Announcements

- Solution to assignment 6 in web page
- Assignment 7 due next Monday
- Midterm next Thursday. Format like last midterm.

It is remarkable that a science which began with the consideration of games of chance should become the most important object of human knowledge ... The most important questions of life are, for the most part, really only problems of probability ...

The theory of probabilities is at bottom nothing but common sense reduced to calculus.

– Pierre Simon de Laplace, Théorie Analytique de Probabilités [1812]

Belief network inference

Main approaches to determine posterior distributions in graphical models:

- Variable Elimination, recursive conditioning: exploit the structure of the network to eliminate (sum out) the non-observed, non-query variables one at a time.
- Stochastic simulation: random cases are generated according to the probability distributions.
- Variational methods: find the closest tractable distribution to the (posterior) distribution we are interested in.
- Bounding approaches: bound the conditional probabilites above and below and iteratively reduce the bounds.
- . . .

Factors

- A factor is a representation of a function from a tuple of random variables into a number.
- We write factor f on variables $X_1, \ldots, X_j$ as $f(X_1, \ldots, X_j)$.
- We can assign some or all of the variables of a factor:
  - $f(X_1{=}v_1, X_2, \ldots, X_j)$, where $v_1 \in domain(X_1)$, is a factor on $X_2, \ldots, X_j$.
  - $f(X_1{=}v_1, X_2{=}v_2, \ldots, X_j{=}v_j)$ is a number that is the value of f when each X_i has value v_i.

  The former is also written as $f(X_1, X_2, \ldots, X_j)_{X_1=v_1}$, etc.

Example factors

$r(X, Y, Z)$:

X	Y	Z	val
t	t	t	0.1
t	t	f	0.9
t	f	t	0.2
t	f	f	0.8
f	t	t	0.4
f	t	f	0.6
f	f	t	0.3
f	f	f	0.7

$r(X=t, Y, Z)$:

Y	Z	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

$r(X=t, Y, Z=f)$:

Y	val
t	0.9
f	0.8

$r(X=t, Y=f, Z=f) = 0.8$

Clicker Question

If $f(W, X, Y, Z)$ is a factor on variables $\{W, X, Y, Z\}$, then $f(W, X = 3, Y = true, Z)$ is a factor on

A $\{W, X, Y, Z\}$

B $\{X, Y\}$

C $\{W, Z\}$

D $\{\}$

E none of the above

Clicker Question

If $f(W, X, Y, Z)$ is a factor on variables $\{W, X, Y, Z\}$, then $f(W = 17, X = 3, Y = true, Z = false)$ is a factor on

A $\{W, X, Y, Z\}$

B $\{X, Y\}$

C $\{W, Z\}$

D $\{\}$

E none of the above

Multiplying factors

The product of factor $f_1(\overline{X}, \overline{Y})$ and $f_2(\overline{Y}, \overline{Z})$, where $\overline{Y}$ are the variables in common, is the factor $(f_1 * f_2)(\overline{X}, \overline{Y}, \overline{Z})$ defined by:

$$(f_1 * f_2)(\overline{X}, \overline{Y}, \overline{Z}) \quad = \quad f_1(\overline{X}, \overline{Y}) f_2(\overline{Y}, \overline{Z}).$$

Multiplying factors example

f_1:

A	B	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

f_2:

B	C	val
t	t	0.3
t	f	0.7
f	t	0.6
f	f	0.4

$f_1 * f_2$:

A	B	C	val
t	t	t	0.03
t	t	f	0.07
t	f	t	0.54
t	f	f	0.36
f	t	t	0.06
f	t	f	0.14
f	f	t	0.48
f	f	f	0.32

Clicker Question

If f is a factor on $\{W, X, Y\}$ and
g is a factor on $\{W, U\}$
$(f * g)$ is a factor on

A $\{W, X, Y, U\}$

B $\{X, Y, U\}$

C $\{W\}$

D $\{f, g, W, X, Y, U\}$

E there is not enough information to tell

Clicker Question

If $f(W{=}3, X{=}4, Y{=}5) = 10$ and
$g(W{=}3, U{=}12) = 15$
$(f * g)(W{=}3, X{=}4, Y{=}5, U{=}12) =$

A a factor on $\{W, X, Y, U\}$

B 25

C 150

D none of the above

E there is not enough information to tell

Summing out variables

We can sum out a variable, say X_1 with domain $\{v_1, \ldots, v_k\}$, from factor $f(X_1, \ldots, X_j)$, resulting in a factor on $X_2, \ldots, X_j$ defined by:

$$\begin{aligned}
&(\sum_{X_1} f)(X_2, \ldots, X_j) \\
&= f(X_1{=}v_1, \ldots, X_j) + \cdots + f(X_1{=}v_k, \ldots, X_j)
\end{aligned}$$

Summing out a variable example

f_3:

A	B	C	val
t	t	t	0.03
t	t	f	0.07
t	f	t	0.54
t	f	f	0.36
f	t	t	0.06
f	t	f	0.14
f	f	t	0.48
f	f	f	0.32

$\sum_B f_3$:

A	C	val
t	t	0.57
t	f	0.43
f	t	0.54
f	f	0.46

Exercise

Given factors:

s:

A	val
t	0.75
f	0.25

t:

A	B	val
t	t	0.6
t	f	0.4
f	t	0.2
f	f	0.8

o:

A	val
t	0.3
f	0.1

What are the following a function of?

i) $s * t$

ii) $\sum_B (s * t)$

iii) $s * o$

iv) $\sum_A s * t * o$

v) $\sum_B (\sum_A s * t * o)$

A $\{A\}$

B $\{B\}$

C $\{A, B\}$

D $\{\}$

E none of the above

Queries and Evidence

- To compute the posterior probability of Z given evidence $E{=}e$:

$$
\begin{aligned}
& P(Z \mid E{=}e) \\
&= \frac{P(Z, E{=}e)}{P(E{=}e)} \\
&= \frac{P(Z, E{=}e)}{\sum_Z P(Z, E{=}e).}
\end{aligned}
$$

- So the computation reduces to the probability of $P(Z, E{=}e)$
- then normalize at the end.

