15-381
ARTIFICIAL INTELLIGENCE
LECTURE 8: BAYESIAN NETWORKS

Fall 2010

REVIEW

REVIEW

Marginal Probability

Joint Probability

SUM RULE

$$p(X) = \sum_{Y} p(X, Y)$$

REVIEW

Marginal Probability

Joint Probability

SUM RULE $$p(X) = \sum_Y p(X, Y)$$

PRODUCT RULE $$p(X, Y) = p(Y|X)p(X)$$

Conditional Probability

REVIEW

Marginal Probability

Joint Probability

SUM RULE

$$p(X) = \sum_{Y} p(X, Y)$$

PRODUCT RULE

$$p(X, Y) = p(Y|X)p(X)$$

Conditional Probability

Likelihood

Prior Probability

BAYES' THEOREM

$$p(X|Y) = \frac{p(Y|X)p(X)}{p(Y)}$$

Posterior Probability

Evidence

BURGLARY

BURGLARY

BURGLARY

BURGLARY

BURGLARY

JOHN

BURGLARY

JOHN

MARY

BURGLARY

EARTHQUAKE

JOHN

MARY

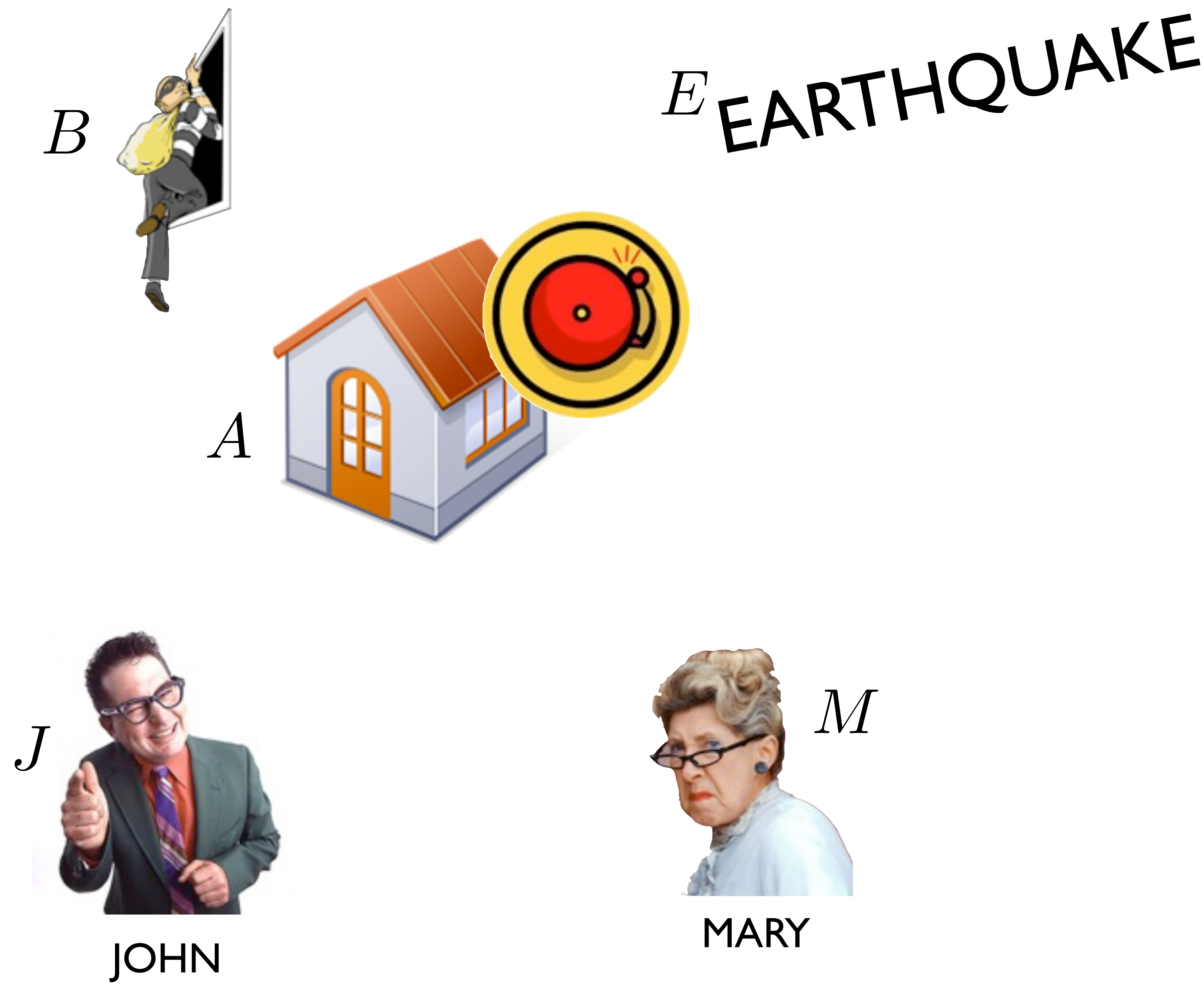
BURGLARY

EARTHQUAKE

JOHN

MARY

BURGLARY



BURGLARY

B

E EARTHQUAKE

A

B=TRUE if house is burgled
E=TRUE is there is an earthquake
A=TRUE if the alarm rings
J=TRUE if John Calls
M=TRUE if Mary Calls

J

JOHN

M

MARY

WILL I GET A CALL IF THERE IS A BURGLARY?

WILL I GET A CALL IF THERE IS AN EARTHQUAKE AND NO BURGLARY?

WILL I GET A CALL IF THERE IS NO EARTHQUAKE AND NO BURGLARY?

WILL I GET A CALL FROM MARY IF THERE IS A BURGLARY?

WILL THE ALARM GO OFF IF THERE IS AN EARTHQUAKE AND NO ALARM?

.

.

.

WILL I GET A CALL IF THERE IS A BURGLARY?

WILL I GET A CALL IF THERE IS AN EARTHQUAKE AND NO BURGLARY?

WILL I GET A CALL IF THERE IS NO EARTHQUAKE AND NO BURGLARY?

WILL I GET A CALL FROM MARY IF THERE IS A BURGLARY?

WILL THE ALARM GO OFF IF THERE IS AN EARTHQUAKE AND NO ALARM?

.

.

.

