CS 4260 and CS 5260
Vanderbilt University

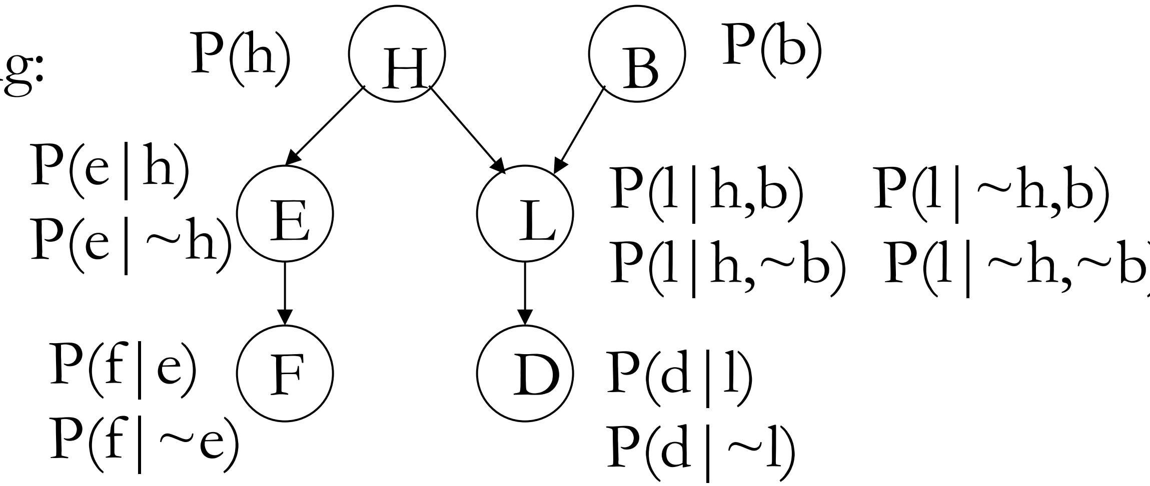
Lecture on Uncertainty
(Inference with Belief Networks)

This lecture assumes that you have
- Read Section 8.1 through 8.3 of ArtInt (though there is some repetition, as well as additional material)
- Seen previous two lectures on probability and belief network semantics

ArtInt: Poole and Mackworth, Artificial Intelligence 2E
 at http://artint.info/2e/html/ArtInt2e.html
to include slides at http://artint.info/2e/slides/ch04/lect1.pdf

Douglas H. Fisher

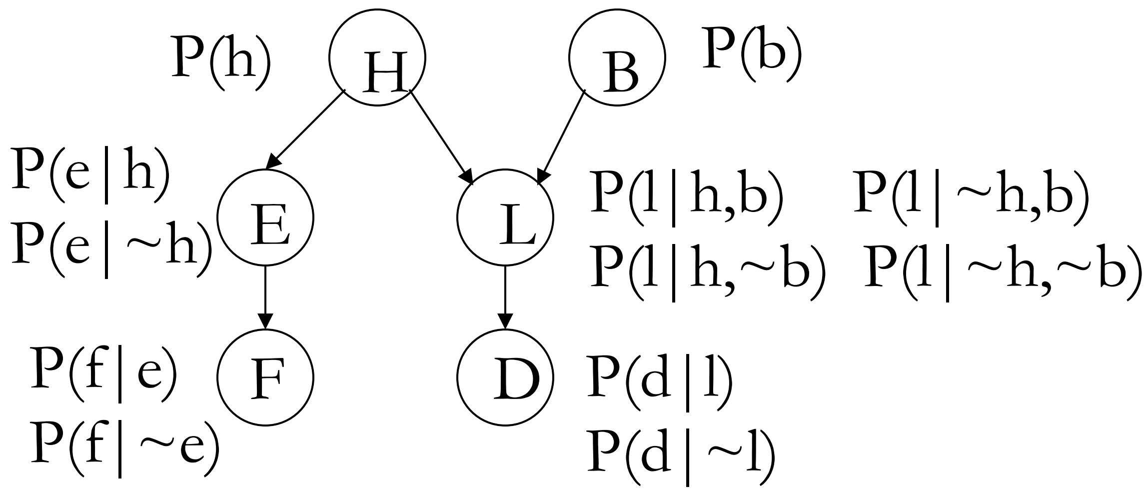
Consider the following:



H: Hardware problems (h) or not (~h)
B: Bugs in code (b) or not (~b)
E: Editor running (e) or not (~e)
L: Lisp interpreter running (l) or not (~l)
F: Cursor flashing (f) or not (~f)
D: prompt displayed (d) or not (~d)

Note that H,B,E,L,F,D order is consistent with the BN (assuming it was constructed using the method of the previous lecture), but so are others such as B,H,L,D,E,F (a node comes after all its ancestors)

Consider the following:



P(h, ~b, e, ~l, f, d)? *Joint probability*

always true

$$= P(h)P(\sim b|h)P(e|h, \sim b)P(\sim l|h, \sim b, e)P(f|h, \sim b, e, \sim l)P(d|h, \sim b, e, \sim l, f)$$

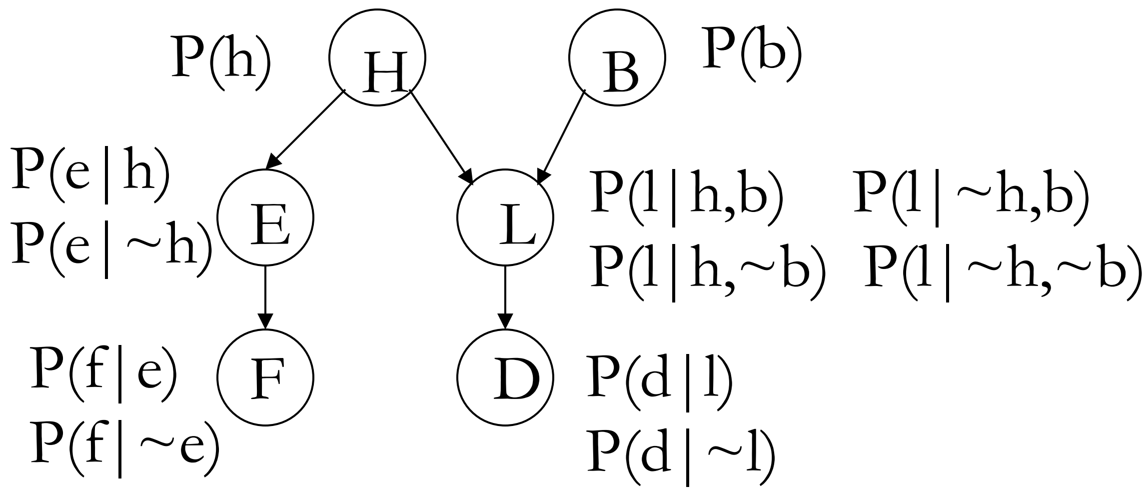
true given BN above

$$= P(h)P(\sim b)P(e|h)P(\sim l|h, \sim b)P(f|e)P(d|\sim l)$$

1-P(b) 1-P(l|h,~b)

Computing a joint probability over specific value assignments to all the variables is straightforward; it is simply the product of all the (conditional) probabilities in the probability tables that reflect those value assignments.

