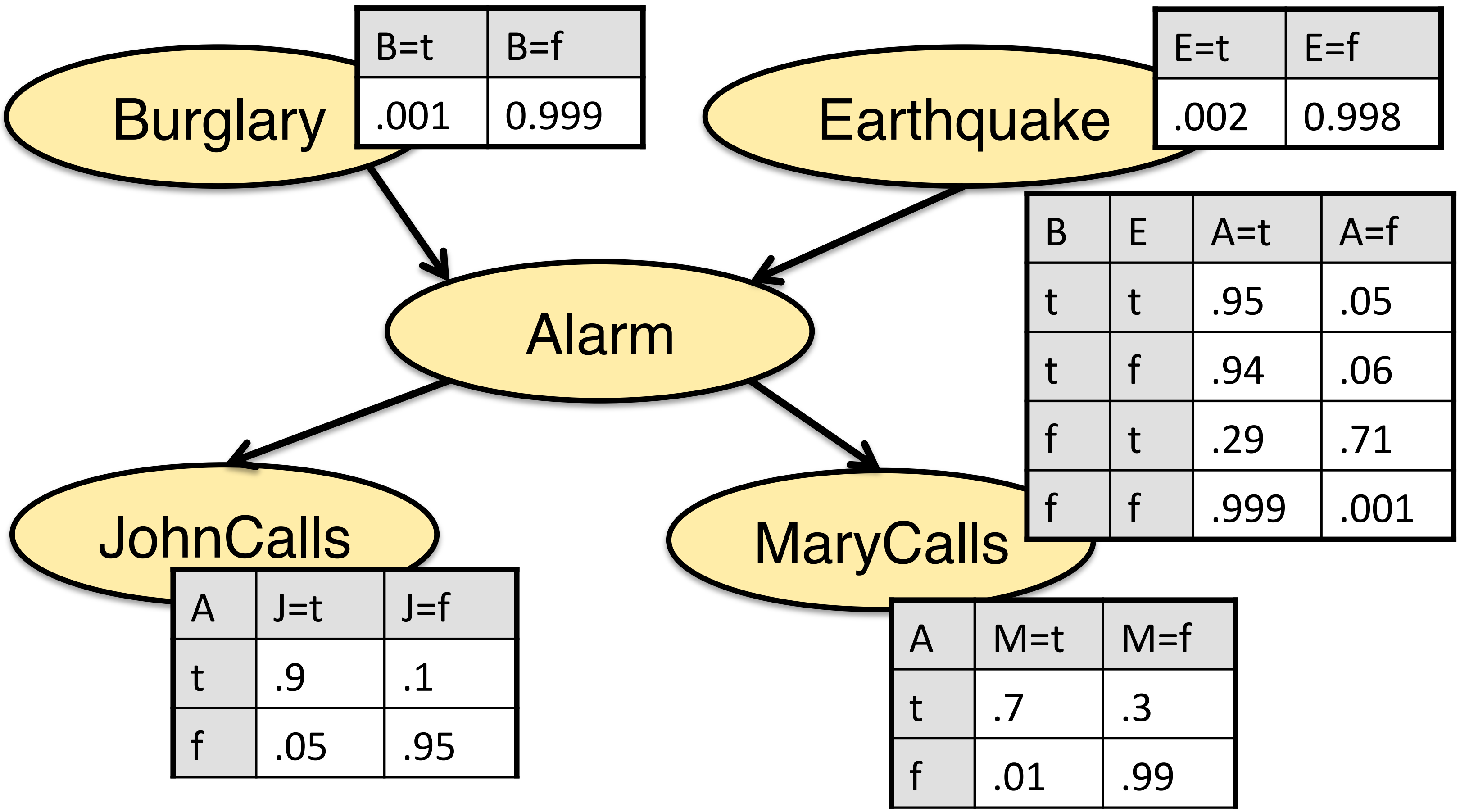
CS440/ECE448: Intro to Artificial Intelligence

