Cmput 651 – Probabilistic Graphical Models

Probabilistic Graphical Models (Cmput 651): Undirected Graphical Models 2

Matthew Brown

20/10/2008

Reading: Koller-Friedman Section 4.5

1

Outline

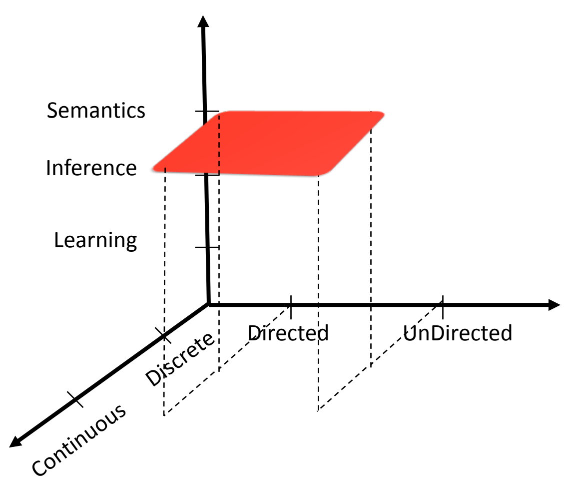
- **Review of previous lecture (other .ppt file)**
- Markov nets vs. Bayes nets
- Markov nets and variable elimination

2

Outline

- Review of previous lecture (other .ppt file)
- **Markov nets vs. Bayes nets**
- Markov nets and variable elimination

Space of Topics

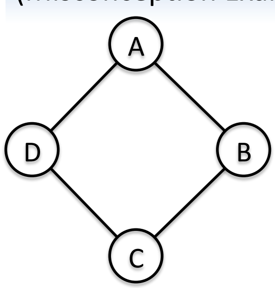


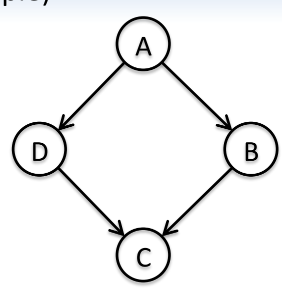
Bayes and Markov Nets Comparison

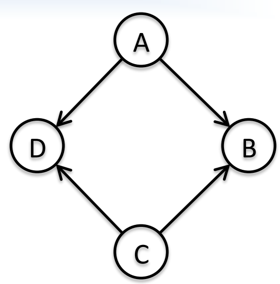
VS

Bayes nets fail on diamond structure

(Misconception Example)







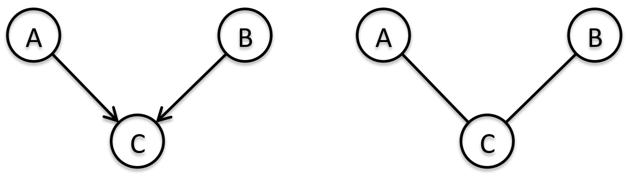
This works:	NO! This implies:	NO! This implies:
$A \perp C \mid B,D$ ✓	$A \perp C \mid B,D$ ✓	$\neg(A \perp C \mid B,D)$ ✗
$B \perp D \mid A,C$ ✓	$\neg(B \perp D \mid A,C)$ ✗	$B \perp D \mid A,C$ ✓

Bayesian networks cannot represent some distributions

- Misconception example from Koller-Friedman (Fig 3.16)

Markov nets fail on V-structure

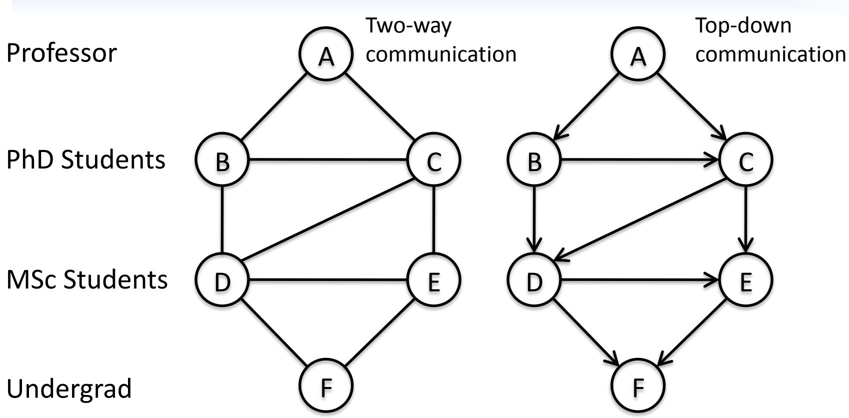
Suppose A and B are independent, uniformly distributed binary variables and that C = XOR(A,B):