$$p(B, E, A, J, M)$$

Joint Probability

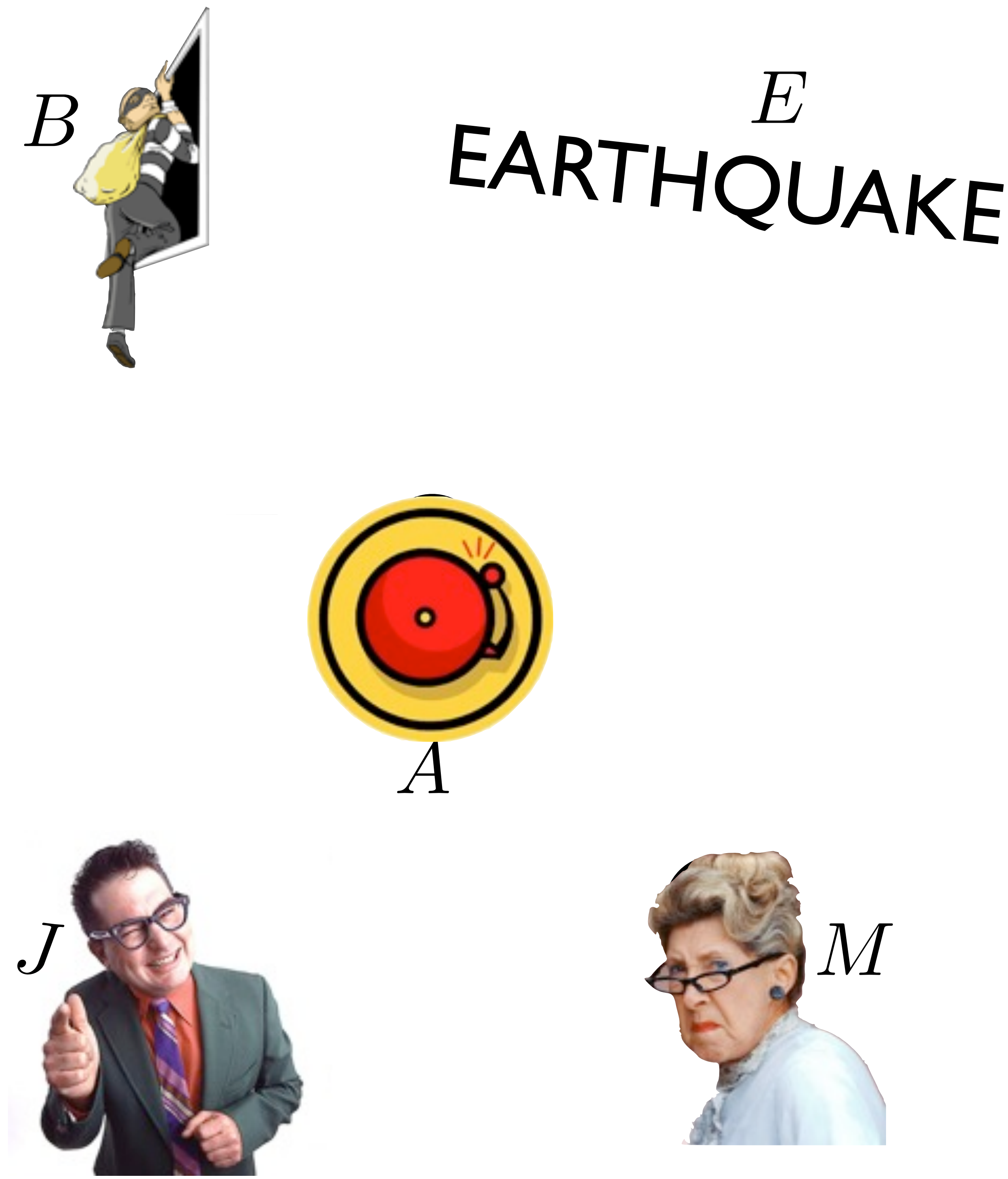
PARAMETERS

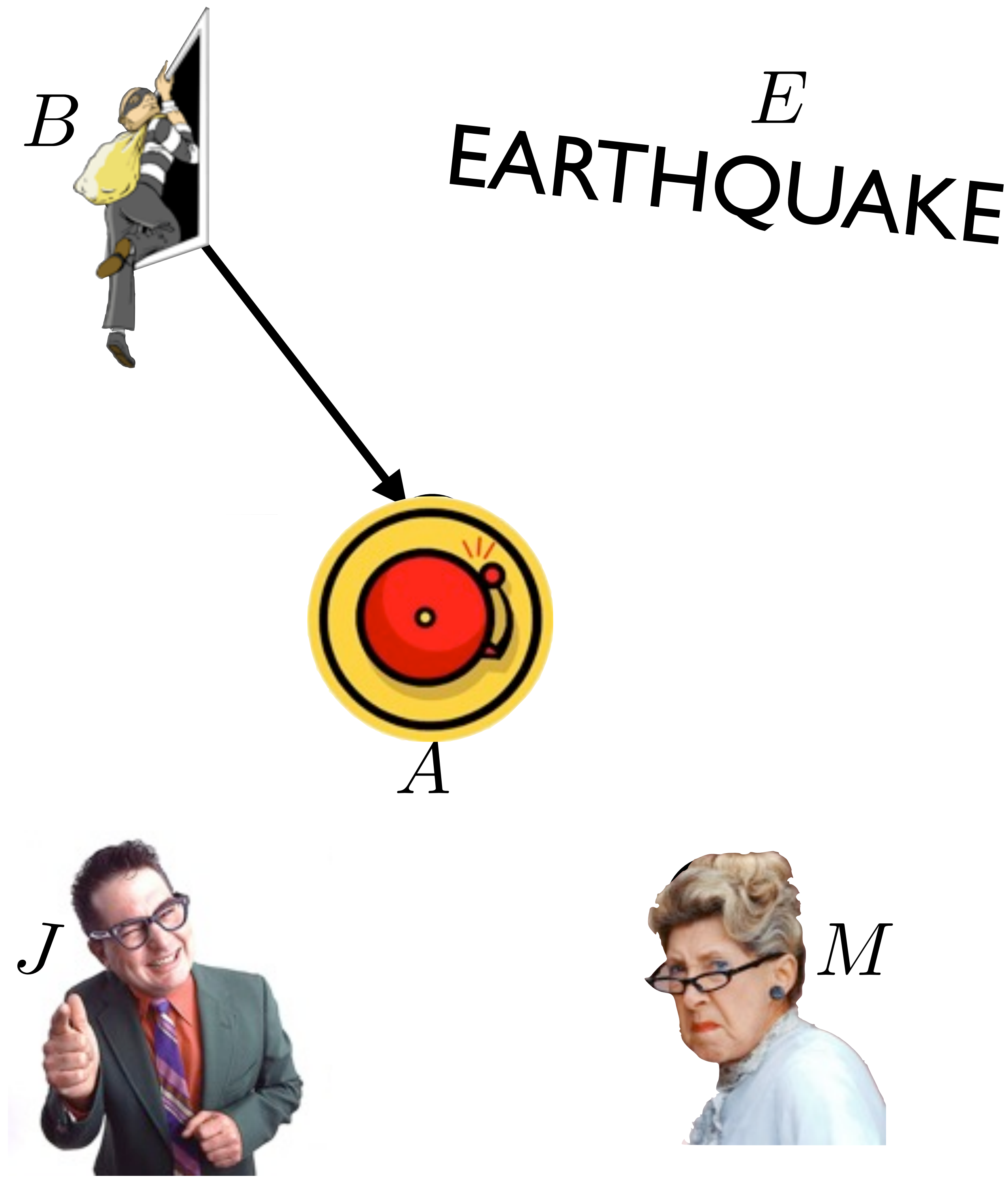
PRODUCT RULE

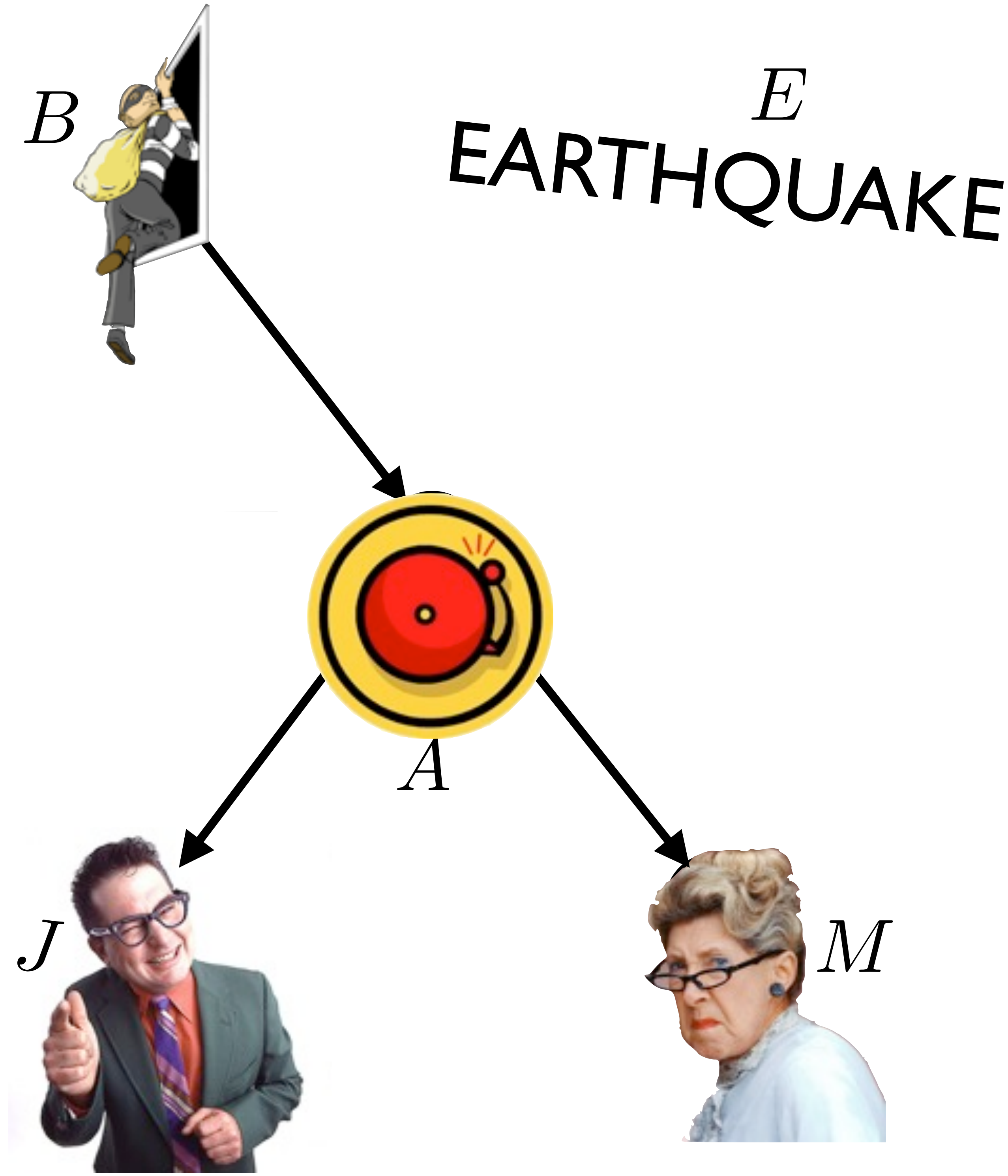
$$p(B,E,A,J,M) = p(B|E,A,J,M)p(E,A,J,M)$$
$$= p(B|E,A,J,M)p(E|A,J,M)p(A,J,M)$$