Probability of a conjunction

- The variables of the belief network are $X_1, \ldots, X_n$.
- The evidence is $Y_1{=}v_1, \ldots, Y_j{=}v_j$
- To compute $P(Z, Y_1{=}v_1, \ldots, Y_j{=}v_j)$:
  we add the other variables,
  $Z_1, \ldots, Z_k = \{X_1, \ldots, X_n\} - \{Z\} - \{Y_1, \ldots, Y_j\}$.
  and sum them out.
- We order the Z_i into an elimination ordering.

$$
\begin{aligned}
& P(Z, Y_1{=}v_1, \ldots, Y_j{=}v_j) \\
&= \sum_{Z_k} \cdots \sum_{Z_1} P(X_1, \ldots, X_n)_{Y_1=v_1,\ldots,Y_j=v_j}. \\
&= \sum_{Z_k} \cdots \sum_{Z_1} \prod_{i=1}^{n} P(X_i \mid parents(X_i))_{Y_1=v_1,\ldots,Y_j=v_j}.
\end{aligned}
$$

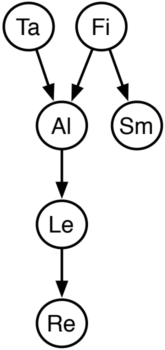
Computing sums of products

Computation in belief networks reduces to computing the sums of products.

- How can we compute $ab + ac$ efficiently?
- Distribute out a giving $a(b + c)$
- How can we compute $\sum_{Z_1} \prod_{i=1}^{n} P(X_i \mid parents(X_i))$ efficiently?
- Distribute out those factors that don't involve Z_1.

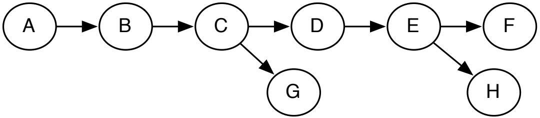
Inference as factorization example

Query: $P(Re \mid SM{=}true)$ write $SM{=}true$ as sm



$$
\begin{aligned}
& P(Re, sm) \\
&= \sum_{Le}\sum_{Al}\sum_{Fi}\sum_{Ta} P(Ta, Fi, Al, sm, Le, Re) \\
&= \sum_{Le}\sum_{Al}\sum_{Fi}\sum_{Ta} P(Ta)P(Fi)P(Al \mid Ta, Fi) \\
&\qquad P(sm \mid Fi)P(Le \mid Al)P(Re \mid Le) \\
&= \sum_{Le}\sum_{Al}\sum_{Fi} P(Fi)P(sm \mid Fi)P(Le \mid Al) \\
&\qquad P(Re \mid Le)\sum_{Ta} P(Ta)P(Al \mid Ta, Fi) \\
&= \sum_{Le} P(Re \mid Le)\sum_{Al} P(Le \mid Al) \\
&\qquad \sum_{Fi} P(Fi)P(sm \mid Fi)\sum_{Ta} P(Ta)P(Al \mid Ta, Fi)
\end{aligned}
$$

Inference as factorization example



Query: $P(G \mid f)$; elimination ordering: A, H, E, D, B, C

$$P(G \mid f) \propto \sum_C \sum_B \sum_D \sum_E \sum_H \sum_A P(A)P(B \mid A)P(C \mid B) P(D \mid C)P(E \mid D)P(f \mid E)P(G \mid C)P(H \mid E)$$

$$= \sum_C \left(\sum_B \left(\sum_A P(A)P(B \mid A) \right) P(C \mid B) \right) P(G \mid C) \left(\sum_D P(D \mid C) \left(\sum_E P(E \mid D)P(f \mid E) \sum_H P(H \mid E) \right) \right)$$

Variable elimination algorithm

To compute $P(Z \mid Y_1{=}v_1 \wedge \ldots \wedge Y_j{=}v_j)$:

- Construct a factor for each conditional probability.
- Set the observed variables to their observed values.
- Sum out each of the non-observed non-query variables (the $\{Z_1, \ldots, Z_k\}$) according to some elimination ordering.
- Multiply the remaining factors.
- Normalize by dividing the resulting factor $f(Z)$ by $\sum_Z f(Z)$.

Summing out a variable

To sum out a variable Z_j from a product $f_1, \ldots, f_k$ of factors:

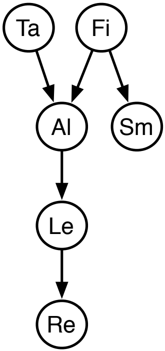
- Partition the factors into
  - those that don't contain Z_j, say $f_1, \ldots, f_i$,
  - those that contain Z_j, say $f_{i+1}, \ldots, f_k$

Then:

$$\sum_{Z_j} f_1 * \cdots * f_k = f_1 * \cdots * f_i * \left(\sum_{Z_j} f_{i+1} * \cdots * f_k \right).$$

- Explicitly construct a representation of the rightmost factor. Replace the factors $f_{i+1}, \ldots, f_k$ by the new factor.

Inference as factorization example



Query: $P(Re \mid SM{=}true)$.
See AIspace.org → Belief and decision network tool → File → Load Sample Problem → Fire Alarm Belief Network → Load → Solve → Make Observation Smoke → Query Report → Verbose