Another joint probability computation



$P(\sim h, b, e, l, \sim f, \sim d)$?

always true

Will typically not have to show this intermediate step, but may be helpful for partial credit, if you are clear

$$= P(\sim h)P(b|\sim h)P(e|\sim h, b)P(l|\sim h, b, e)P(\sim f|\sim h, b, e, l)P(\sim d|\sim h, b, e, l, \sim f)$$

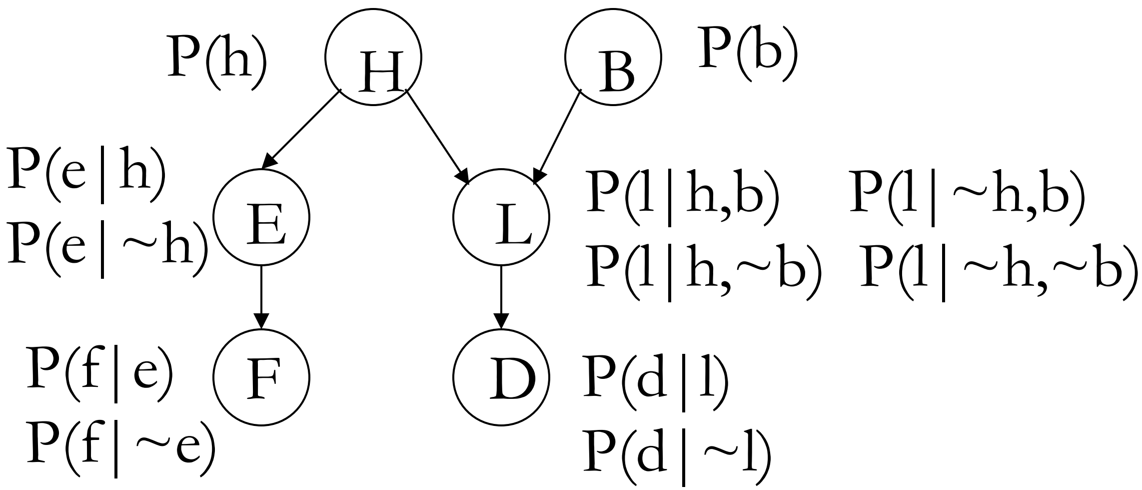
true given BN above

$$= P(\sim h)P(b)P(e|\sim h)P(l|\sim h,b)P(\sim f|e)P(\sim d|l)$$

1-P(h)   1-P(f|e)   1-P(f|e)

On an exam, you need not rewrite negated P(~x|…) as 1-P(x|…), unless I say otherwise. In most of what follows, I will not do that rewriting.

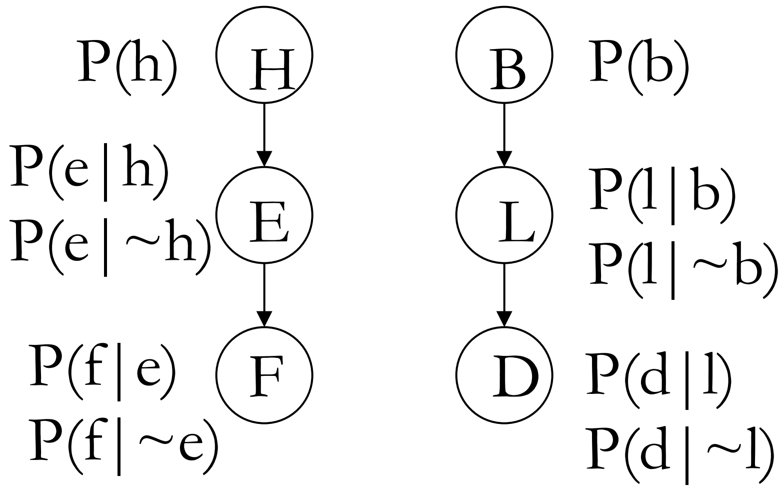
What about joint probabilities over a subset of variables



P(~h, e, ~f)?	P(~h, e, ~f) = P(~h)P(e|~h)P(~f|e)
P(~h, e)?	P(~h, e) = P(~h)P(e|~h)
P(h, b, ~l, d)?	P(h, b, ~l, d) = P(h)P(b)P(~l|h, b)P(d|~l)
P(h, ~b, l)?	P(h, ~b, l) = P(h)P(~b)P(l|h, ~b)
P(h, ~b)?	P(h, ~b) = P(h)P(~b)

If a variable is included in the set, then all of its parents are too

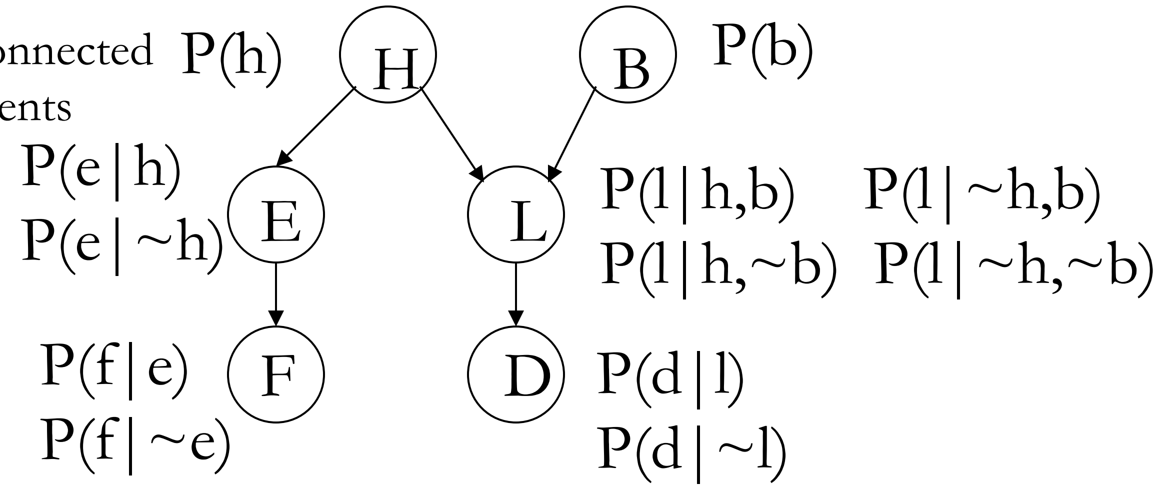
Douglas H. Fisher



A network can be disconnected, but computation again straightforward if each variable with a value in the joint probability has assigned values for its parents in the set too.

$$P(e, h, \sim l, b) = P(e|h)P(h)P(\sim l|b)P(b)$$

What about joint probabilities of connected variables, some with unassigned parents



We must use only probabilities in tables (or trivially computed).

Expand scope of involved variables to include the ancestors of assigned variables, and "sum out" these unassigned variables.