This works: $\neg(A \perp B \mid C)$ ✓

NO! This wrongly implies: $A \perp B \mid C$ ✗

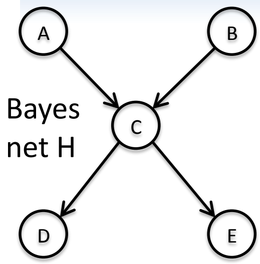
Sometimes Bayes and Markov nets both work



Some distributions can be represented by both Bayes and Markov nets. In this case, the Bayes and Markov nets are **chordal graphs** (more on this to come).

Bayes nets to Markov nets step 1/2

(Also see KF Section 4.5.1.)



Task: Given a PDF P that factorizes based on a Bayes net H, find a Markov net M that factorizes P.

Start with factorization based on Bayes net H:

$$P(A,B,C,D,E) = P(A)P(B)P(C \mid A,B)P(D \mid C)P(E \mid C)$$

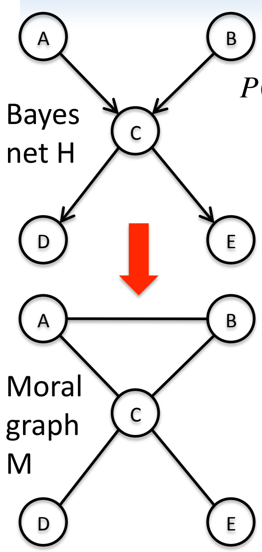
1st, we use Bayes nets' CPDs as factors in a Gibbs distribution. (Recall that PDFs and CPDs are valid examples of factors that just happen to be normalized.)

Gibbs distribution:

$$P(A,B,C,D,E) = \phi_1(A)\phi_2(B)\phi_3(A,B,C)\phi_4(C,D)\phi_5(C,E)$$
$$\phi_1(A) = P(A), \phi_2(B) = P(B), \phi_3(A,B,C) = P(C \mid A,B)$$
$$\phi_4(C,D) = P(D \mid C), \phi_5(C,E) = P(E \mid C)$$

Bayes nets to Markov nets step 2/2

(Also see KF Section 4.5.1.)



Once we have the Gibbs distribution

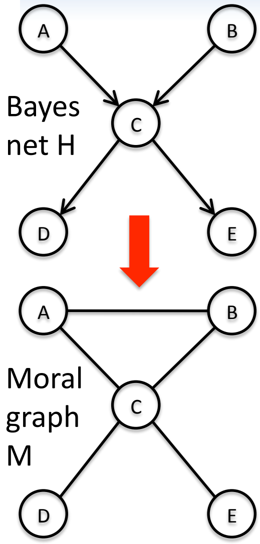
$$P(A,B,C,D,E) = \phi_1(A)\phi_2(B)\phi_3(A,B,C)\phi_4(C,D)\phi_5(C,E)$$

We build a Markov net based on the Bayes net and making sure that the cliques match the scopes (input variables) of the factors in the Gibbs distribution.

Thus, we **moralize** the Bayes net to create a **moral graph** (eg: figure on left).
A moral graph M is an undirected graph with the same skeleton as a Bayes net H except that any unconnected parents in the Bayes net are joined with an edge in the moral graph.

Bayes nets to Markov nets

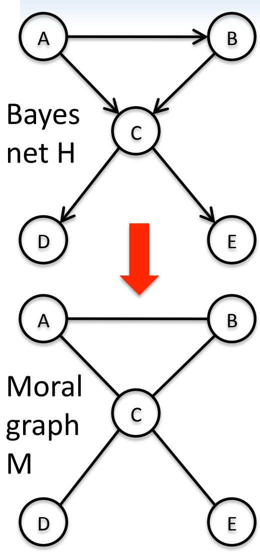
(Also see KF Section 4.5.1.)



The moral graph M of a Bayes net H is always a minimal I-map for H but not necessarily a perfect map for H (recall: a perfect map encodes precisely the same independencies). In example on left, H encodes the independency $(A \perp B \mid \{\})$ but M does not (because of the connection between A and B).

Bayes nets to Markov nets

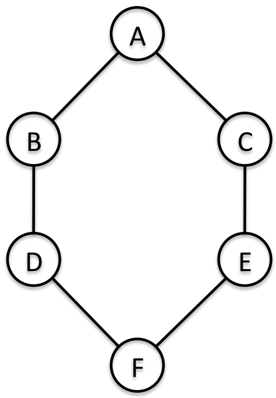
(Also see KF Section 4.5.1.)