$$p(B,E,A,J,M) = \underbrace{p(B|E,A,J,M)}_{2^4=16}\underbrace{p(E|A,J,M)}_{8}\underbrace{p(A|J,M)}_{4}\underbrace{p(J|M)}_{2}\underbrace{p(M)}_{1}$$

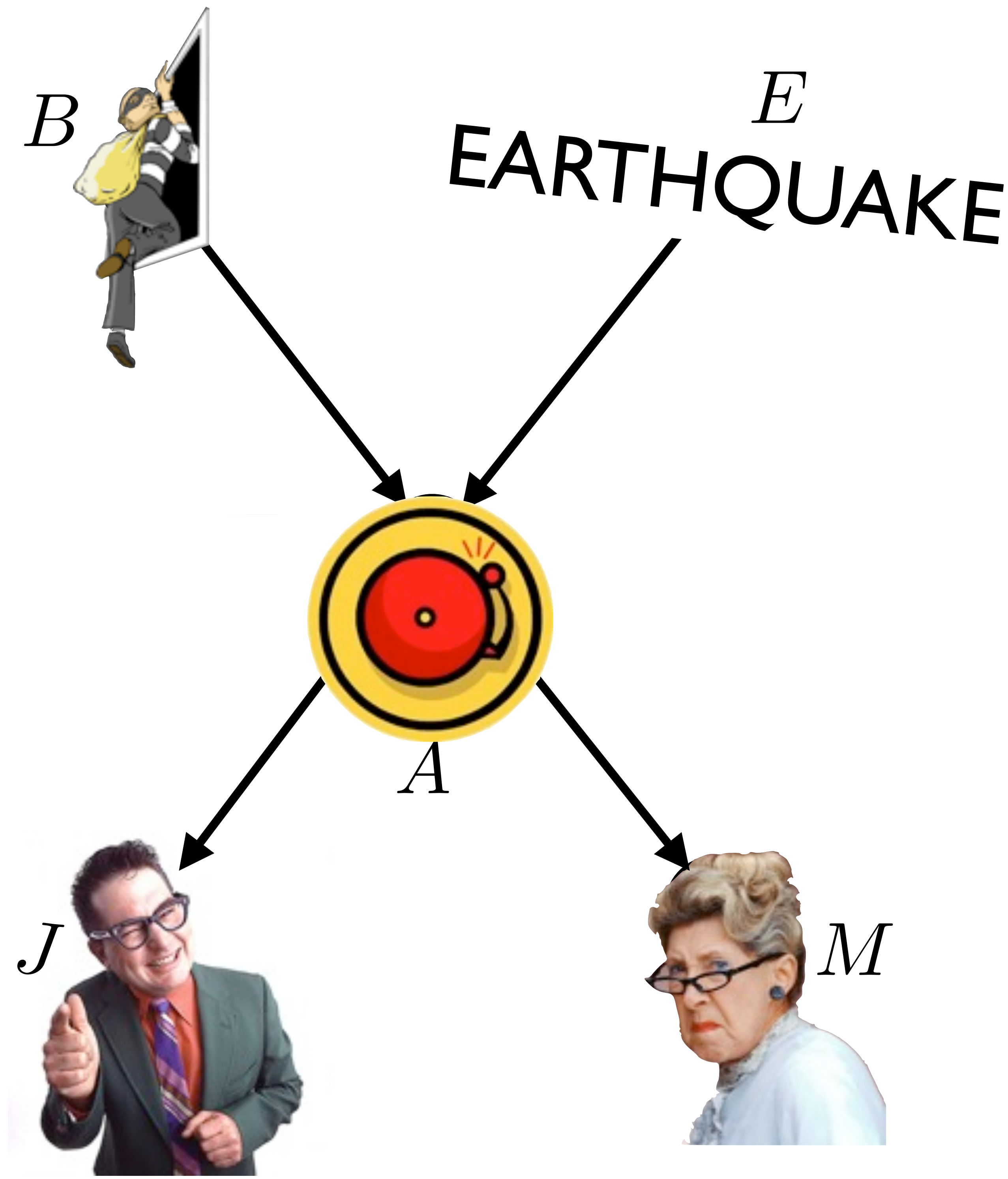
TOTAL # OF PARAMETERS: 31


B
E
EARTHQUAKE
A
J
M