P(e, ~f)? H, the parent of E, isn't assigned

P(e, ~f)? P(e, ~f) = P(e)P(~f|e) But P(e) not anywhere in tables

(e ∧ h) and (e ∧ ~h) are mutually exclusive

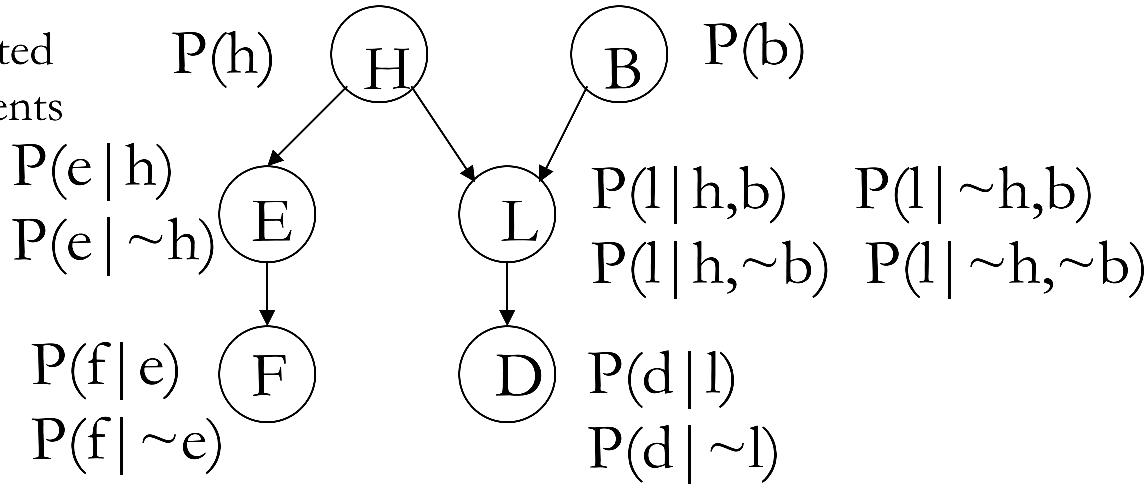
P(e) = P(e, h) + P(e, ~h)
= P(e|h)P(h) + P(e|~h)P(~h)

P(e, ~f) = P(e)P(~f|e) = [P(e|h)P(h) + P(e|~h)P(~h)]P(~f|e)

P(e, ~f) expressed in terms of ONLY probabilities found in BN tables (or trivially computed from them). This is the answer I am typically looking for.

Douglas H. Fisher

Another joint probability of connected variables, some with unassigned parents



$P(\sim l, \sim d)$?

$P(\sim l, \sim d) = P(\sim l)P(\sim d|\sim l)$

$= [(P(\sim l,h,b)+P(\sim l,h,\sim b)+P(\sim l,\sim h,b)+P(\sim l,\sim h,\sim b)]\ P(\sim d|\sim l)$

$= [(P(\sim l|h,b)P(h)P(b) + P(\sim l|h,\sim b)P(h)P(\sim b) + P(\sim l|\sim h,b)P(\sim h)P(b) + P(\sim l|\sim h,\sim b)P(\sim h)P(\sim b)]\ P(\sim d|\sim l)$

Expressed in terms of ONLY probabilities found in BN tables (or trivially computed from them).

Expand scope of involved variables to include the ancestors of assigned variables, and "sum out" these unassigned variables.

Or we could have written

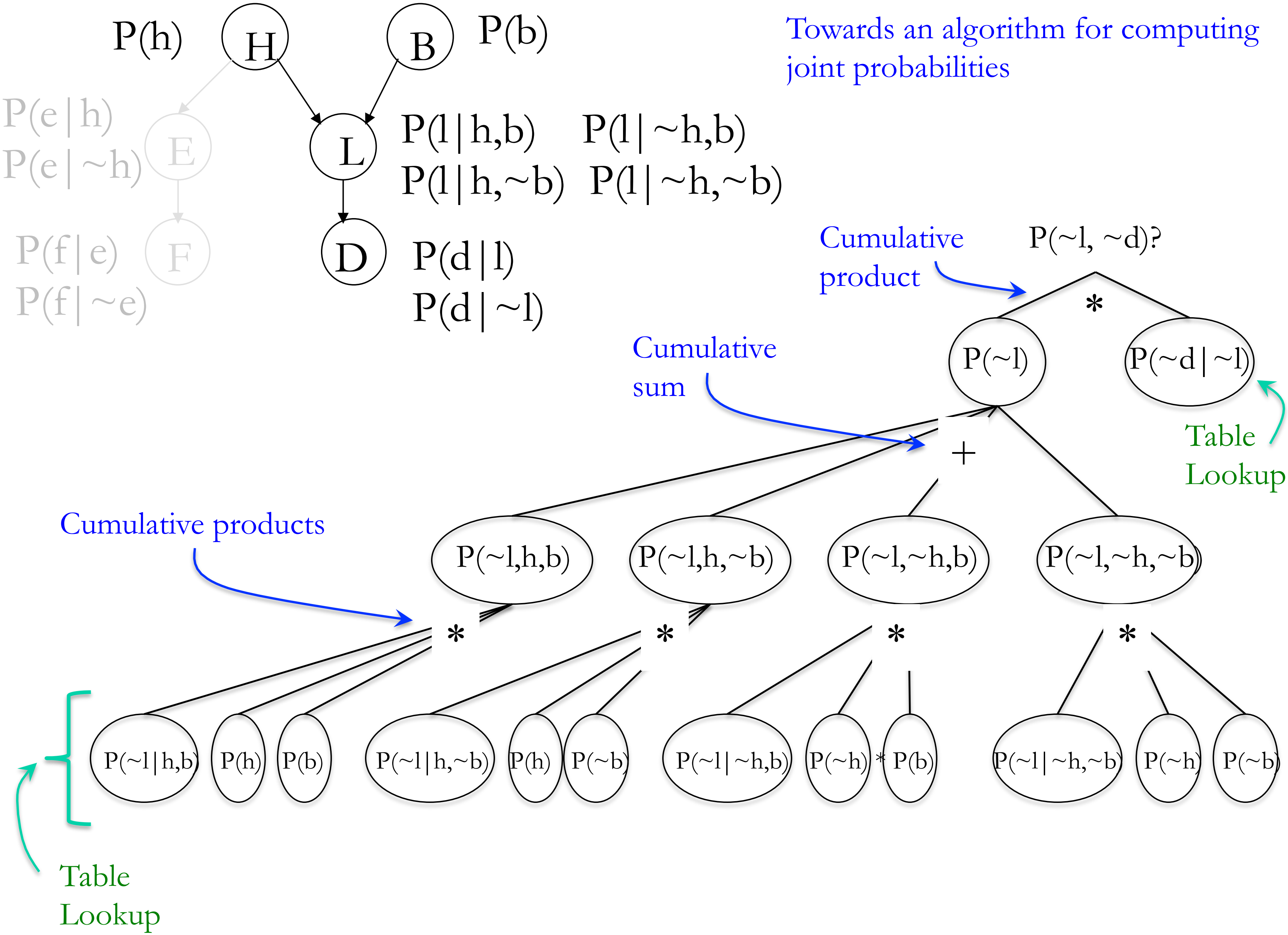
$P(\sim l, \sim d) = P(h,b,\sim l,\sim d) + P(h,\sim b,\sim l,\sim d) + P(\sim h,b,\sim l,\sim d) + P(\sim h,\sim b,\sim l,\sim d)$

$= P(\sim l|h,b)P(h)P(b)P(\sim d|\sim l) + \ldots + P(\sim l|\sim h,\sim b)P(\sim h)P(\sim b)P(\sim d|\sim l)$

$= [(P(\sim l|h,b)P(h)P(b) + P(\sim l|h,\sim b)P(h)P(\sim b) + P(\sim l|\sim h,b)P(\sim h)P(b) + P(\sim l|\sim h,\sim b)P(\sim h)P(\sim b)]\ P(\sim d|\sim l)$