If H is moral to begin with (i.e. if H has no unconnected parents), then the moral graph M built from H is a perfect map of H (i.e. M encodes precisely the same independencies as H, not just a subset of the independencies).

This is illustrated on the left.

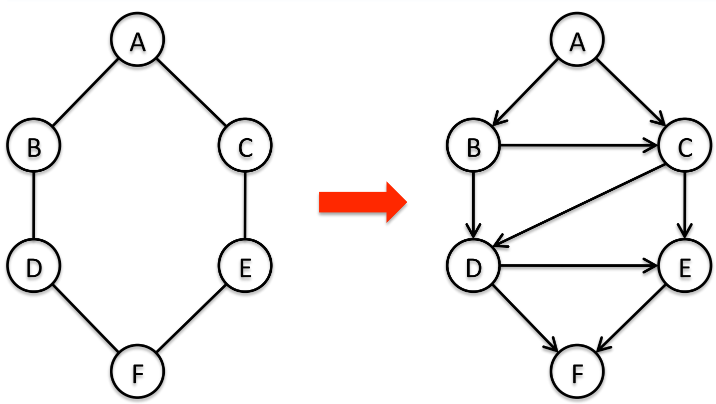
Markov nets to Bayes nets: triangulation
(also see KF 4.5.2)



Task: Given a PDF P that factorizes over a Markov net M, build a Bayes net H that factorizes P.

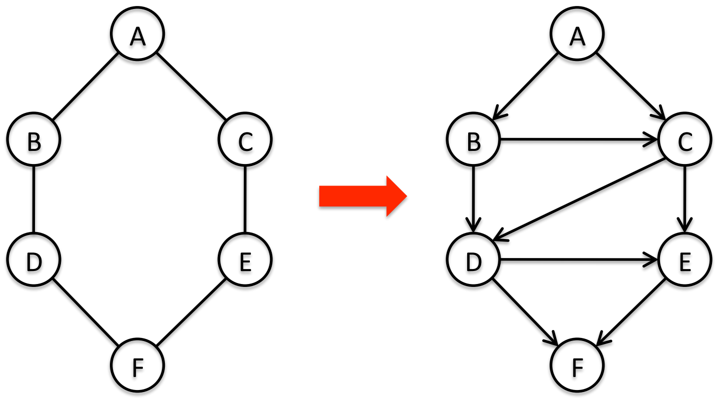
We've seen previously (lecture RG ??, KF Section 3.4.1) how to build a Bayes net given a PDF:
1. Define an ordering on the variables
2. Define parent set for each nodes according to the independencies in the PDF.

Markov nets to Bayes nets: triangulation
(also see KF 4.5.2)



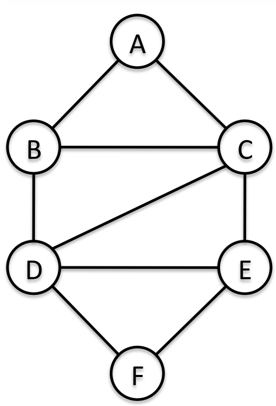
If we start with ordering A,B,C,D,E,F, for example.
Must add B-C arc because C still depends on B even after conditioning on A. Use similar arguments for C-D arc. etc. etc.

Markov nets to Bayes nets: triangulation
(also see KF 4.5.2)



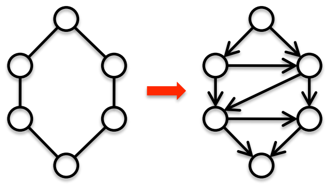
Bayes net produced by triangulation has (potentially many) extra edges compared to the Markov net, meaning the Bayes net fails to capture some independencies implied by the Markov net.
eg: $(B \perp C \mid A, F)$

Chordal Graphs (also see KF Definition 4.5.10)



Consider the loop A-B-D-F-E-C-A:
The edges B-C, C-D, and D-E are all **chords** in this loop.
A **chordal graph** is one in which any loop of length ≥ 4 has a chord. I.e. the longest minimal loop (one without a shortcut) is a triangle (has length three). Chordal graphs are also called triangular.

