Exercise

Plate Interpretation

Consider this plate model with edges removed. What might possibly represent an instance of X in the grounded model?

[You may select 1 or more options]



a) Whether a specific teacher T taught a specific course C at school S

b) Whether someone with expertise E taught something of difficulty D at a place in location L

c) Whether a specific teacher T is a tough grader

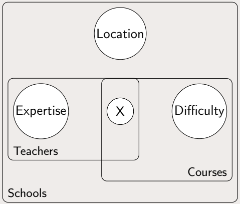
d) None

e) Whether a teacher with expertise E taught a course of difficulty D

Exercise

Grounded Plates

Consider this plate model and assume that there are s schools, t teachers and c courses in each school. How many instances of the Difficulty variable are there?



a) c

b) $s \cdot c$

c) Not enough information to answer

d) $s \cdot t$

Plate models

Probabilistic Graphical Models

Jerónimo Hernández-González

Plate models

Limitations

- cannot have edges between two instances of the same variable

  (e.g., position a time t depends on position at time $t-1$)
- cannot have edges between particular pairs selected by some other relation

  (e.g., Genotype(U1) depends on Genotype(U2), where U2 is mother of U1

Alternatives

- Dynamic Bayesian Networks (DBNs)

  Specific to repetitions over time
- Probabilistic relational models

  More flexibility

Template models

What is a template model?

- $\boldsymbol{X}$ takes different values at each (discrete) time step
  $\boldsymbol{X}(t)$ is the random variable at time t
- Markov assumption

$$\boldsymbol{X}(t+1) \perp\!\!\!\perp \boldsymbol{X}(0), \ldots, \boldsymbol{X}(t-1) \mid \boldsymbol{X}(t)$$

- Stationary assumption (Time invariance or homogenous)

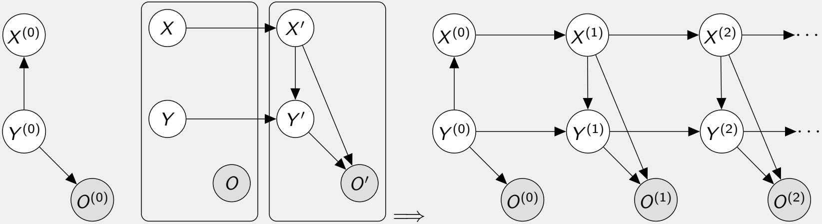
$$P(\boldsymbol{X}(t+1) \mid \boldsymbol{X}(t)), \text{ the same for all } t$$

- Use **conditional** Bayesian network to define $P(\boldsymbol{X}(t+1)|\boldsymbol{X}(t))$
  2-time slice Bayesian network, Dynamic Bayesian network, Hidden Markov models

Temporal models

Dynamic Bayesian Network

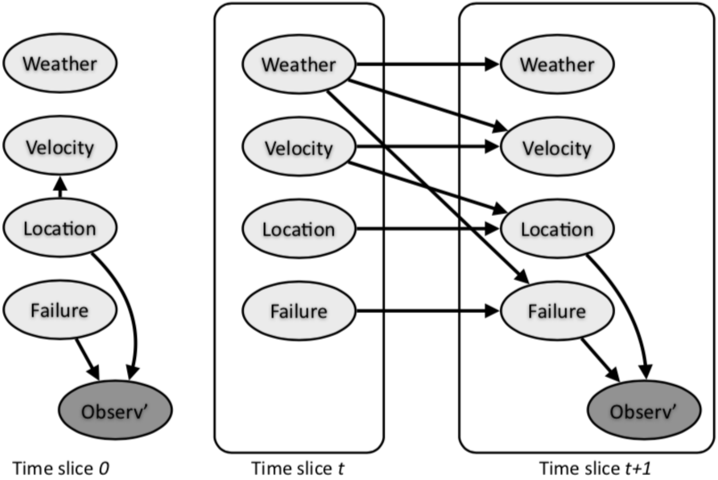
1. Bayesian network over $\boldsymbol{X}(0)$
2. Conditional BN for $\boldsymbol{X}(t+1)$ given $\boldsymbol{X}(t)$ (2-time-slice)



(1) (2) Grounded (unrolled) BN

Temporal models

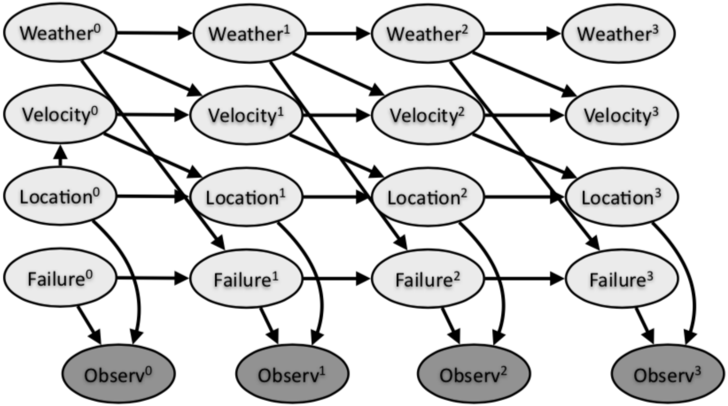
Example: Dynamic Bayesian Network (DBN) for vehicle position



$$P(W', V', L', F', O' | W, V, L, F) = p(O' | F', L') p(F' | F, W) p(L' | L, V) p(V' | V, W) p(W' | W)$$

Temporal models

Example: Dynamic Bayesian Network (DBN) for vehicle position



Grounded (unrolled) BN

Exercise

Markov Assumption

If a dynamic system X satisfies the Markov assumption for all time $t \geq 0$, which statement is true?

a) $X^{(t+1)} \perp\!\!\!\perp X^{(t)}$

b) $X^{(t+1)} \perp\!\!\!\perp X^{(t)} | X^{(t-1)}$

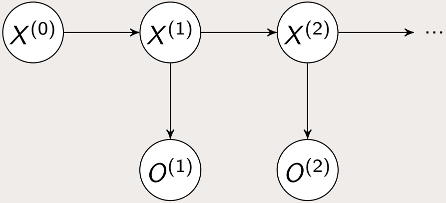
c) $X^{(t+1)} \perp\!\!\!\perp X^{(0:(t-1))} | X^{(t)}$

Exercise

Independencies in DBNs

In this DBN, which of these independence statements are true?

[You may select 1 or more options]



a) $O^{(t)} \perp\!\!\!\perp X^{(t+1)} | X^{(t)}$

b) $O^{(t)} \perp\!\!\!\perp X^{(t-1)} | X^{(t)}$

c) $O^{(t)} \perp\!\!\!\perp O^{(t-1)}$

d) $O^{(t)} \perp\!\!\!\perp O^{(t-1)} | X^{(t)}$

Exercise

Applications of DBNs

For which of the following applications might one use a DBN?

a) Modeling data taken at different locations along a road, where the data at each location is influenced by the data at many other locations.

b) Predicting the probability that today will be a snow day (school will be closed because of the snow), when this probability depends only on whether yesterday was a snow day.

c) Predicting the probability that today will be a snow day (school will be closed because of the snow), when this probability depends only on whether yesterday, the day before yesterday, and 2 Mondays ago were snow days.

d) Modeling time-series data, where the events at each time-point are influenced by only the events at the one time-point directly before it

Plate models

Probabilistic Graphical Models

Jerónimo Hernández-González