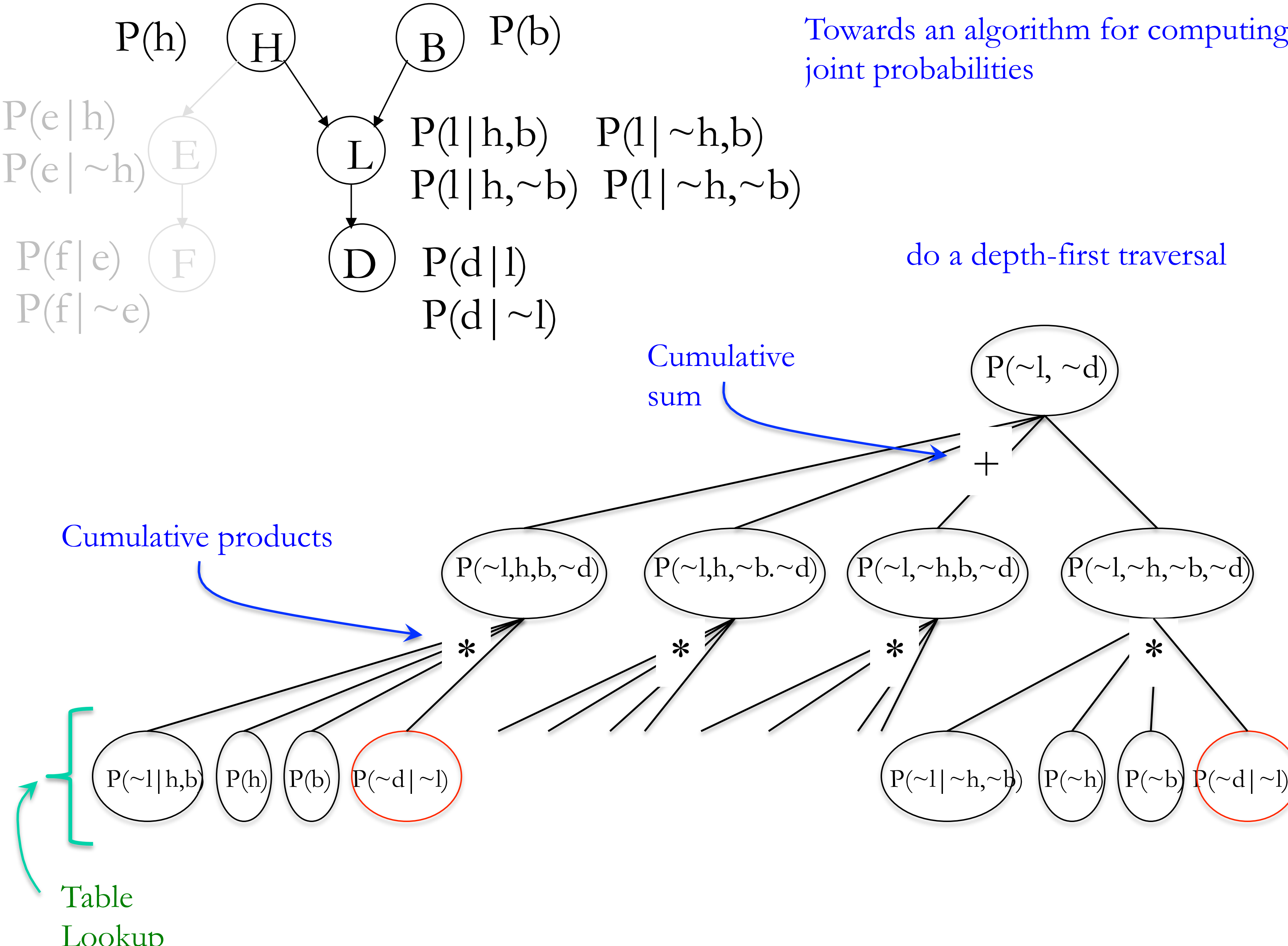
Douglas H. Fisher

Towards an algorithm for computing joint probabilities



Douglas H. Fisher

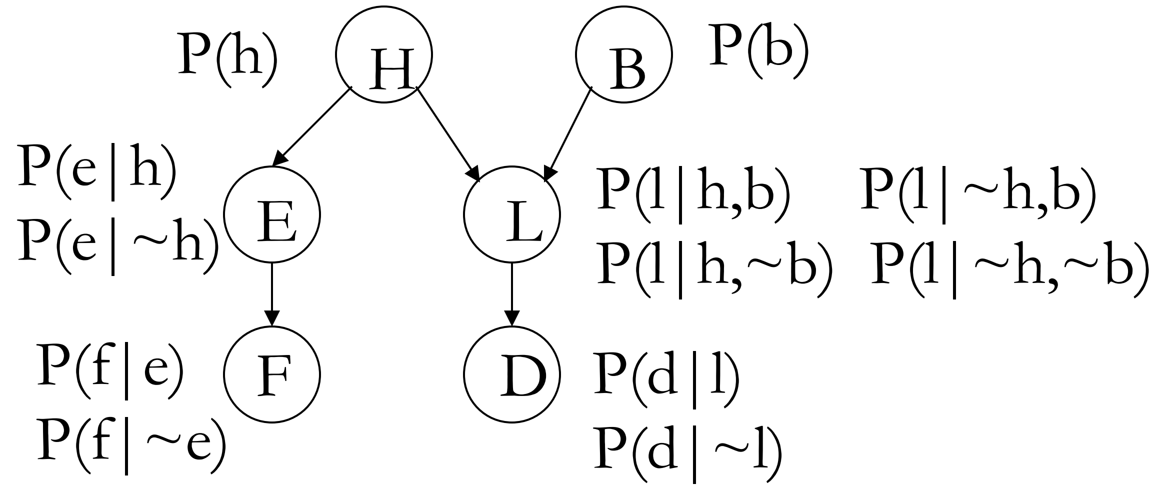
Towards an algorithm for computing joint probabilities



Douglas H. Fisher

P(b, ~l, ~d)?

This is very much like the previous problem, but one of L's parents is assigned (i.e., B), and thus we need only sum over values of H



$$P(b, \sim l, \sim d) = P(b, \sim l)P(\sim d \mid \sim l)$$

$$= [(P(\sim l, h, b) + P(\sim l, \sim h, b)]\ P(\sim d \mid \sim l)$$

$$= [(P(\sim l \mid h, b)P(h)P(b) + P(\sim l \mid \sim h, b)P(\sim h)P(b)]\ P(\sim d \mid \sim l)$$

Expressed in terms of ONLY probabilities found in BN tables (or trivially computed from them).