Lecture 18: Bayesian Networks continued

Prof. Julia Hockenmaier
juliahmr@illinois.edu

http://cs.illinois.edu/fa11/cs440

The *Burglary* example



B=t	B=f
.001	0.999

E=t	E=f
.002	0.998

B	E	A=t	A=f
t	t	.95	.05
t	f	.94	.06
f	t	.29	.71
f	f	.999	.001

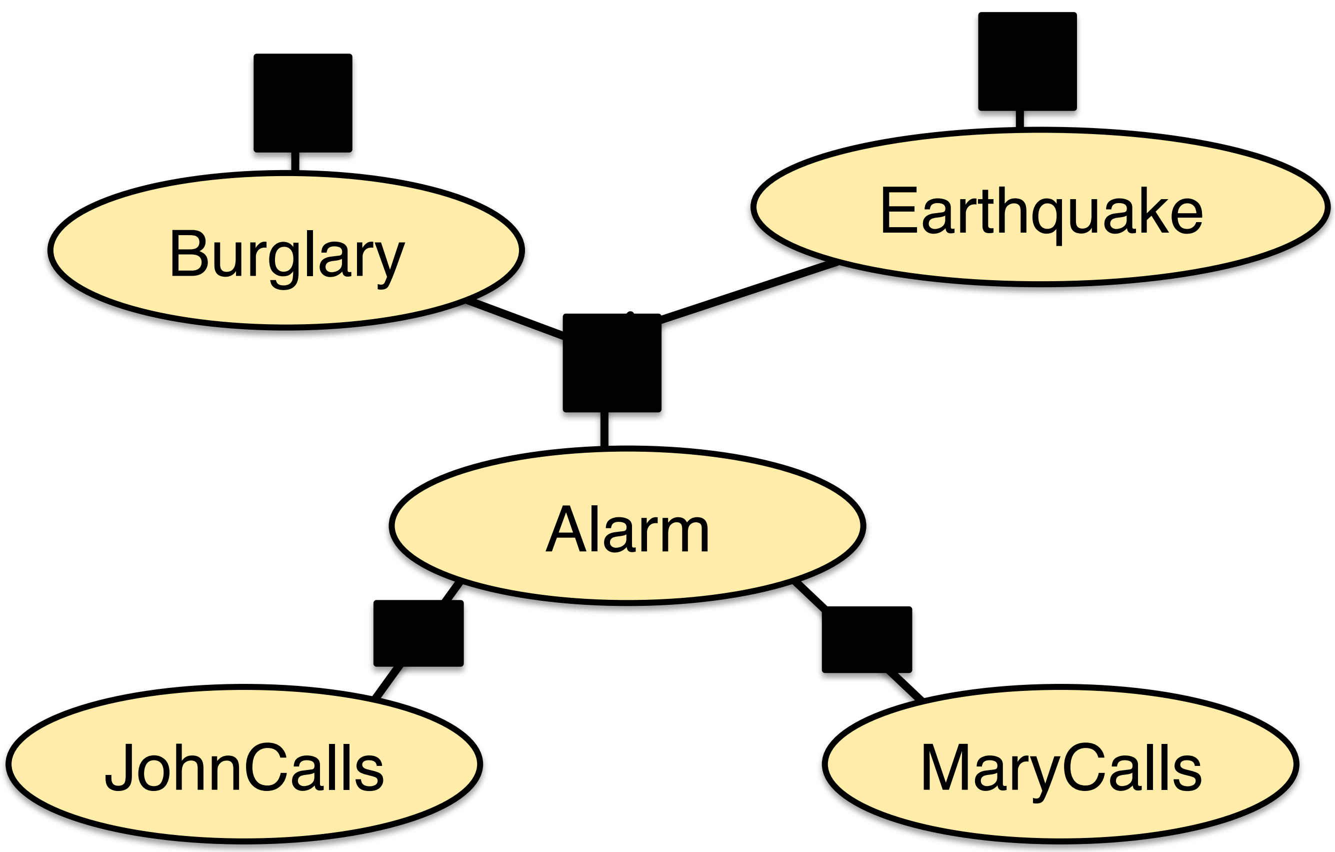
A	J=t	J=f
t	.9	.1
f	.05	.95

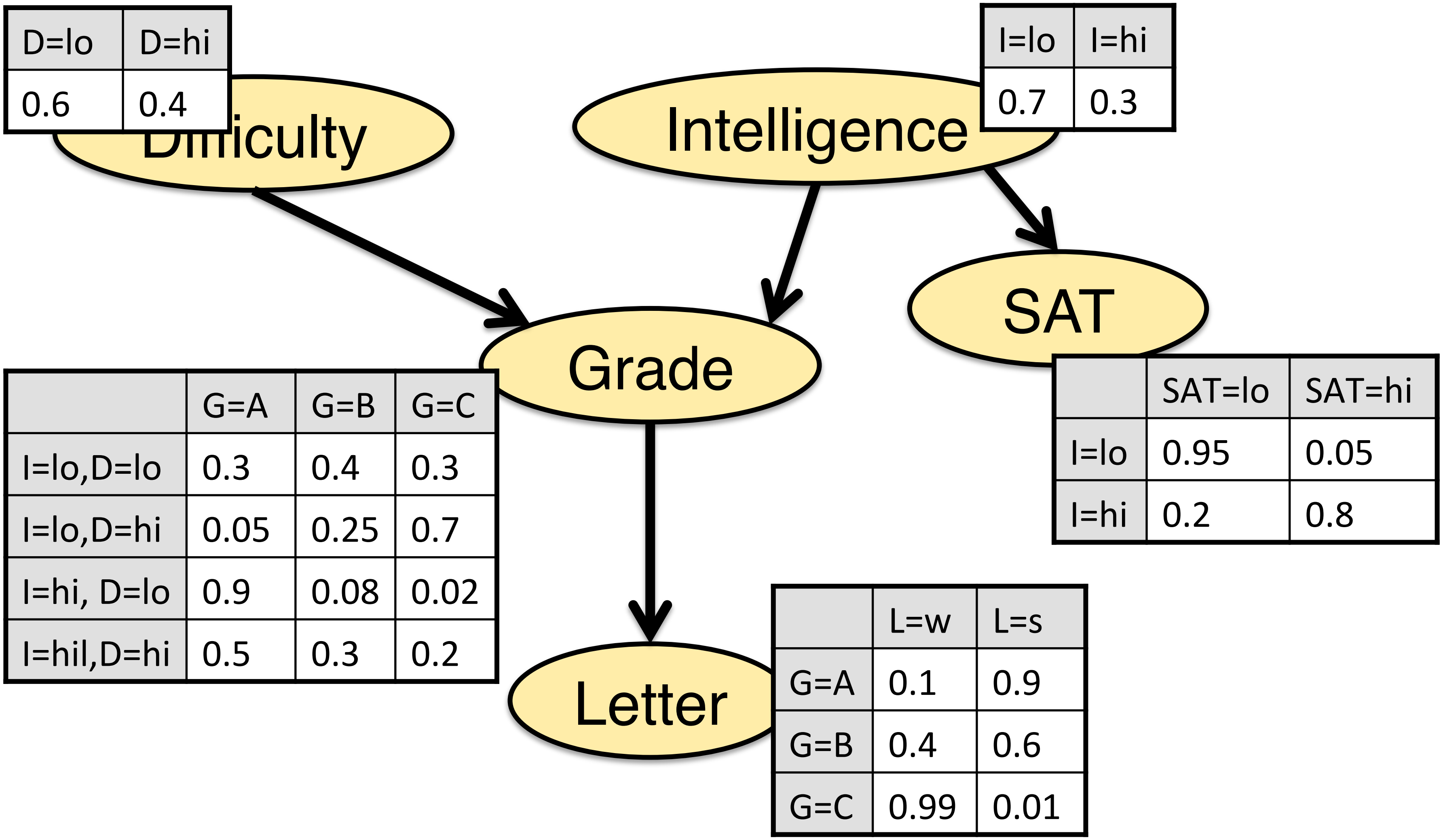
A	M=t	M=f
t	.7	.3
f	.01	.99

What is the probability of a burglary if John and Mary call?