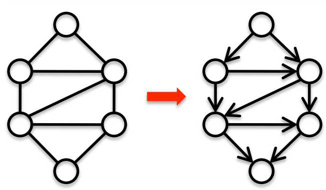
Markov nets to Bayes nets: triangulation
(also see KF 4.5.2)



Theorem: If M is a Markov net and H is a Bayes net that is a minimal I-map for M, then:
i) H can have no immoralities, and therefore
ii) H must be chordal

Interpretation: even though building a Bayes net from a Markov net using triangulation can add lots of extra edges and the Bayes net can miss some independencies, that Bayes net is the minimal I-map, so we cannot do any better with some other method.

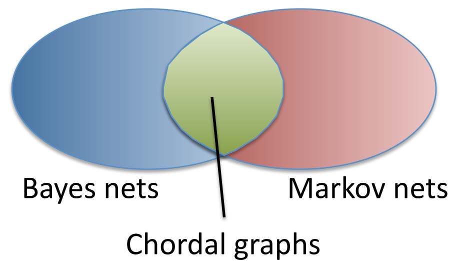
Markov nets to Bayes nets: triangulation
(also see KF 4.5.2)



When does triangulation of a Markov net M produce a Bayes net H that is a perfect map for M (i.e. I(M)=I(H))?
Answer: When M is chordal to begin with. In this case, triangulation does not add any extra edges that were not already present in M.

Chordal graphs = Bayes ∧ Markov

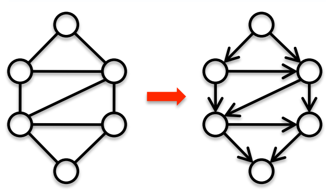
(Also see KF Section 4.5.3)



The set of chordal graphs is the intersection of the set of Bayes nets and the set of Markov nets. The independence assumptions represented by a chordal graph can be represented perfectly by either a Bayes net or Markov net.

Chordal Markov nets to chordal Bayes nets

(also see KF 4.5.3)

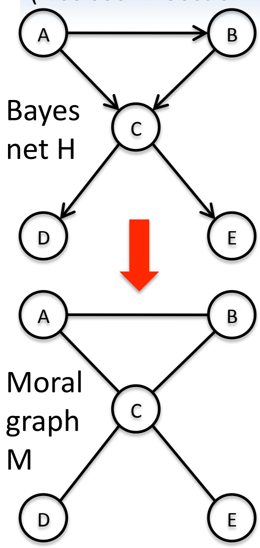


Theorem: If a Markov net M is non-chordal, then no Bayes net H is a perfect map for M.

Proof: Recall that any Bayes net H that is a minimal I-map for M must be chordal. So this Bayes net H would have extra edges not present in the (non-chordal) Markov net M. Adding edges removes independence assumptions, implying that H cannot encode all of I(M).

Chordal Bayes nets to chordal Markov nets

(Also see KF Section 4.5.3)



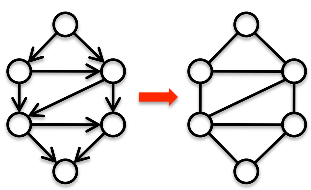
Recall this from earlier:

Only for a moral Bayes net H can we find a Markov net M that is a perfect map for H (i.e. I(M)=I(H)).

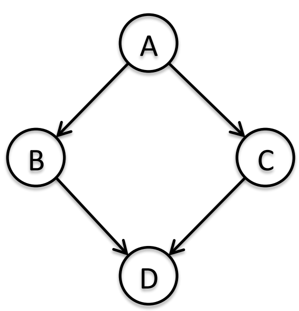
(continue to next slide)

Chordal Bayes nets to chordal Markov nets

(Also see KF Section 4.5.3)



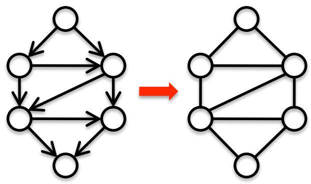
A moral Bayes net is necessarily chordal, implying that for a chordal Bayes net, we can find a Markov net that is a perfect map for the Bayes net.



Non-chordal=>non-moral:
To be non-chordal, a Bayes net must have a loop of length ≥4 that lacks a chord. This requires an immorality (recall, Bayes nets must be DAGs).

Chordal Bayes nets to chordal Markov nets

(Also see KF Section 4.5.3)



Theorem: If a Bayes net H is non-chordal, then no Markov net M is a perfect map for H. (Proof in KF Section 4.5.3.)

Outline

- Review of previous lecture (other .ppt file)
- Markov nets vs. Bayes nets
- **Markov nets and variable elimination**

Markov nets and variable elimination

Variable elimination in Markov nets is the same as in Bayes nets. In both cases, we work with factors, and the algorithm does not care where the factors come from. We have already covered variable elimination in a previous lecture, and we won't go over it again here.