Outline

- review of dependencies; info in BNs; [lecture II, III] and introduce running example

- Why trees are nice (variable elimination 
  \rightarrow \text{message passing from a child to its single parent})

- e.g., as in the "Job" node, you can get blowup in when "Job" is to
  be solved for, you get a blowup in factor sizes

- tree organization: "above" vs. "at" and below

- in HMMs, traditionally a slightly diff organization
  ("forward" includes "at" 
  \& "above")

  which leads to \text{travis-graph-organized dynamic program}

  <use square nodes, to avoid confusing \text{a graphical model}>

- can also be used to compute Viterbi paths
(can we use inbuilt display in order to show on the board?)

I: graphical model for variables relevant to a student taking a particular student.

What affects $G$, a student's grade in the course?
In this model, core grade is affected by the student's intelligence; the difficulty of the course.

Difficulty comes from how coherent the lecturer is, but if we know how difficult the course is, knowing $G$, coherence doesn't give us any extra info about the grade.

Grades affect how good your letter of recommendation is:
Your letter of rec. affects what job you get.
Your grade + your job affect your happiness.

clique q II.

4: 15% (A) is false: if I knew your $G$, I have some info about your SAT score, a more accurate guess @ your SAT score than I did before.

12: 46% (B) is true: if I knew your intelligence, my accuracy @ guessing your grade isn't going to change if I'm also told your SAT score.

2: 8% (C) is false, since I know you've got a good job.

8: 31% (so D is false)

1. (A)+(B): children are not indep, but are cond indep given parent.
2. (C) parents are not indep given children.
Correct answer to question 3: false. 
\[ P(D) = \sum_c P(D | C = c) P(C = c) \]

For question 4, we have: 
\[ P(Y, Z) = \sum_{X} P(Y, Z | X) \text{ where } X \text{ is a variable not of interest.} \]

Exercise: 
To see how structure affects our computations, let's try to compute: 
\[ P(J) = \sum_{C, D, I, \ldots, H} P(C, D, I, \ldots, H) \text{ for all except } J \]

Correct answer to question 5: a leaf below J.

14% said disappear
4% said become subscript
7% said remains an argument

If we don't know happiness, it doesn't matter. If we guess, we eliminate happiness...

Correct answer to question 6: an orphan with 1 child.

2% disappears
24% said yes
16% said remains an arg. (probably wasn't made clear to me) - cross off coherence.
in lecture, I skipped the elimination of D, which, naturally, gives you a different result when you eliminate F. But you still get a multi-argument m.

Can similarly break off orphan D

\[ \sum_{m} \sum_{\text{all \omega} = \text{a}, \text{b}, \text{c}, \text{d}} \frac{m}{(\omega, D)} \frac{1}{P(D, \omega I)} \sum_{\text{all \ I}} P(G, D, \omega I) \]

\[ = \sum_{m} \sum_{\text{all \omega} = \text{a}, \text{b}, \text{c}, \text{d}} \frac{m}{(\omega, D)} \frac{1}{P(D, \omega I)} \sum_{\text{all \ I}} P(G, D, \omega I) \]

\[ \text{mathematically, equals } P(G, I), \text{ given by table.} \]
In an HMM looks like this:

What about work? demand?

What is $P(C)$?

So working upwards is great!

But, wait. This doesn't look like what we studied yesterday. It seems like we are marginally over the observed.

Make sure $X, Y, Z$ are observed.

Eliminate root: yuck...

$$\sum_A P(w|A)P(b|A)P(a)$$

1st eliminate W:

$$\sum_{w} P(w|A)$$

Now eliminate root:

$$\sum_A P(b|A)P(a)$$

now eliminate root:

$$m_a(b)$$

now eliminate leaf $X$: $\sum_{x} P(x|b)$$

now eliminate chain root $B$:

$$\sum_{b} P(c|b) m_a(b)$$

Eliminate leaf $X$, leaf $Y$, leaf $Z$.

$$m_b(c) \quad m_b(c) = \sum_{b} P(c|b) m_a(b)$$

$$P(A, B)$$

(Continue this)
So, let's talk about computing $P(B, \hat{Z}, \hat{X}, X, Y, Y, Y, Z)$.

Bottom up:

Eliminate the "leaf" $C$

$$
\sum_{C} P(c|B) P(z|c) P(Z|c) \\
= \sum_{C} P(c|B) P(z|c) \frac{m_c(B, Z|z)}{m_c(B, Z|z)} \\
= \sum_{C} P(c|B) \frac{m_c(B, Z|z)}{m_c(B, Z|z)}
$$

Eliminate the root $C$

Now we have to do the root

$$
\sum_A P(B|A) P(X|x|A) P(A) \\
= \sum_A P(B|A) \frac{m_A(B, X|x)}{m_A(B, X|x)}
$$

and we have one more leaf

$P(Y|y|B)$

prob the 2nd emission is $Y$ given 2nd state is $B$.

\[ \text{which we can write as} \]
\[ \beta_2(b) = \hat{P}(Z|y|B, b) \]

\[ \text{almost from:} \]
\[ P(\text{2nd state is } B; \text{1st emission is } X), \]

\[ \hat{\alpha}_2(b) = \hat{P}(X=x, B=b) \]

\[ \text{which we can write as a} \]
\[ \"Sridharan \alpha\" \]

Sridharan's $\alpha_2$

Sridharan's $\beta_2$

Combines the $\alpha_2$, the $\beta_2$, $\hat{P}$ of transitioning from $B$ to $C$ given $B$, prob of $B$ path $Y$ given $B$, and you get the prob of the whole.