Factor graphs

Each CPT is one factor (= black box)





D=lo	D=hi
0.6	0.4

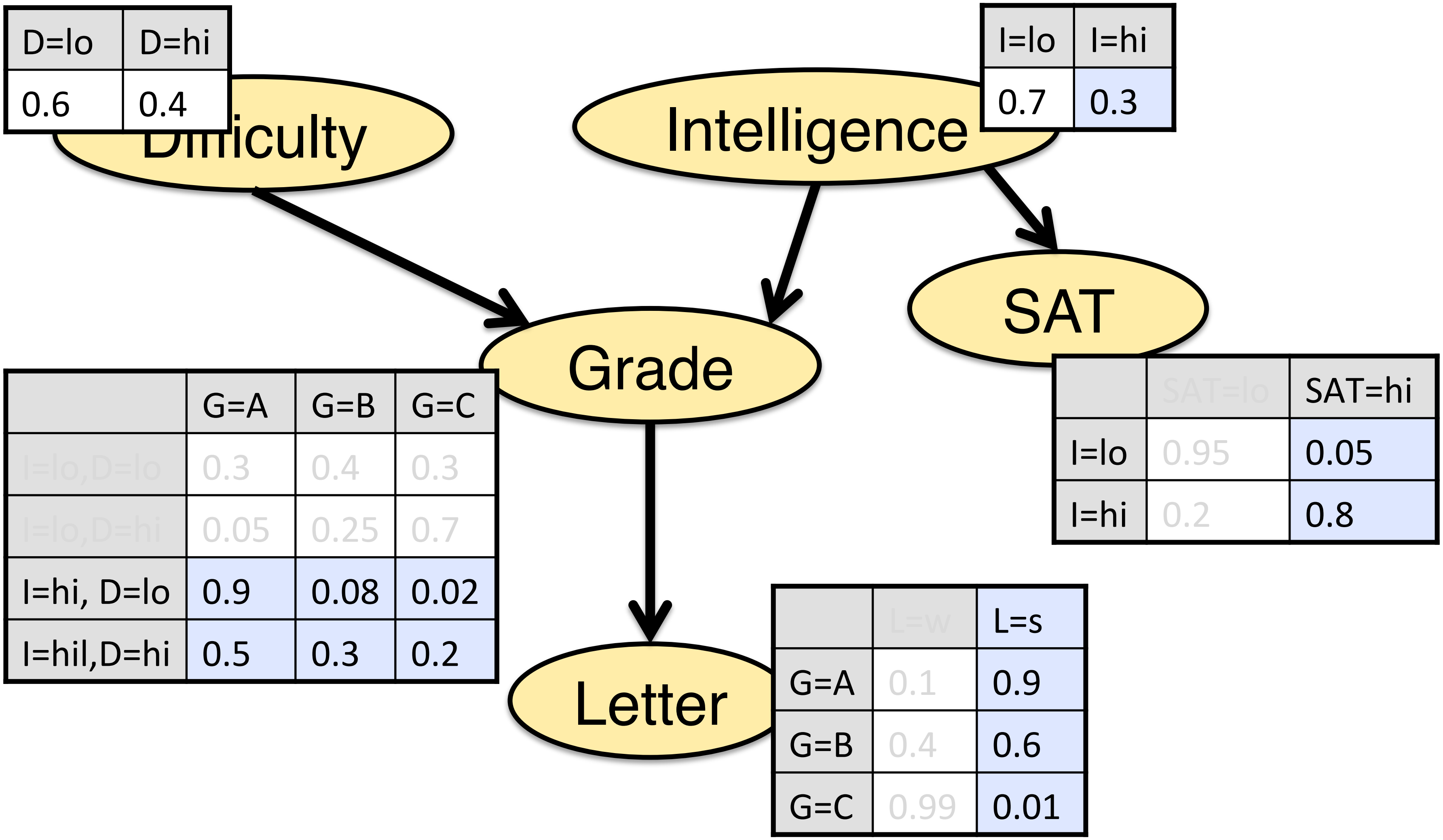
I=lo	I=hi
0.7	0.3

	G=A	G=B	G=C
I=lo,D=lo	0.3	0.4	0.3
I=lo,D=hi	0.05	0.25	0.7
I=hi, D=lo	0.9	0.08	0.02
I=hil,D=hi	0.5	0.3	0.2

	SAT=lo	SAT=hi
I=lo	0.95	0.05
I=hi	0.2	0.8

	L=w	L=s
G=A	0.1	0.9
G=B	0.4	0.6
G=C	0.99	0.01

What’s the probability that the student is intelligent if he got a good letter and a high SAT test?