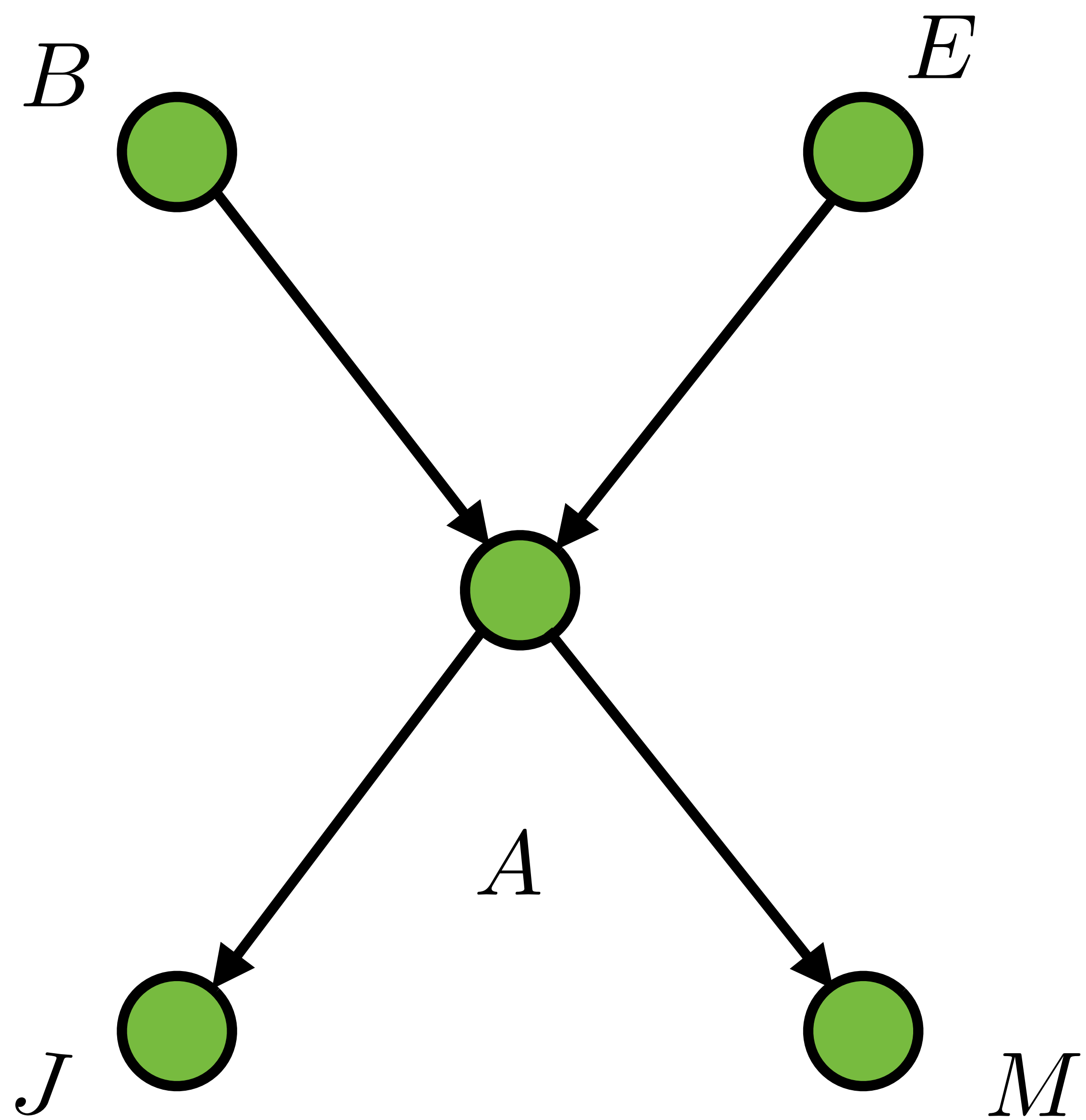

B
E
EARTHQUAKE
A
J
M



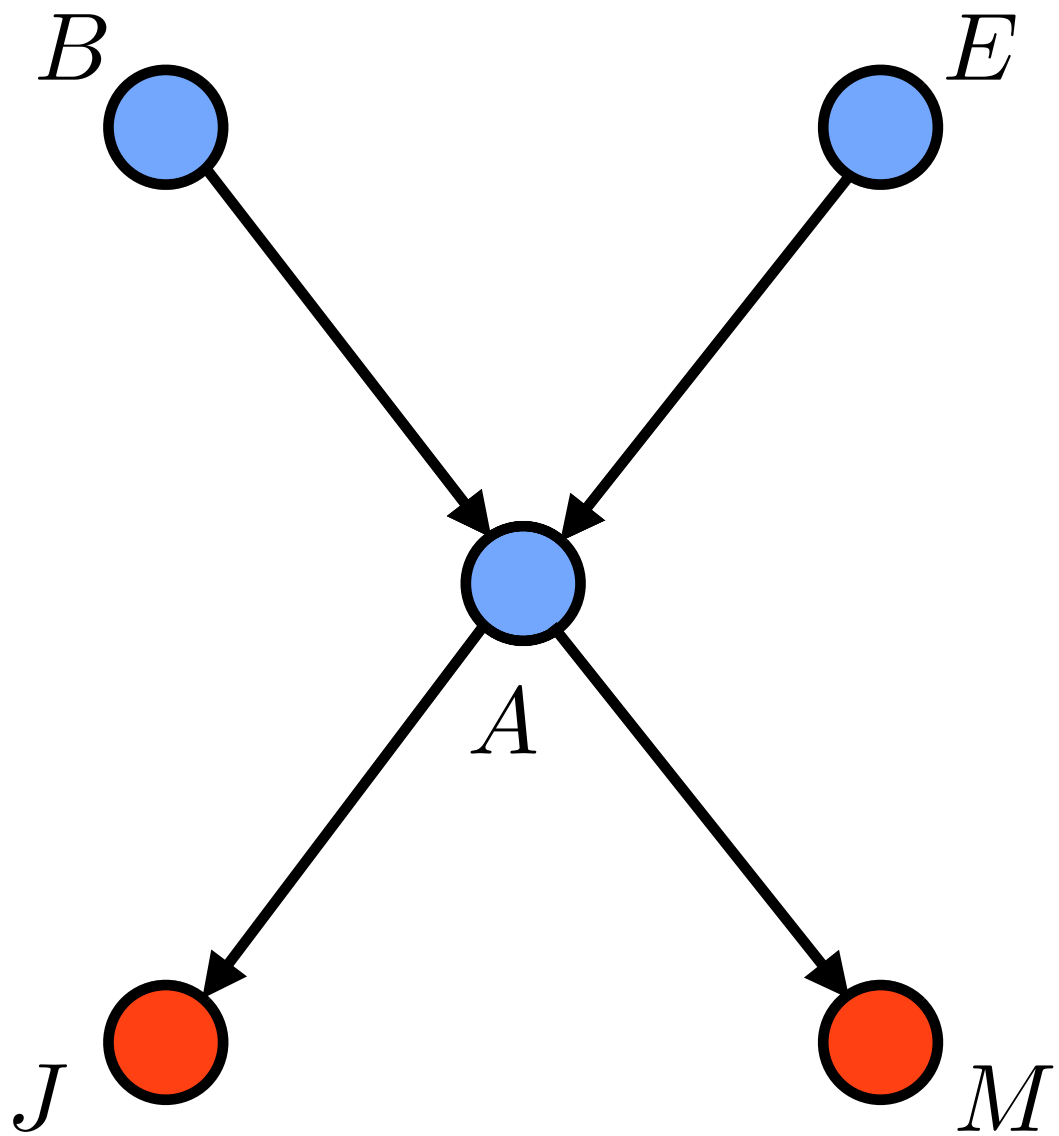
B
E
EARTHQUAKE
A
J
M



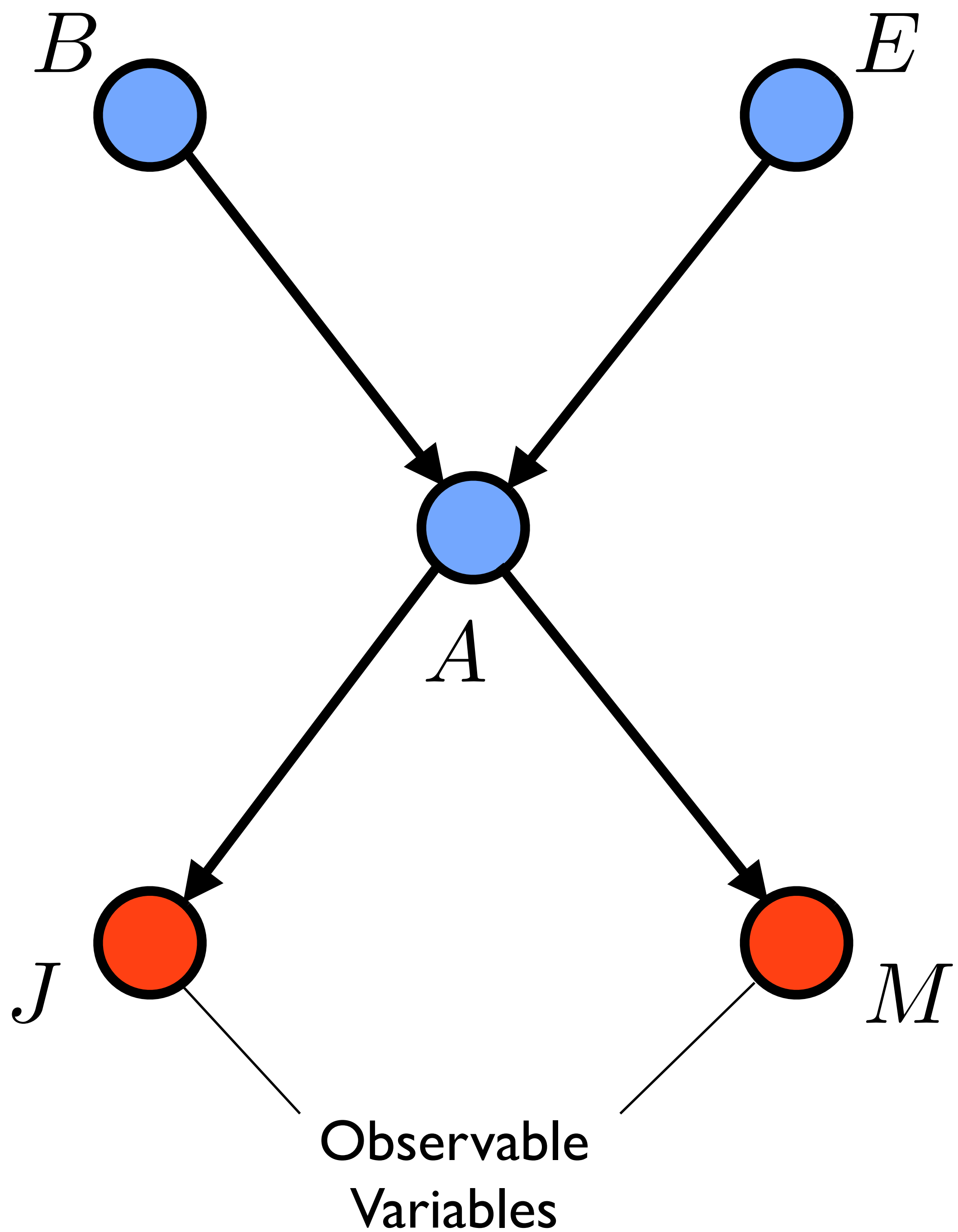
B
E
EARTHQUAKE
A
J
M


B
E
A
J
M

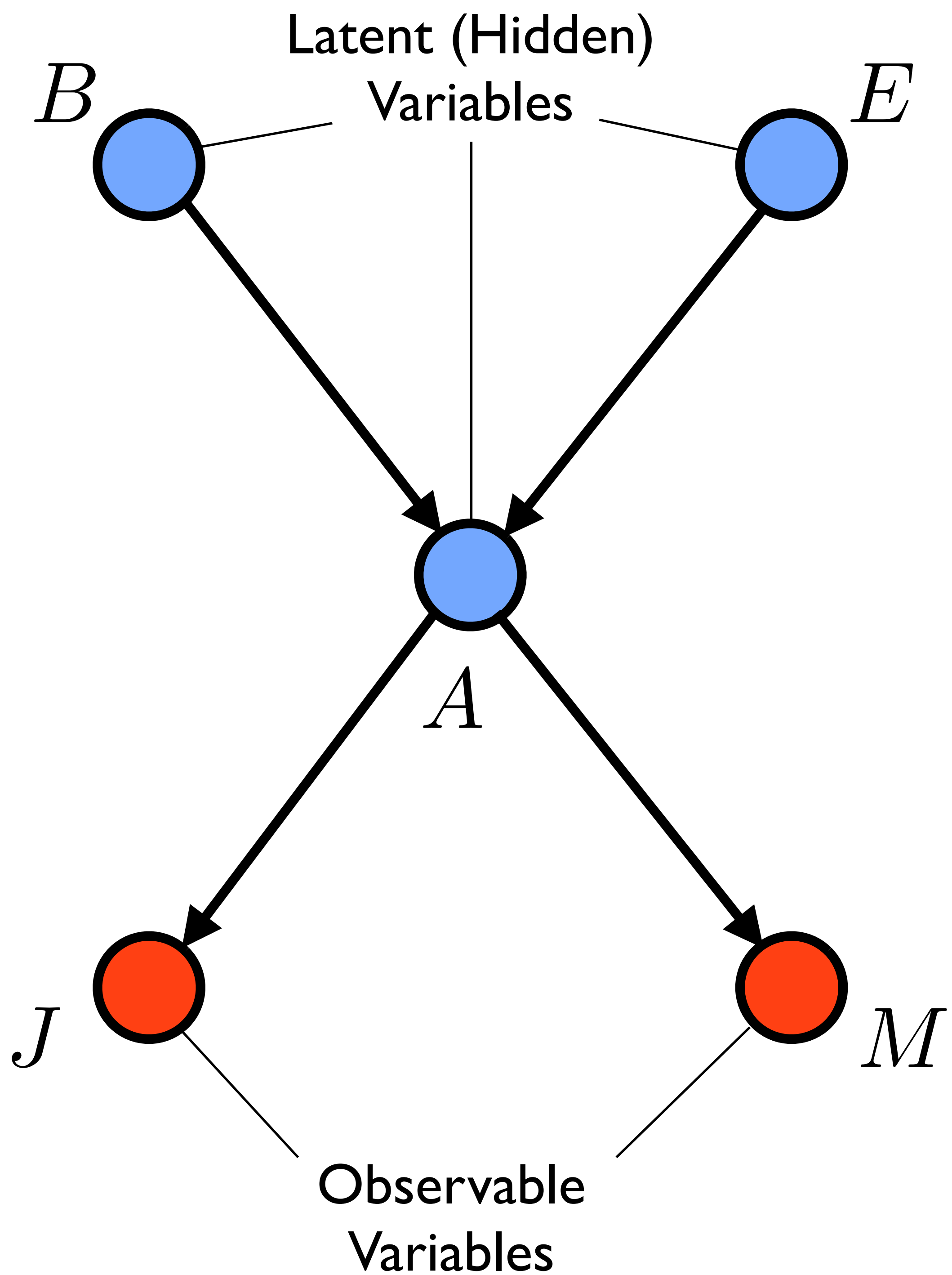
OBSERVABLE AND LATENT VARIABLES