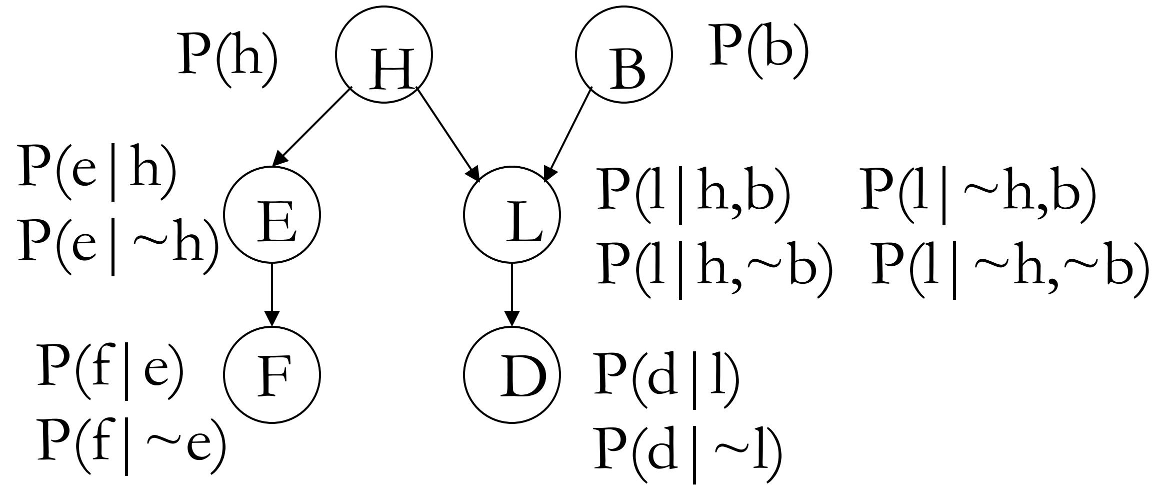
if more explanation needed

Given BN structure

$$P(b, \sim l, \sim d) = P(b)\ P(\sim l \mid b)P(\sim d \mid b, \sim l) = P(b, \sim l)P(\sim d \mid \sim l)$$

a general equality

Douglas H. Fisher



P(e)?

We must use only probabilities in tables (or trivially computed).

Expand scope to involve E's parents, and "sum out" all the values of E.

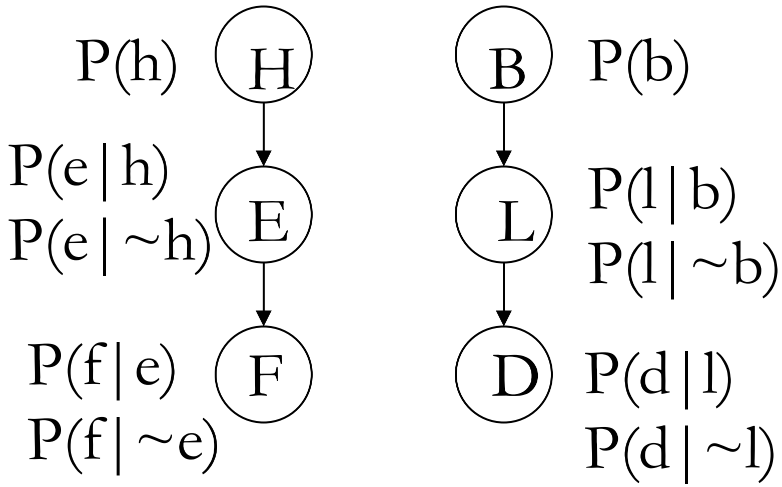
P(e) = P(e, h) + P(e, ~h) = P(e|h)P(h) + P(e|~h)P(~h)

P(~f)? You do this one.

How about computing a joint probability of nodes not directly connected? Sum over their ancestor nodes.

P(e, l) = P(e, h, l, b) + P(e, h, l, ~b) + P(e, ~h, l, b) + P(e, ~h, l, ~b) = finish it

A network can be disconnected, but same rule of summing over ancestors applies.



$$P(e, l) = P(e, l, h, b) + P(e, l, h, \sim b) + P(e, l, \sim h, b) + P(e, l, \sim h, \sim b)$$

$$= P(e, l, h, b) + P(e, l, h, \sim b) + P(e, l, \sim h, b) + P(e, l, \sim h, \sim b)$$

$$= P(e|h)P(h)P(l|b)P(b) + P(e|h)P(h)P(l|\sim b)P(\sim b) + P(e|\sim h)P(\sim h)P(l|b)P(b) + P(e|\sim h)P(\sim h)P(l|\sim b)P(\sim b)$$

if more explanation needed

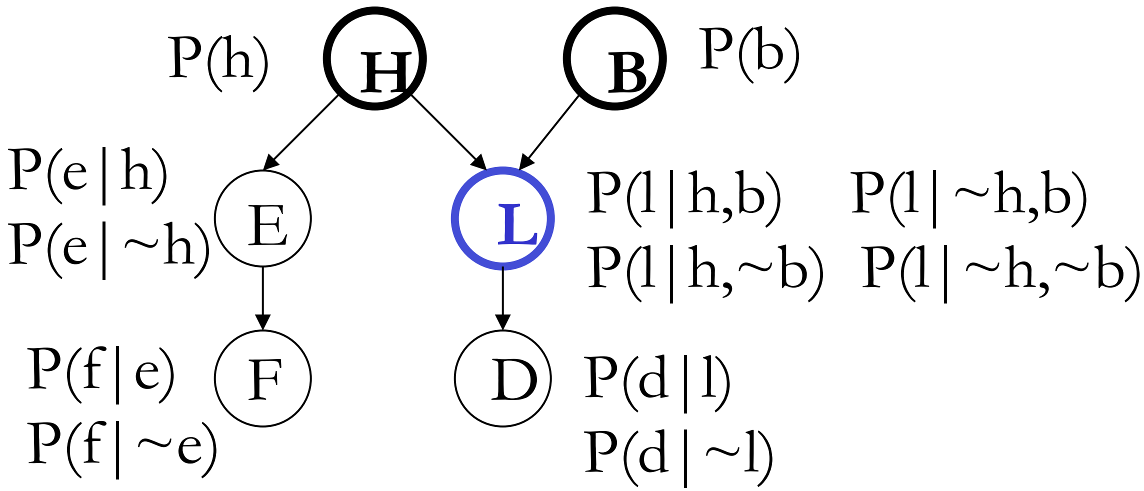
Expressed in terms of ONLY probabilities found in BN tables (or trivially computed from them).

$$P(e, l, h, b) = P(h)P(b|h)P(l|h, b)P(e|h, b, l) = P(h)P(b)P(l|b)P(e|h)$$

Chain rule BN structure

Douglas H. Fisher

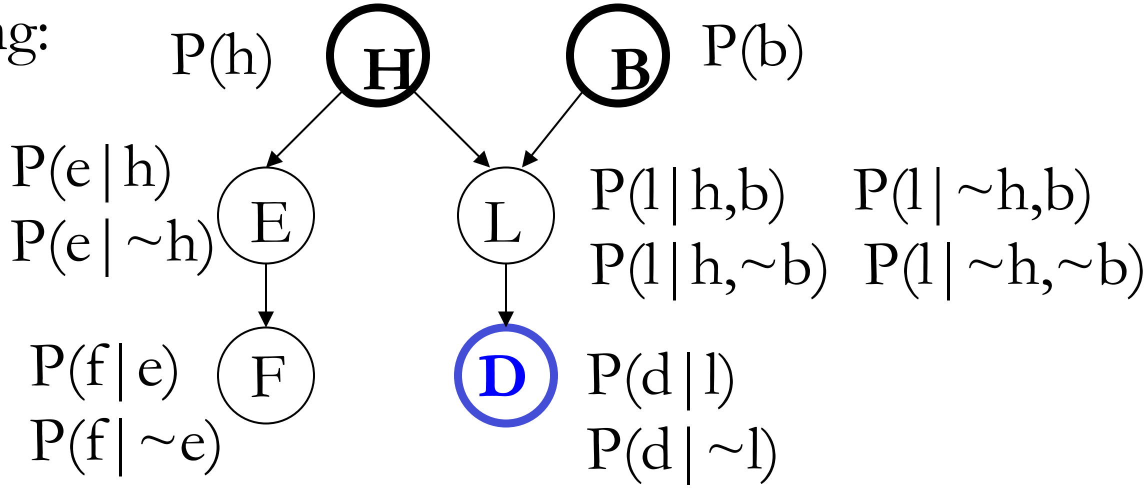
We know how to compute joint probabilities, so how about computing conditional probabilities?