D=lo	D=hi
0.6	0.4

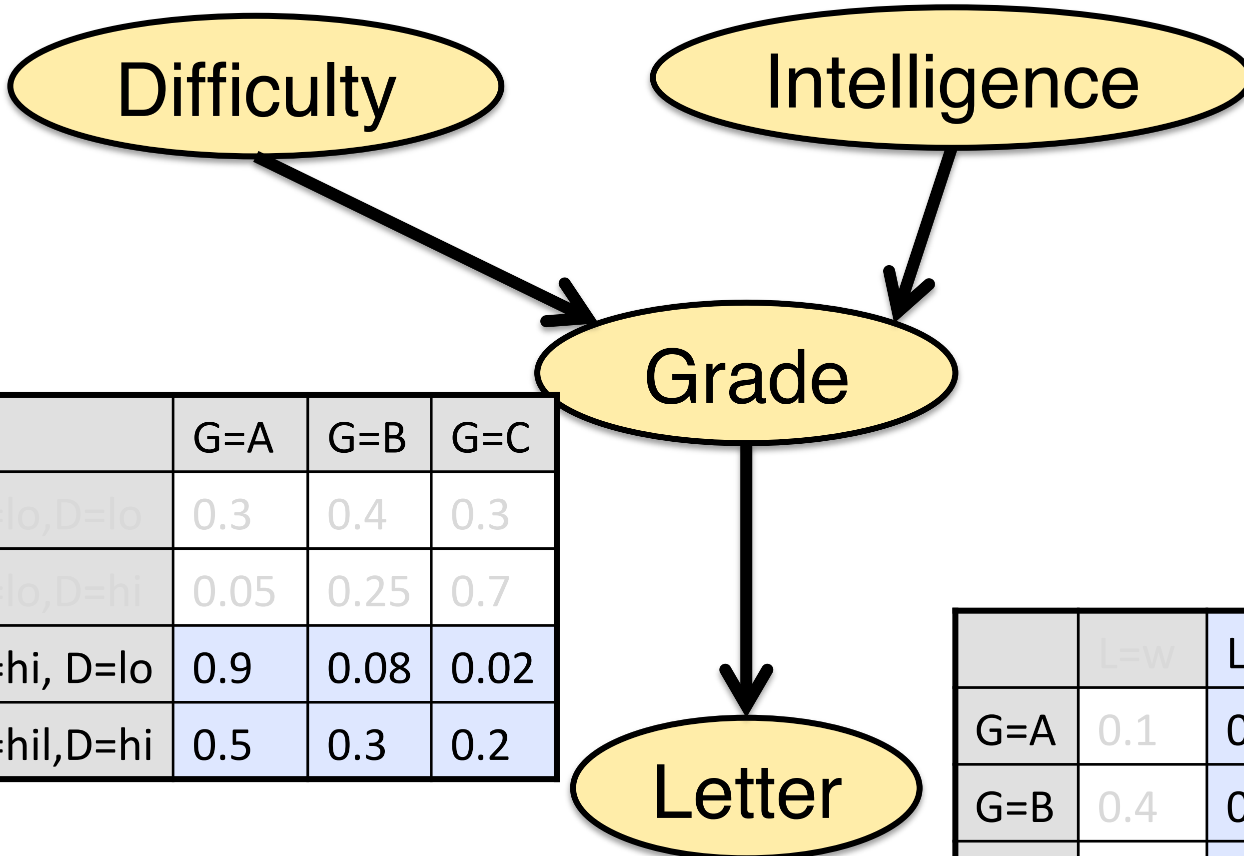
I=lo	I=hi
0.7	0.3

	G=A	G=B	G=C
I=lo,D=lo	0.3	0.4	0.3
I=lo,D=hi	0.05	0.25	0.7
I=hi, D=lo	0.9	0.08	0.02
I=hil,D=hi	0.5	0.3	0.2

	SAT=lo	SAT=hi
I=lo	0.95	0.05
I=hi	0.2	0.8

	L=w	L=s
G=A	0.1	0.9
G=B	0.4	0.6
G=C	0.99	0.01

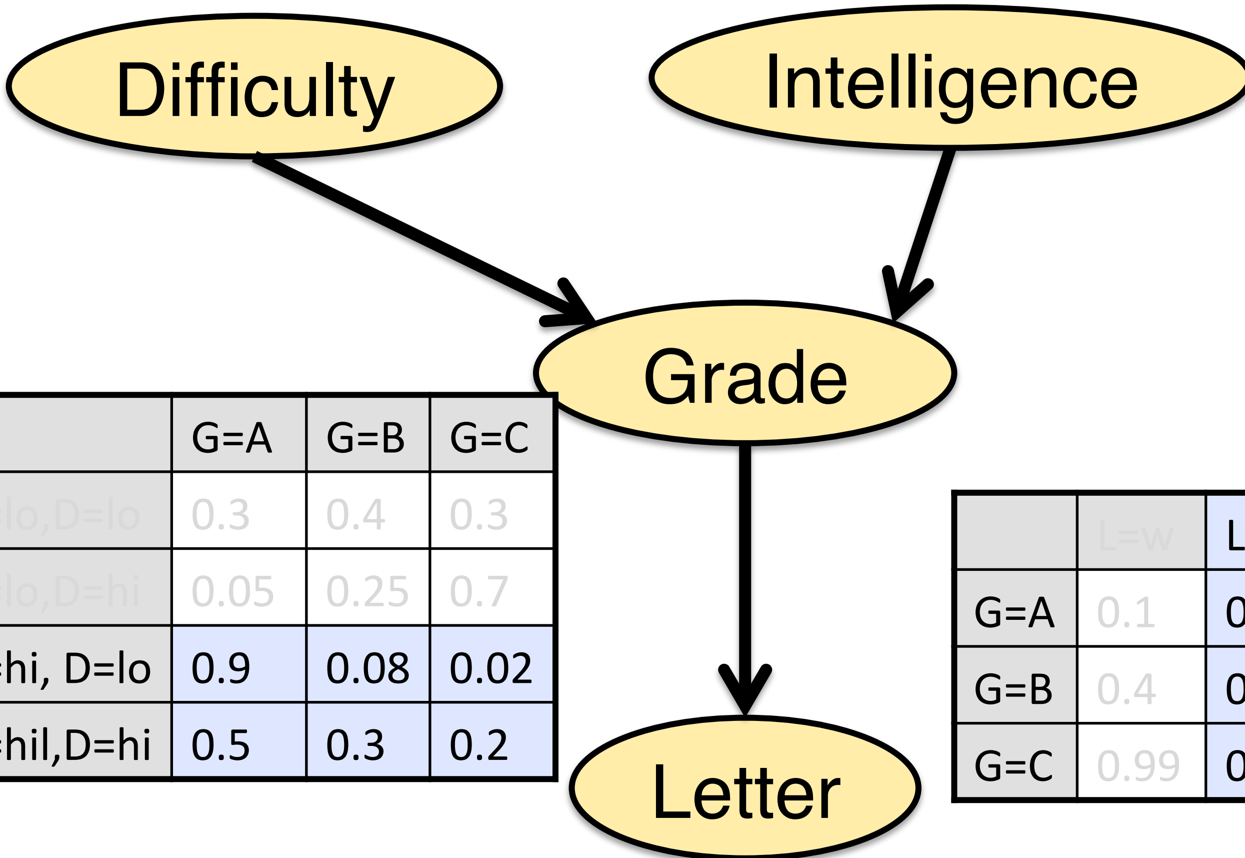
What's the probability that the student is intelligent if he got a good letter and a high SAT test?



	G=A	G=B	G=C
I=lo,D=lo	0.3	0.4	0.3
I=lo,D=hi	0.05	0.25	0.7
I=hi, D=lo	0.9	0.08	0.02
I=hil,D=hi	0.5	0.3	0.2