OBSERVABLE AND LATENT VARIABLES



OBSERVABLE AND LATENT VARIABLES



PROBABILITIES

B

E EARTHQUAKE

PROBABILITIES

B

E EARTHQUAKE

$$p(B = \texttt{TRUE}) = 0.001$$
$$p(B = \texttt{FALSE}) = 0.999$$

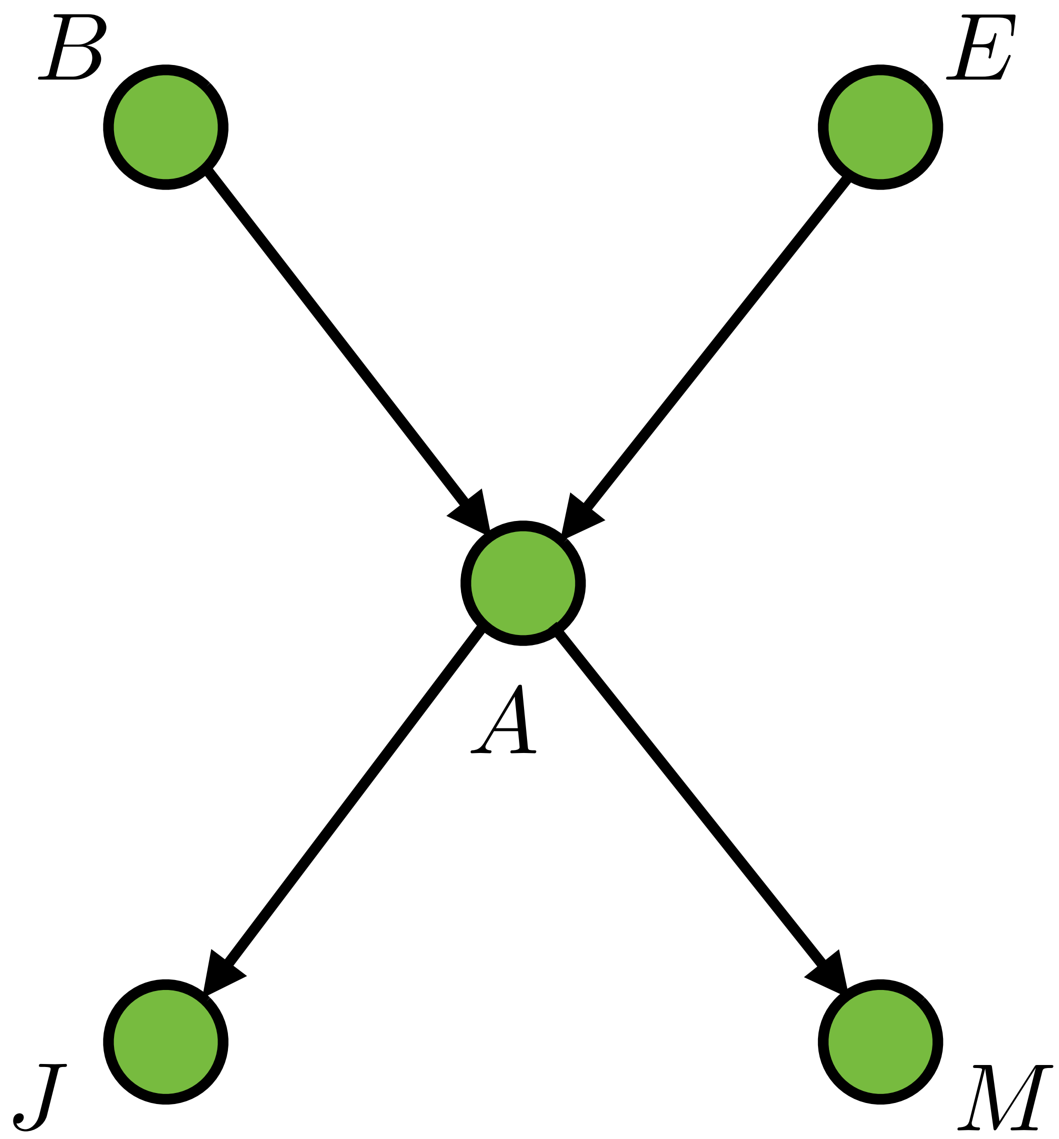