P(~l|h,~b)?   1 – P(l|h,~b)   (essentially table lookup in this example)

If the network was constructed using an ordering of variables from causes to manifestations, then this is inference of manifestation (and "hidden" or intermediate) variables from causes.

Consider the following:



We must use only probabilities in tables (or trivially computed). *Expand scope of involved variables to fill gaps between assigned variables, and "sum out" these unassigned variables.*

P(d|h,~b)?

A conditional probability can be expressed as two joint probabilities

= P(d, h, ~b) / P(h,~b))

Note that P(h, ~b) isn't referenced in final expression of P(d|h,~b), because their values have been fixed, and (h,~b) assumed true (with probability 1.0)

= [P(d, l,h,~b)+P(d,~l,h,~b)]/P(h,~b)

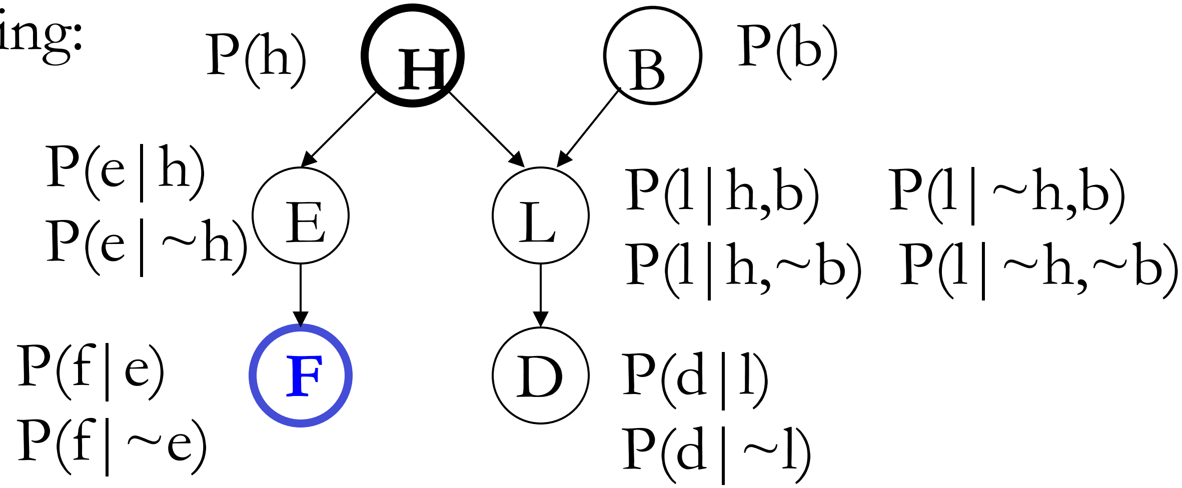
= [P(d|l,h,~b)P(l,h,~b)+P(d|~l,h,~b) P(~l,h,~b)]/P(h,~b)

= [P(d|l)P(l|h,~b)P(h,~b)+P(d|~l) P(~l|h,~b)P(h,~b)]/P(h,~b)

= P(d|l)P(l|h,~b)+P(d|~l) P(~l|h,~b)

Douglas H. Fisher

Consider the following:



$P(f|h)$?

$= P(f, h) \ / \ P(h))$

$= [P(f, e, h) + P(f, \sim e, h)]/P(h)$

$= [P(f|h, e)P(e|h)P(h) + P(f|h, \sim e)\ P(\sim e|h)P(h)]/P(h)$

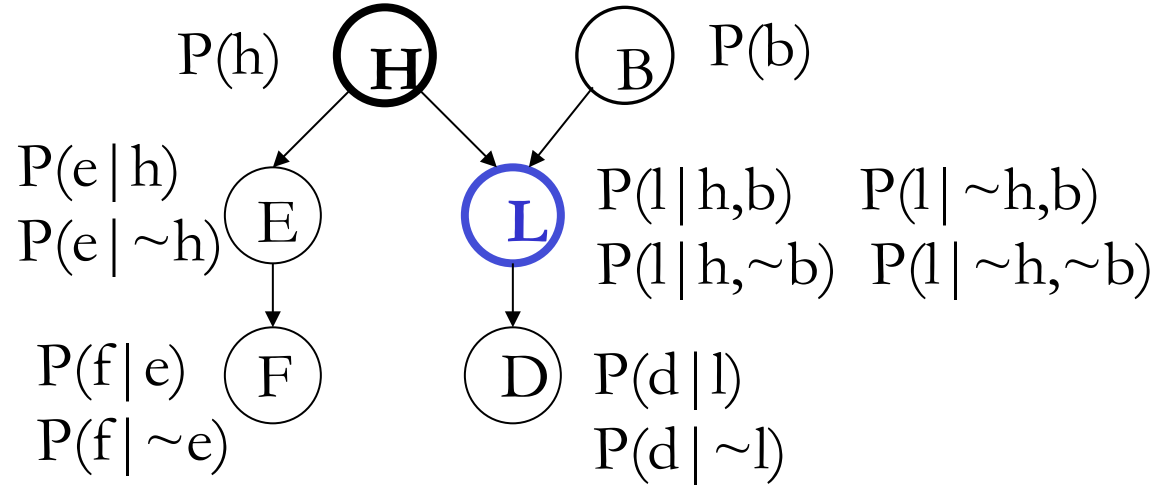
$= [P(f|e)P(e|h)P(h) + P(f|\sim e)P(\sim e|h)P(h)]/P(h)$

Note that P(h) isn't referenced in final expression of P(f|h), because their values have been fixed, and (h) assumed true (with probability 1.0)

$= P(f|e)P(e|h) + P(f|\sim e)\ P(\sim e|h)$

Douglas H. Fisher

Consider the following:



$P(\sim l \mid h)?$

$= P(\sim l, h) / P(h)$

$= [P(\sim l, h, b) + P(\sim l, h, \sim b)] \;/\; P(h)$ *We must involve B*

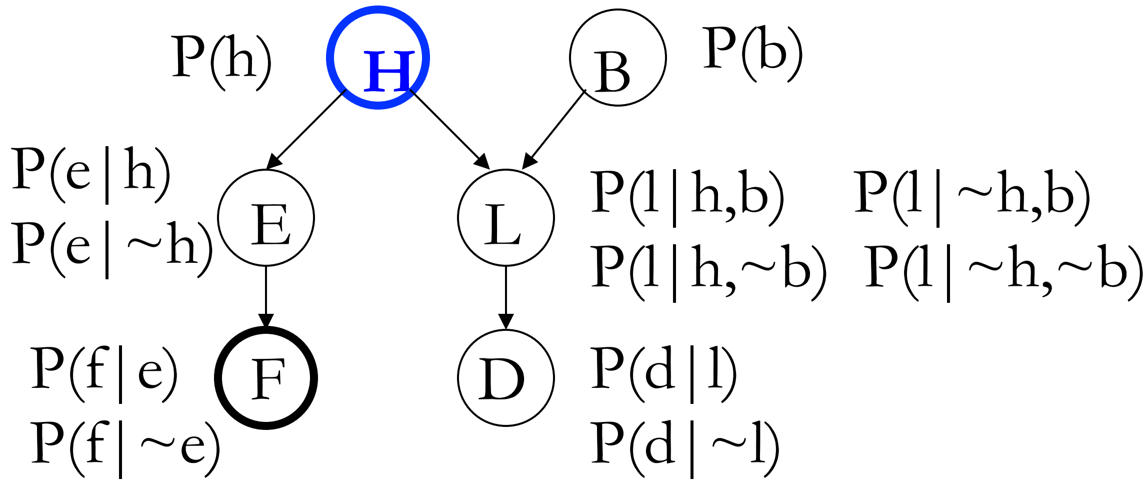
$= [P(\sim l \mid h, b)P(h, b) + P(\sim l \mid h, \sim b)P(h, \sim b)] \;/\; P(h)$