	L=w	L=s
G=A	0.1	0.9
G=B	0.4	0.6
G=C	0.99	0.01

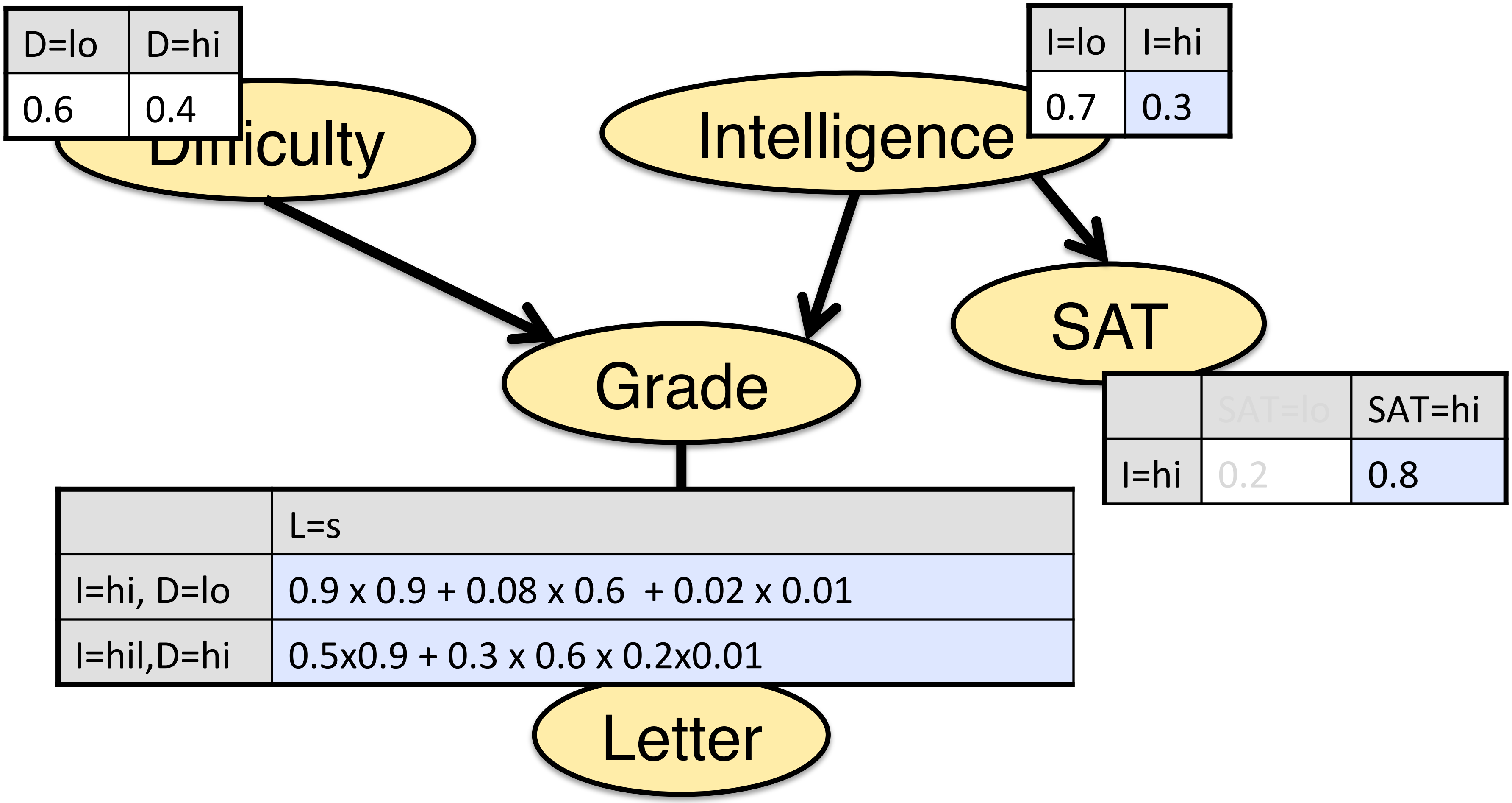
What's the probability that the student is intelligent if he got a good letter and a high SAT test?



	G=A	G=B	G=C
I=lo,D=lo	0.3	0.4	0.3
I=lo,D=hi	0.05	0.25	0.7
I=hi, D=lo	0.9	0.08	0.02
I=hil,D=hi	0.5	0.3	0.2

	L=w	L=s
G=A	0.1	0.9
G=B	0.4	0.6
G=C	0.99	0.01

	L=s
I=hi, D=lo	0.9 x 0.9 + 0.08 x 0.6 + 0.02 x 0.01
I=hil,D=hi	0.5x0.9 + 0.3 x 0.6 x 0.2x0.01