PROBABILITIES

B

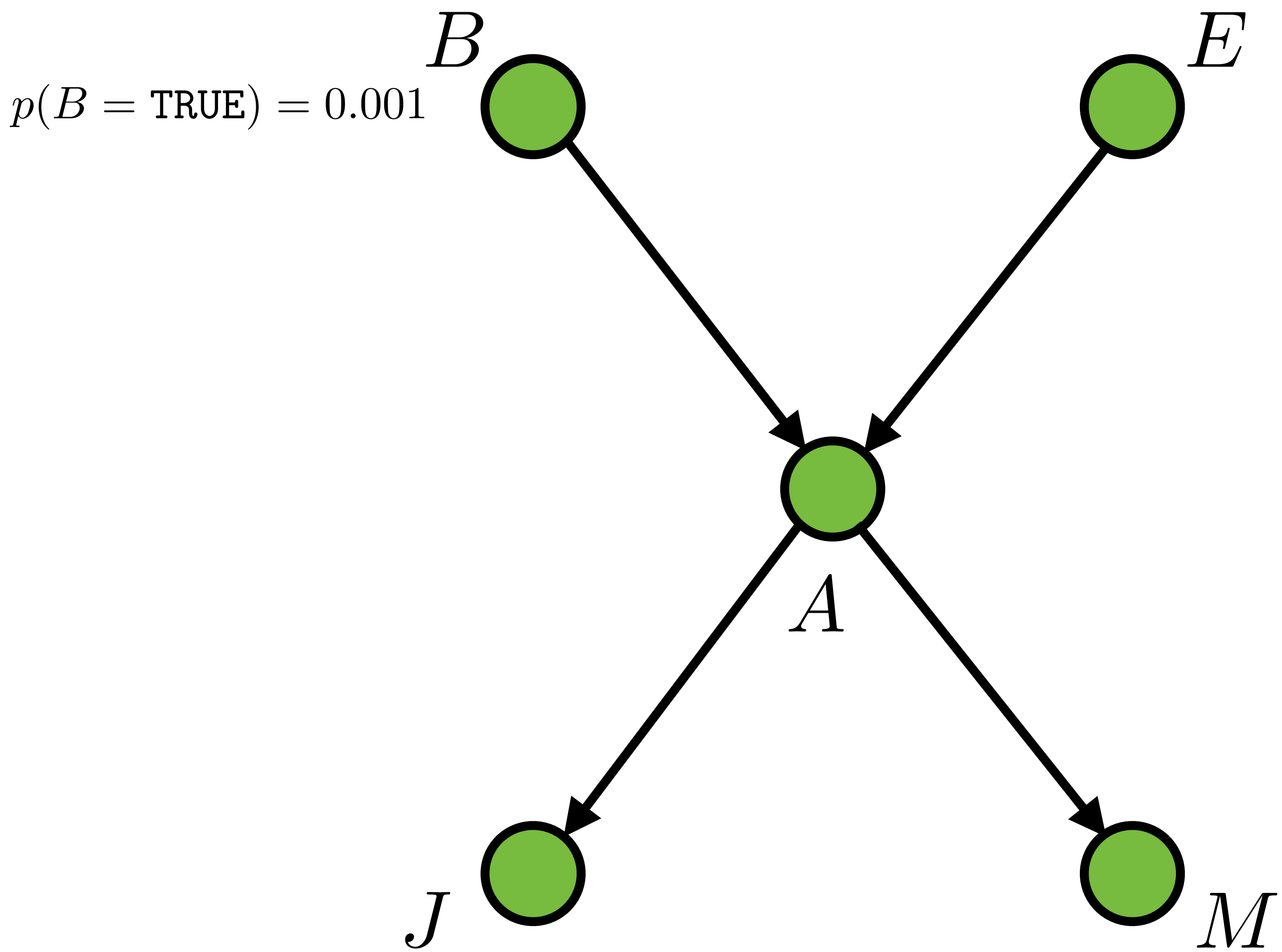
E EARTHQUAKE

$$p(B = \texttt{TRUE}) = 0.001$$
$$p(B = \texttt{FALSE}) = 0.999$$

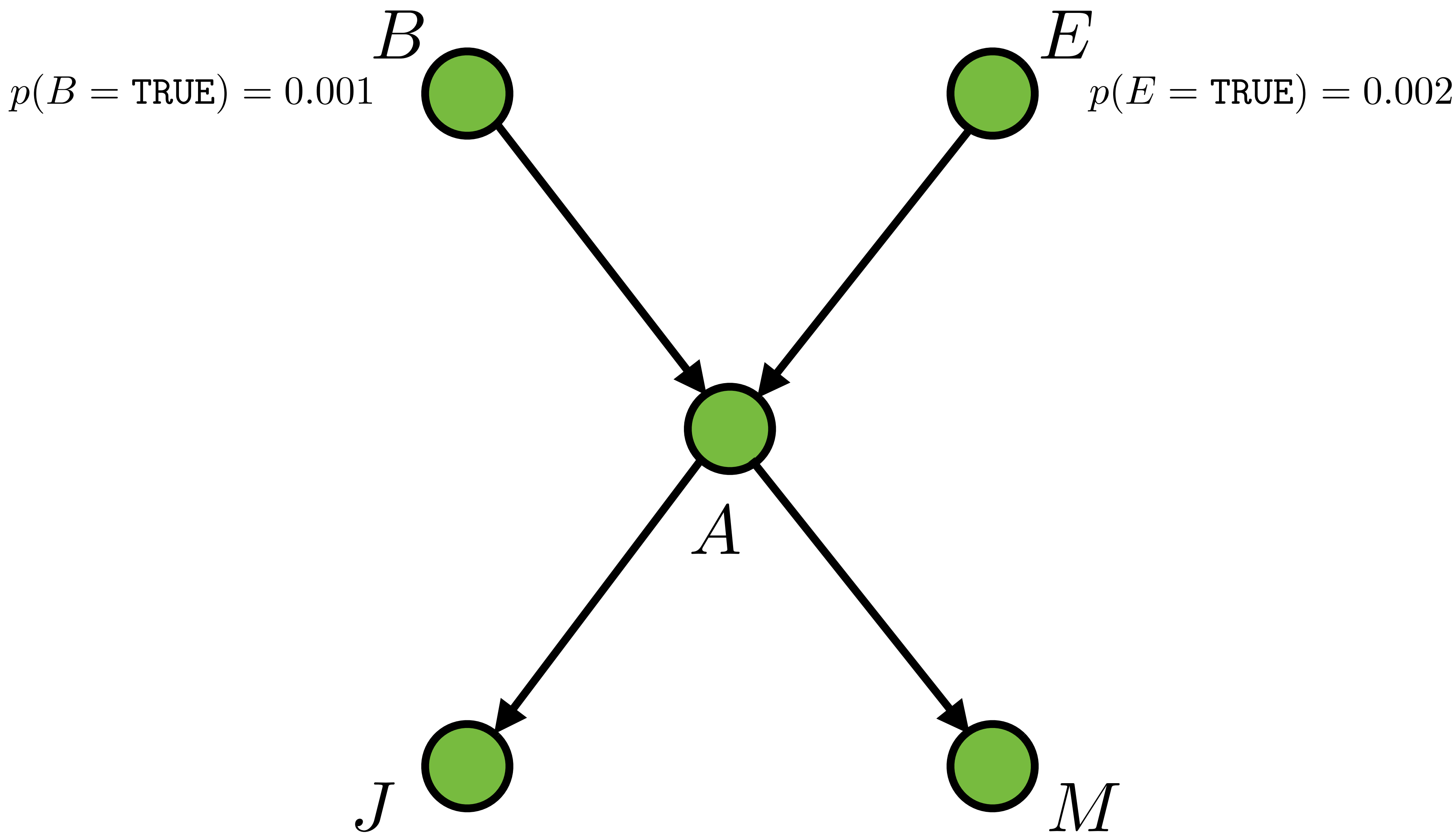
$$p(E = \texttt{TRUE}) = 0.002$$
$$p(E = \texttt{FALSE}) = 0.998$$