$= [P(\sim l \mid h, b)\textcolor{red}{P(h)}P(b) + P(\sim l \mid h, \sim b)\textcolor{red}{P(h)}P(\sim b)] \;/\; \textcolor{red}{P(h)}$

$= P(\sim l \mid h, b)P(b) + P(\sim l \mid h, \sim b)P(\sim b)$ *Final expression does not reference P(h), reflecting that H is fixed*

Douglas H. Fisher

We can reason about causes from manifestations:



P(h|f)?

$= P(h, f)/P(f)$

$= [P(h,e,f)+P(h,\sim e,f)]/P(f)$ *We must involve E – fill the gap*

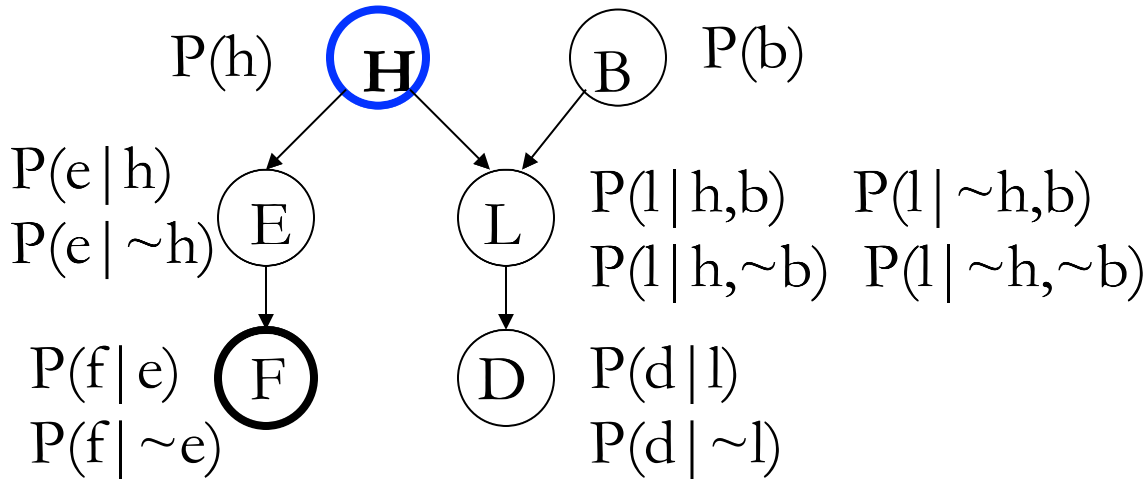
$= [P(f|h,e)P(e|h)P(h) + P(f|h,\sim e)P(\sim e|h)P(h)]/P(f)$

$= [P(f|e)P(e|h)P(h) + P(f|\sim e)P(\sim e|h)P(h)]/P(f)$

$= [P(f|e)P(e|h)P(h) + P(f|\sim e)P(\sim e|h)P(h)]/[P(f,e,h)+P(f,e,\sim h)+P(f,\sim e,h)+P(f,\sim e,\sim h)]$

$= [P(f|e)P(e|h)P(h) + P(f|\sim e)P(\sim e|h)P(h)]$
$/ [P(f|e)P(e|h)P(h) + P(f|e)P(e|\sim h)P(\sim h) + P(f|\sim e)P(\sim e|h)P(h) + P(f|\sim e)P(\sim e|\sim h)P(\sim h)]$

Notice that there is some repeated computation



$P(h|f)$?

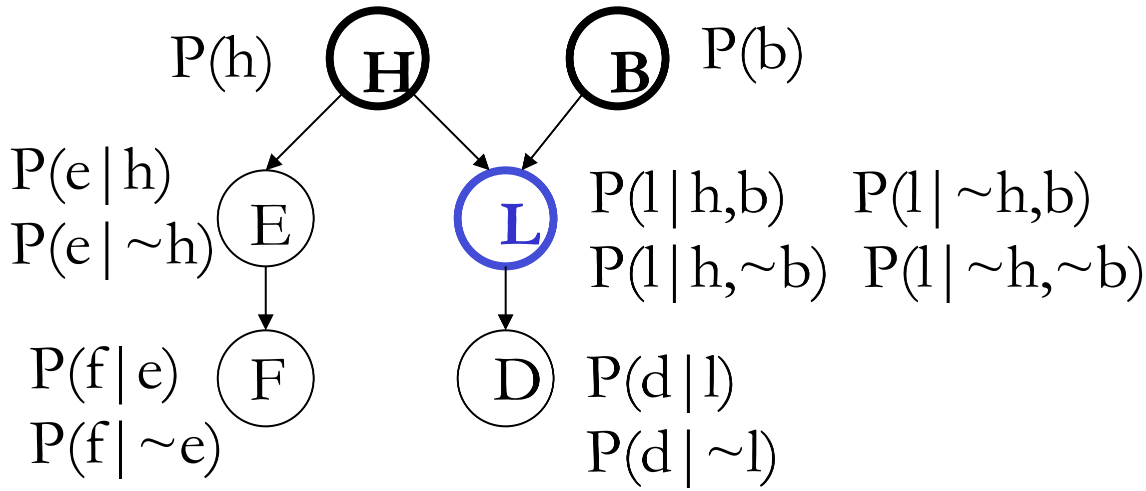
$= P(h, f)/P(f)$

$= [P(h,e,f)+P(h,\sim e,f)]/P(f)$

$= [P(f|e)P(e|h)P(h) + P(f|\sim e)P(\sim e|h)P(h)]/[P(f,e,h)+P(f,e,\sim h)+P(f,\sim e,h)+P(f,\sim e,\sim h)]$

$= [P(f|e)P(e|h)P(h) + P(f|\sim e)P(\sim e|h)P(h)]$
$/ [P(f|e)P(e|h)P(h) + P(f|e)P(e|\sim h)P(\sim h) + P(f|\sim e)P(\sim e|h)P(h) + P(f|\sim e)P(\sim e|\sim h)P(\sim h)]$

L induces *conditional dependence* on H and B



P(h|l)?