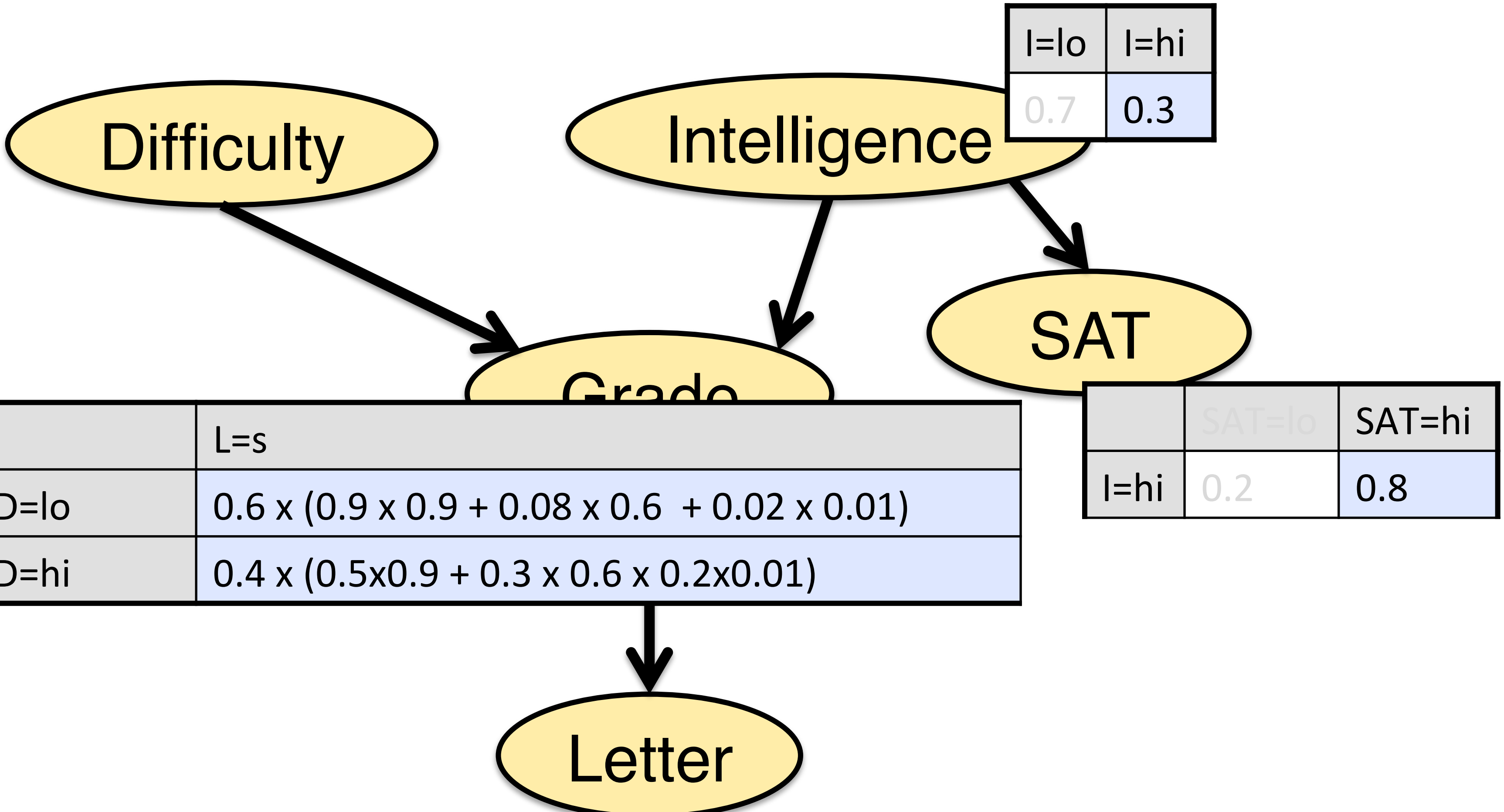
D=lo	D=hi
0.6	0.4

I=lo	I=hi
0.7	0.3

	SAT=lo	SAT=hi
I=hi	0.2	0.8

	L=s
I=hi, D=lo	0.9 x 0.9 + 0.08 x 0.6 + 0.02 x 0.01
I=hil,D=hi	0.5x0.9 + 0.3 x 0.6 x 0.2x0.01

What's the probability that the student is intelligent if he got a good letter and a high SAT test?



I=lo	I=hi
0.7	0.3

	SAT=lo	SAT=hi
I=hi	0.2	0.8

	L=s
D=lo	0.6 x (0.9 x 0.9 + 0.08 x 0.6 + 0.02 x 0.01)
D=hi	0.4 x (0.5x0.9 + 0.3 x 0.6 x 0.2x0.01)

What's the probability that the student is intelligent if he got a good letter and a high SAT test?

0.3 x 0.8 x [0.6x(0.9 x 0.9 + 0.08 x 0.6 + 0.02 x 0.01) + 0.4x(0.5x0.9 + 0.3 x 0.6 x 0.2x0.01)]

Complexity of variable elimination

Polytree networks:
only one (undirected) path between each pair of nodes.

Complexity is linear in the size of the network (number of CPT entries)

Complexity of variable elimination

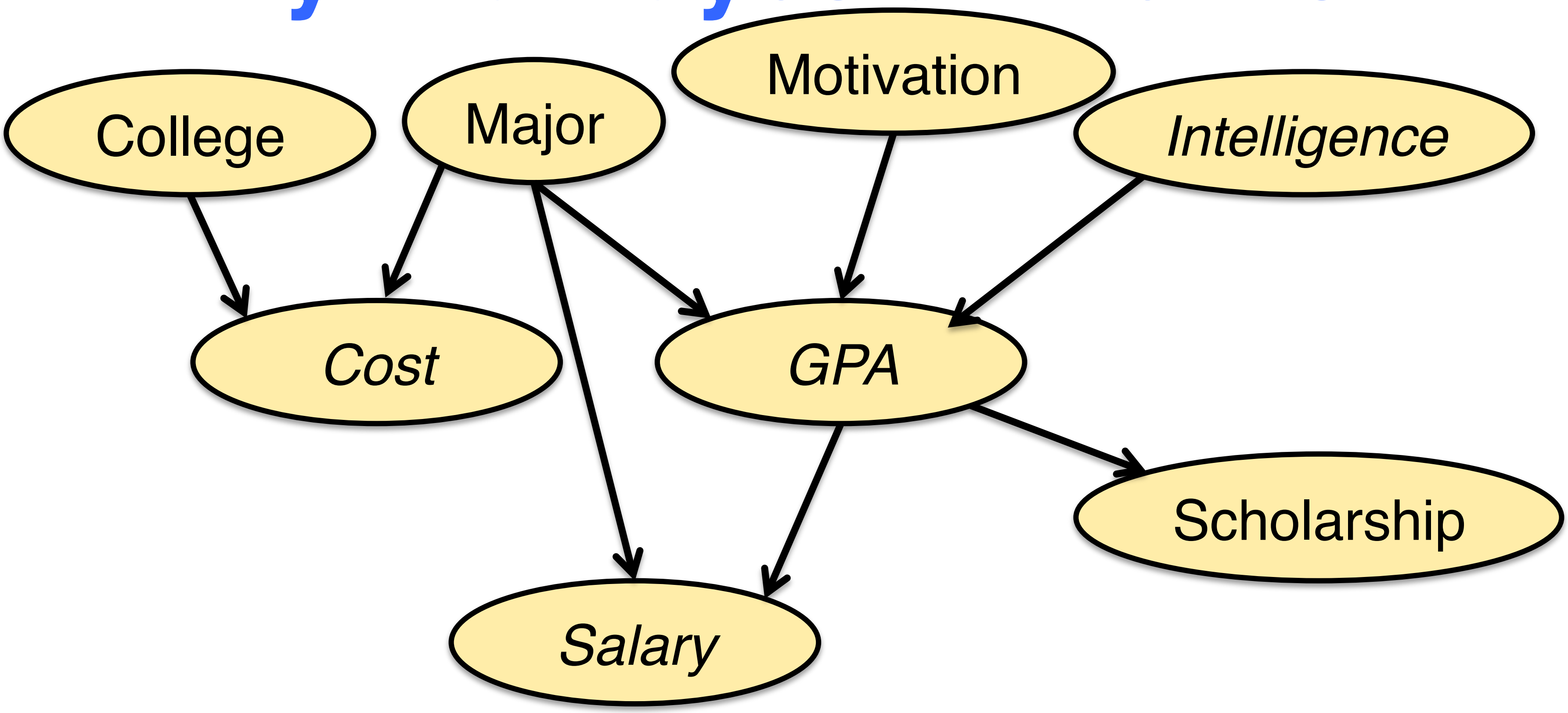
Multiply connected networks:
multiple (undirected) paths between some pair of nodes.

Complexity is **exponential** in the size of the network (number of CPT entries)

- (Similar to finding number of models in propositional logic)

Continuous variables

A hybrid Bayesian Network



The cost of your degree (tuition and living) depends on your choice of college and major. Your GPA depends on your motivation (high/normal/low), your IQ and your major.
Whether you get a scholarship depends on your GPA.
Your starting salary after you graduate depends on your major and your GPA.