B
E
A
J
M



B
$p(B = \texttt{TRUE}) = 0.001$
E
A
J
M


B
$p(B = \texttt{TRUE}) = 0.001$
E
$p(E = \texttt{TRUE}) = 0.002$
A
J
M

CONDITIONAL PROBABILITY TABLE

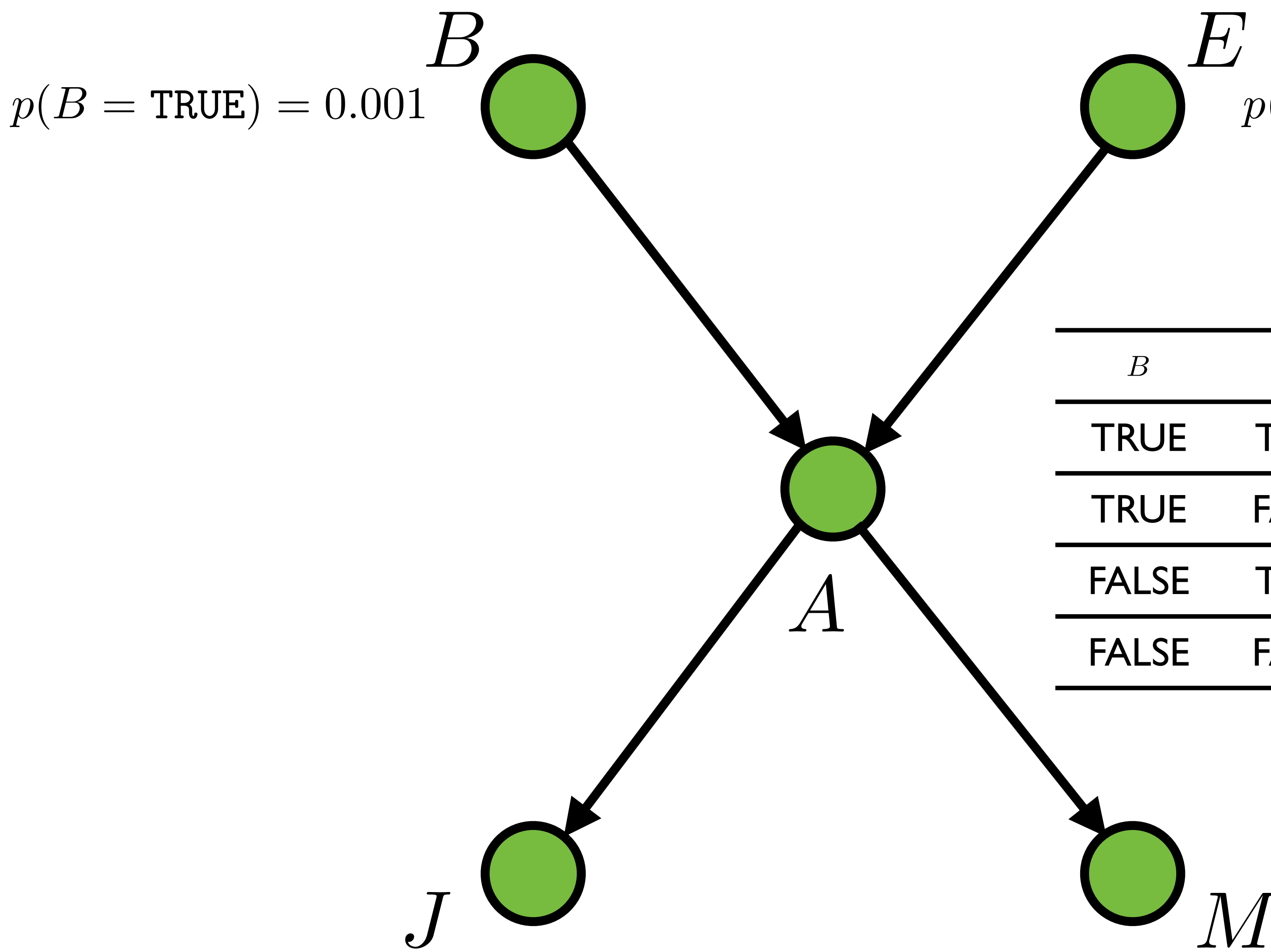
		$p(A\|B,E)$	
B	E	TRUE	FALSE
TRUE	TRUE	0.95	0.05
TRUE	FALSE	0.95	0.05
FALSE	TRUE	0.29	0.71
FALSE	FALSE	0.001	0.999