= P(h, l)/P(l)

= [P(l|h,b)P(h)P(b) + P(l|h,~b)P(h)P(~b)] / P(l)

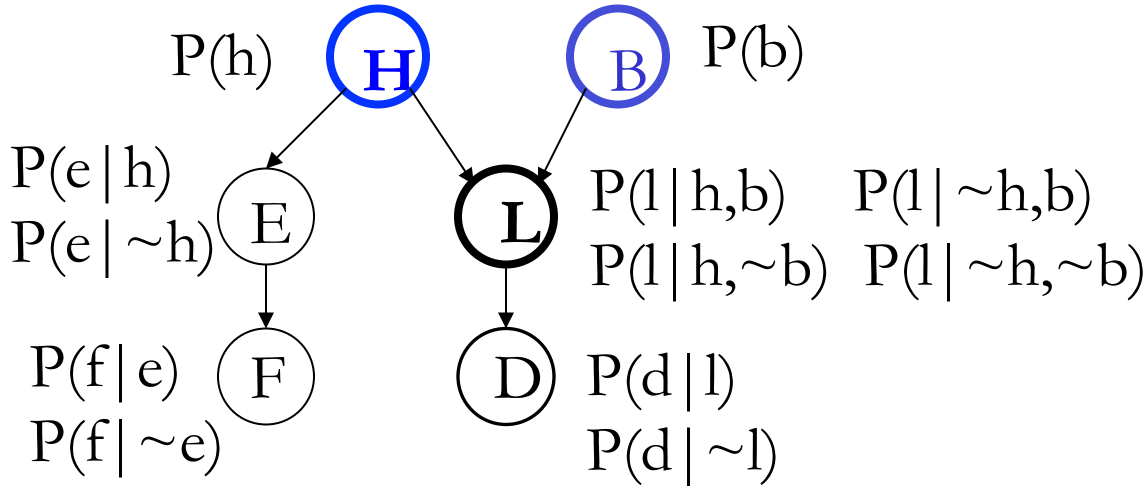
Don't need to expand P(l) out to make a point about conditional dependence

P(b|l)?

= P(b, l)/P(l)

= [P(l|h,b)P(h)P(b) + P(l|~h,b)P(~h)P(b)] / P(l)

L induces *conditional dependence* on H and B



$P(h, b|l)$?

$= P(h, b, l)/P(l)$

$= [P(l|h,b)P(b)P(h) / P(l)$

Does $P(h, b \mid l) = P(h \mid l) * P(b \mid l)$? from previous slide

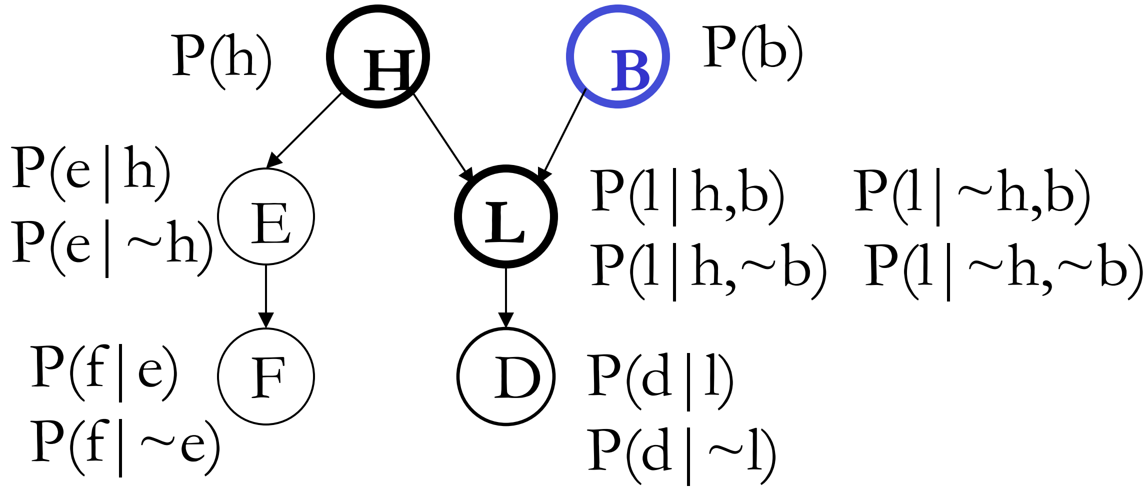
$P(h|l, b) = P(h|l)$?

$P(b|l, h) = P(b|l)$? $P(b|h,l)$ on next slide

Equivalent definitions of conditional independence

These equalities are not true; knowing L induces *dependence* on H and B. For example, if L is known, and we are additionally told H, this changes the probability of B (and B would change H)

Consider the following:



$P(b|h, l)$?

$$= P(b,h,l)/P(h,l)$$

$$= P(l|h,b)P(b)P(h) \;/\; [P(l|h,b)P(b)P(h) + P(l|h, {\sim}b)P({\sim}b)P(h)]$$

$$= P(l|h,b)P(b)P(h) \;/\; [P(l|h,b)P(b)P(h) + P(l|h, {\sim}b)P({\sim}b)P(h)]$$

$$= P(l|h,b)P(b) \;/\; [P(l|h,b)P(b) + P(l|h, {\sim}b)P({\sim}b)]$$

If H and B both have causal influence on L (and D), then a known value of one (e.g., H), will typically alter the probability of the other (e.g., B) conditioned on H's value. For example. If h and b each make l more likely, then if we know h

Aside: Show equivalence between definitions of conditional independence. For example, show that

$P(h, b | l) = P(h | l) * P(b | l)$ implies $P(b | l, h) = P(b | l)$

$P(h, b \mid l) = P(h \mid l) * P(b \mid l)$

Multiply both sides by P(l)

$P(l)P(h, b \mid l) = P(l)P(h \mid l) * P(b \mid l)$

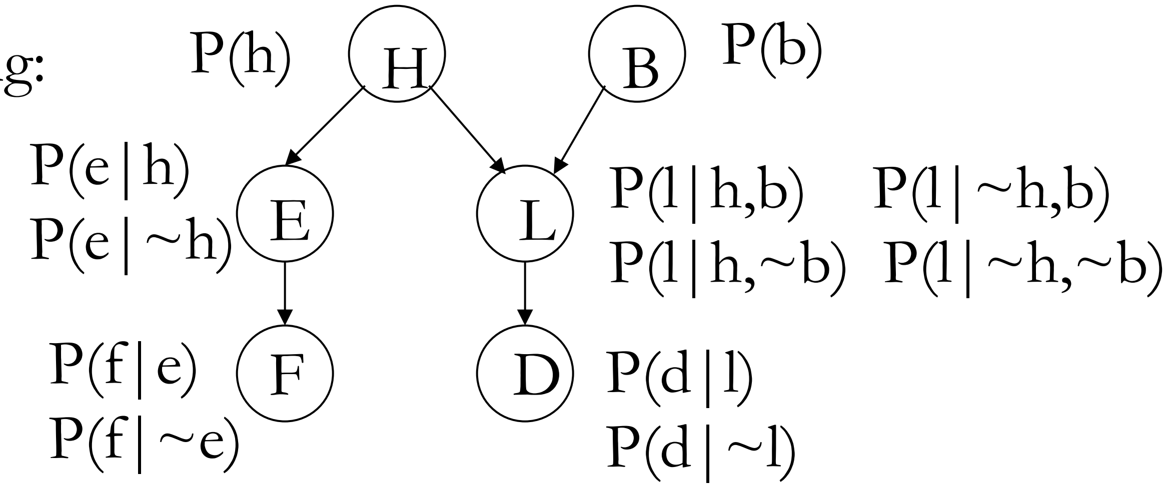
$P(h, b, l) = P(h, l) * P(b | l)$

Divide both sides by P(h, l)

$P(h, b, l)/P(h, l) = P(b | l)$

$P(b | h, l) = P(b | l)$

Consider the following:



P(f | l) = ?

P(f | b, l) = ?

P(f | l, h) = ?

P(d | h) = ?

P(f | b) = ?

P(h|l)? = ?

Douglas H. Fisher