Continuous variables

Continuous R.V.s (without parents)

e.g. *Temperature* (may not depend on anything)

Continuous R.V.s with discrete parents

e.g. *Tip* might depend on *CustomerIsHappy*

Continuous R.V.s with continuous parents

e.g. *SATscore* might depend on *IQ**

**(it's easier to treat integers like reals)*

Discrete R.V.s with continuous parents

e.g. *CustomerIsHappy* might depend on *Temperature*

R.V.s with mixed (discrete&continuous) parents

Continuous random variables

Many quantities of interest are numerical:

- intelligence: IQ score
- SAT: SAT score
- Temperature,…

We can quantize them:

- intelligence: low = IQ < 85; avg: 85 < IQ < 115

Or we can deal with them directly

- if we are talking about a scale, it's often easier to treat them as real numbers

Continuous random variables

If X is a real-valued random variable, what is the probability that *X=1.2, 1.21, 1.211, …?*

We can't assign probabilities to individual outcomes anymore.

But we can assign probabilities to the event that X falls into a specific interval:
What is the probability that $1.20 < X < 1.21$?

Probability density functions

A **probability density function** $P(X)$ **(PDF)** for a real-valued random variable X is a function such that

– $P(x) \geq 0$ for all $-\infty \leq x \leq +\infty$ (P is non-negative)

– $\int P(x)dx = 1$

– For any interval A: $P(x \in A) = \int_A P(x)dx$

The **cumulative distribution function** $F(n)$ is the probability that $X \leq n$: $F(n) = P(X \leq n) = \int^n P(x)dx$

Expectation (Mean)

The **expectation** $E(X)$ of a real-valued random variable X with PDF $P(X)$ is defined as

$$E(X) = \int xP(x)dx$$

$E(X)$ is also called the **mean** μ

The **expectation** $E(f(X))$ of a function of X with PDF $P(X)$ is defined as:

$$E(X) = \int f(x)P(x)dx$$

Variance

The **variance** $Var(X) = \sigma^2$ is defined as

$$Var(X) = E[(X - \mu)^2]$$

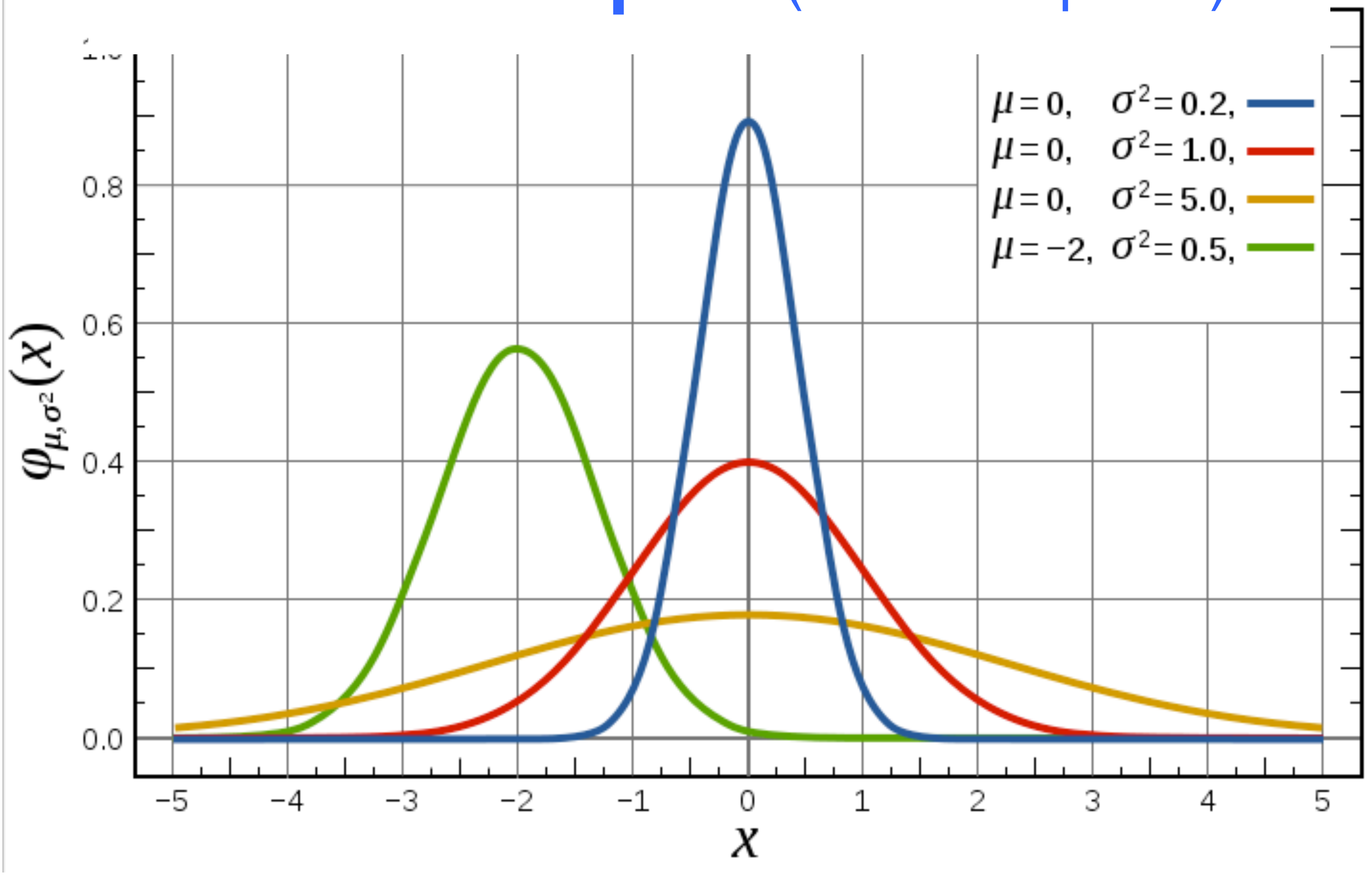
The **standard deviation** σ is the square root of the variance.

The normal (Gaussian) distribution

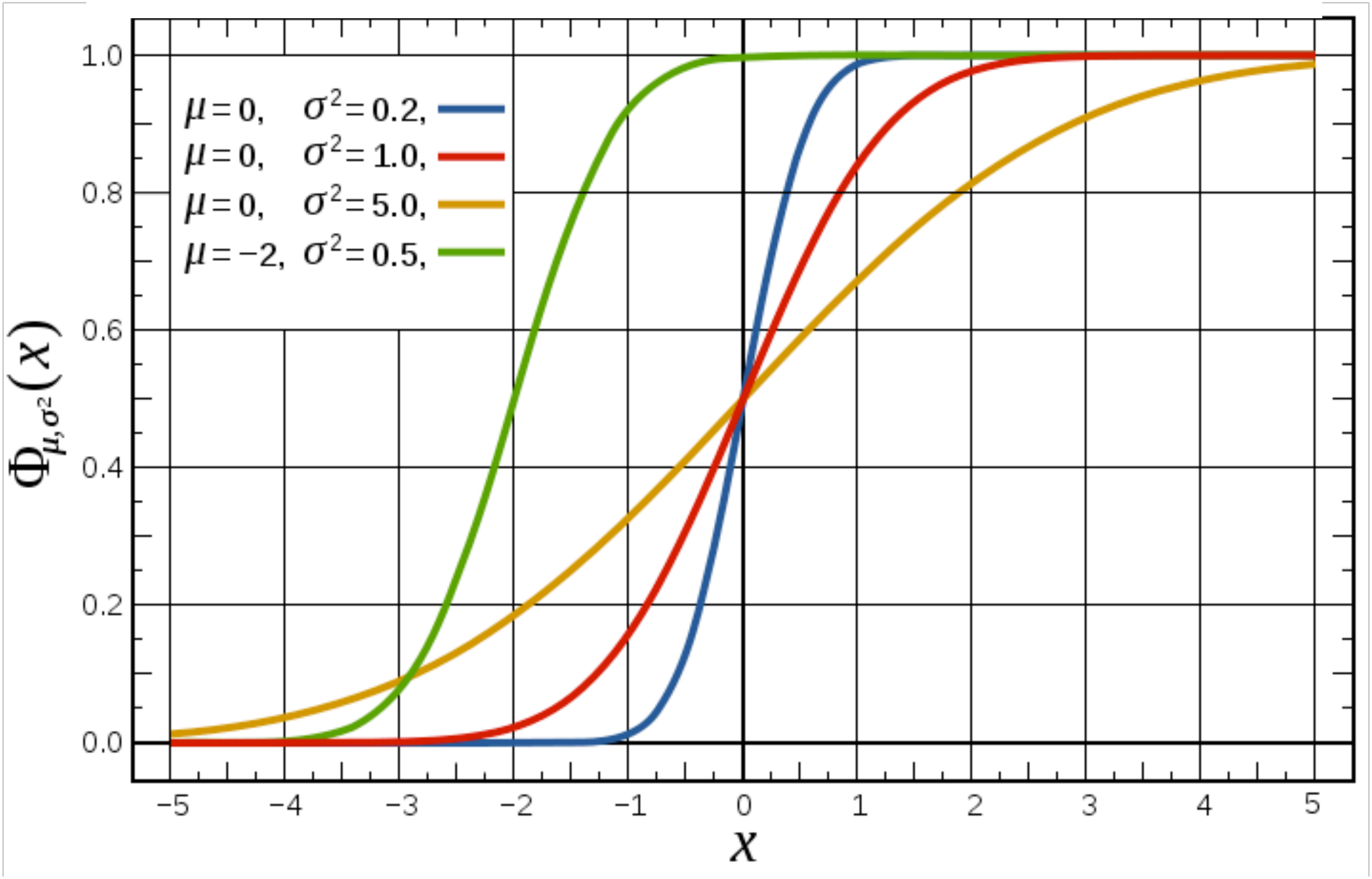
Gaussian (Normal) distribution with mean μ and variance σ^2

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{(2\sigma^2)}}$$

A few examples (from Wikipedia)

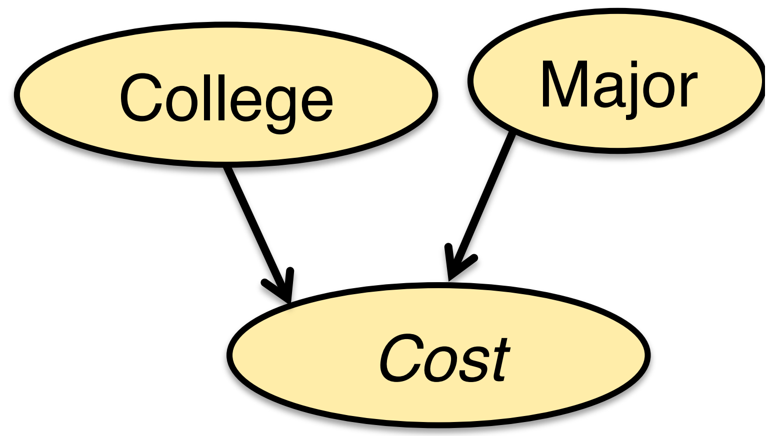


The CDFs (also from Wikipedia)



Continuous R.V. with discrete parents: a set of Gaussians

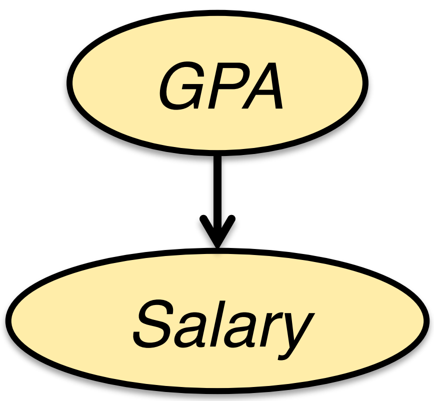
Each possible outcome of the set of discrete parents = one distribution



College	Major	Cost
U of I	CS	$N(\mu_{UofI,CS}, \sigma^2_{UofICS})$
U of I	English	$N(\mu_{UofI,E}, \sigma^2_{UofIE})$
UChicago	CS	$N(\mu_{UC,CS}, \sigma^2_{UC,CS})$

Continuous R.V.s with continuous parents: linear Gaussians

Salary is a Gaussian whose mean depends (linearly) on *GPA:* $\mu_{Salary} = a \times GPA + b$
We assume a fixed variance σ^2 .
$P(Salary \mid GPA) = N(a \times GPA + b, \sigma^2)$



$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{(2\sigma^2)}}$$

Discrete (binary) R.V.s with continuous parents

The probability of getting a scholarship depends on your GPA:

- If your GPA is very low, it is close to zero.
- If your GPA is very high, it is close to one (you'd hope).
- Inbetween, the probability varies smoothly.

We need a threshold function

- Probit: CDF of Standard Normal (Gaussian with mean 0 and variance 1)
- Logit: uses logistic function $1/(1+e^{-x})$