CONDITIONAL PROBABILITY TABLE

		$p(A\|B, E)$	
B	E	TRUE	FALSE
TRUE	TRUE	0.95	0.05
TRUE	FALSE	0.95	0.05
FALSE	TRUE	0.29	0.71
FALSE	FALSE	0.001	0.999

CONDITIONAL PROBABILITY TABLE

B	E	$p(A\|B,E)$
TRUE	TRUE	0.95
TRUE	FALSE	0.95
FALSE	TRUE	0.29
FALSE	FALSE	0.001



B	E	$p(A\|B, E)$
TRUE	TRUE	0.95
TRUE	FALSE	0.95
FALSE	TRUE	0.29
FALSE	FALSE	0.001

PROBABILITIES

J

JOHN

M

MARY

PROBABILITIES

J

JOHN

M

MARY

A	$p(J \mid A)$
TRUE	0.90
FALSE	0.05

PROBABILITIES

J

JOHN

M

MARY

A	$p(J\|A)$
TRUE	0.90
FALSE	0.05

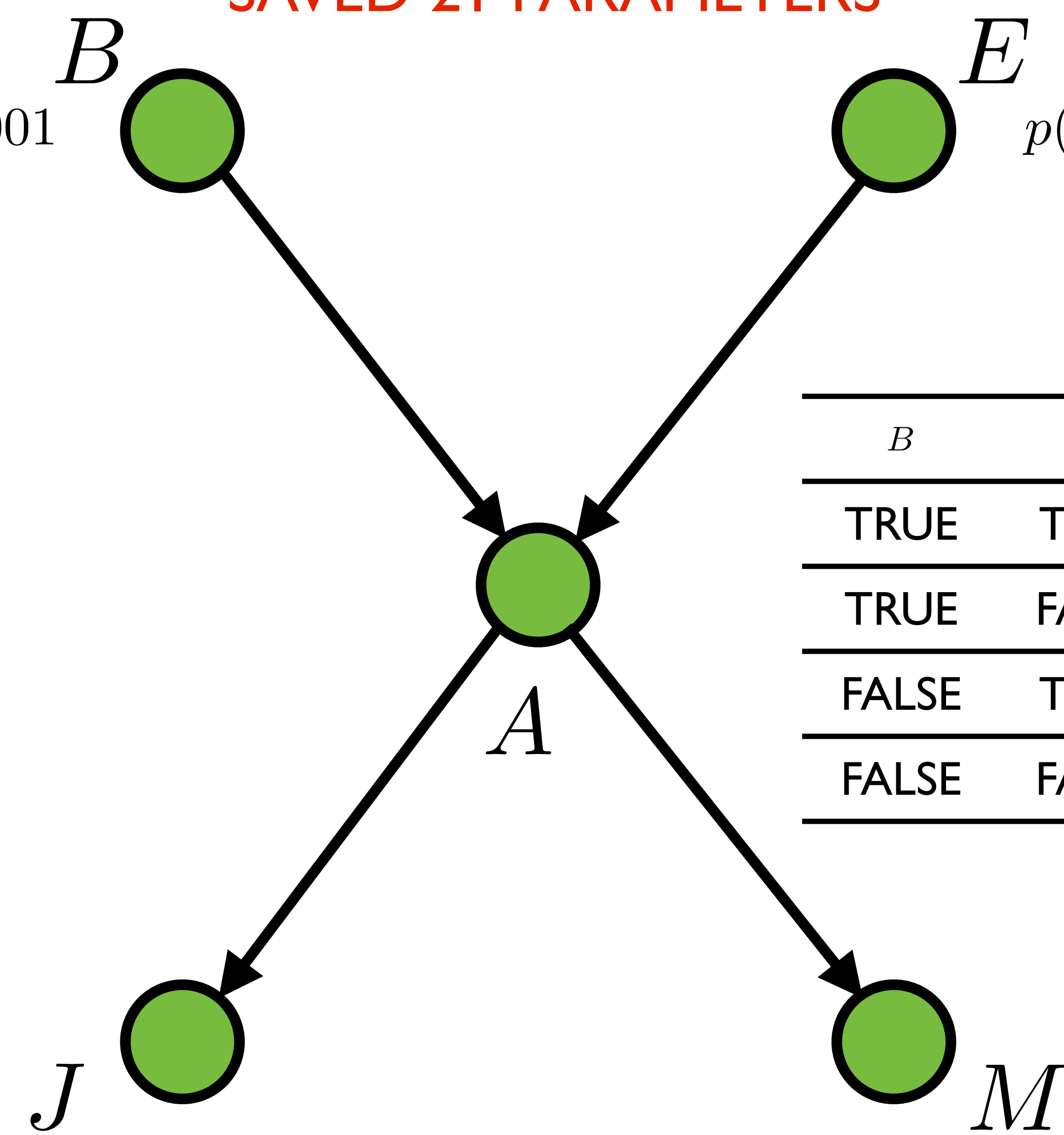
A	$p(M\|A)$
TRUE	0.70
FALSE	0.01

TOTAL # OF PARAMETERS: 10

DOMAIN KNOWLEDGE
SAVED 21 PARAMETERS

$p(B = \texttt{TRUE}) = 0.001$

$p(E = \texttt{TRUE}) = 0.002$

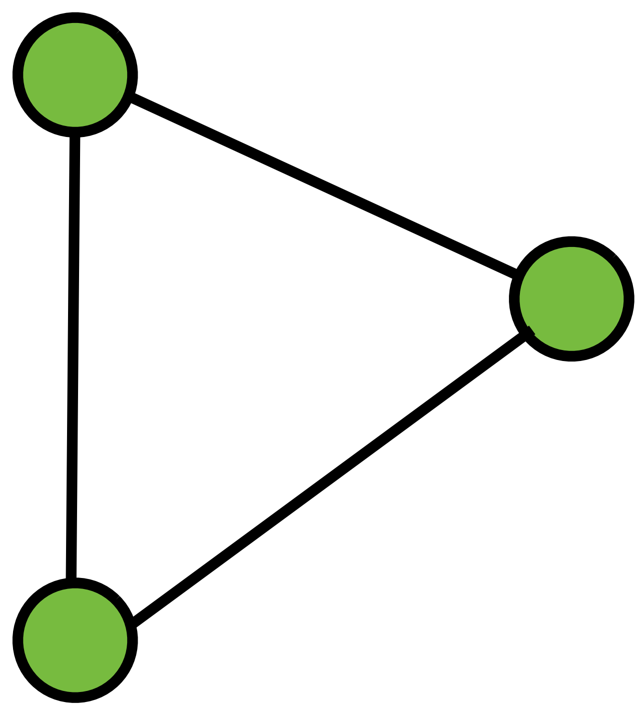


B	E	$p(A\|B, E)$
TRUE	TRUE	0.95
TRUE	FALSE	0.95
FALSE	TRUE	0.29
FALSE	FALSE	0.001

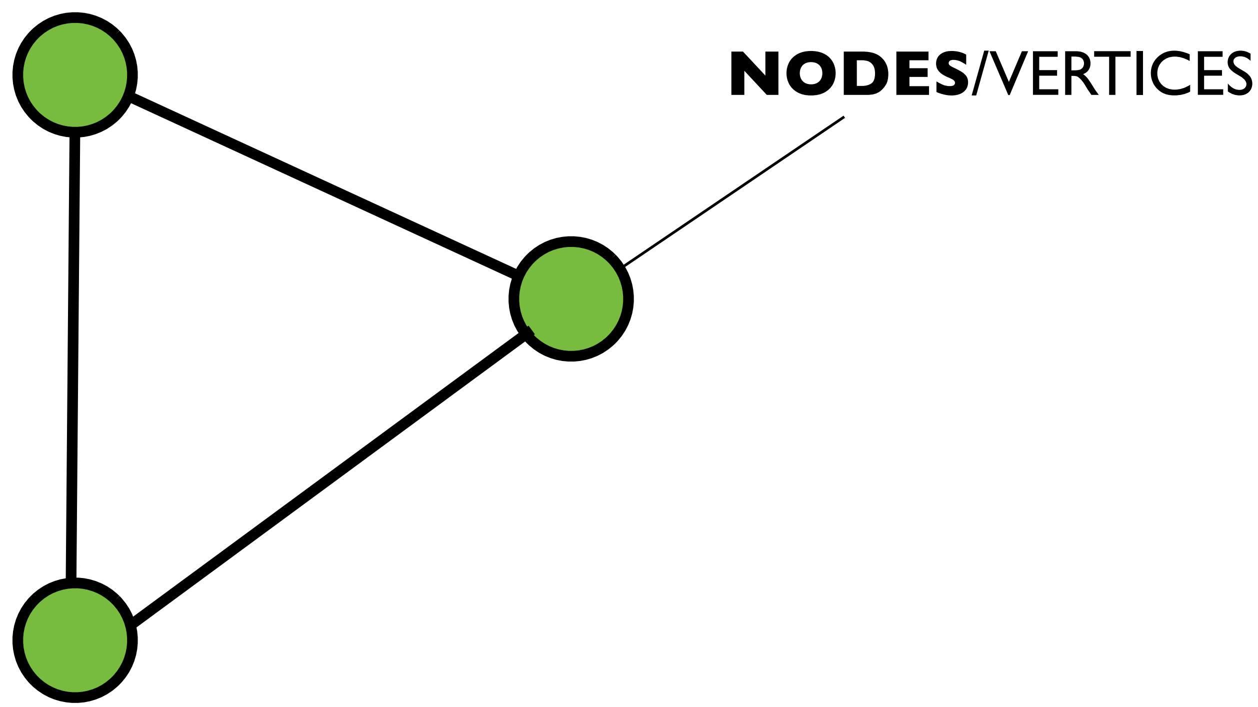
A	$p(J\|A)$
TRUE	0.90
FALSE	0.05

A	$p(M\|A)$
TRUE	0.70
FALSE	0.01

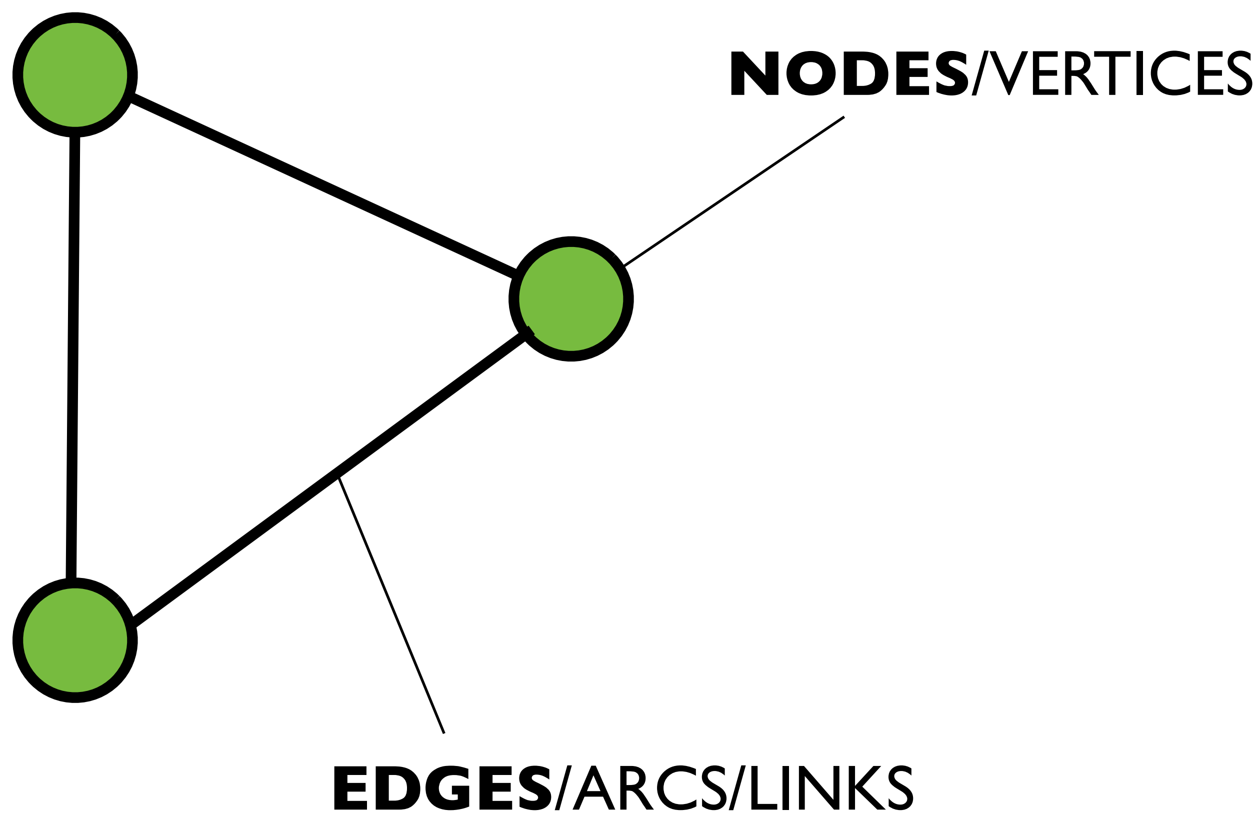
GRAPHS



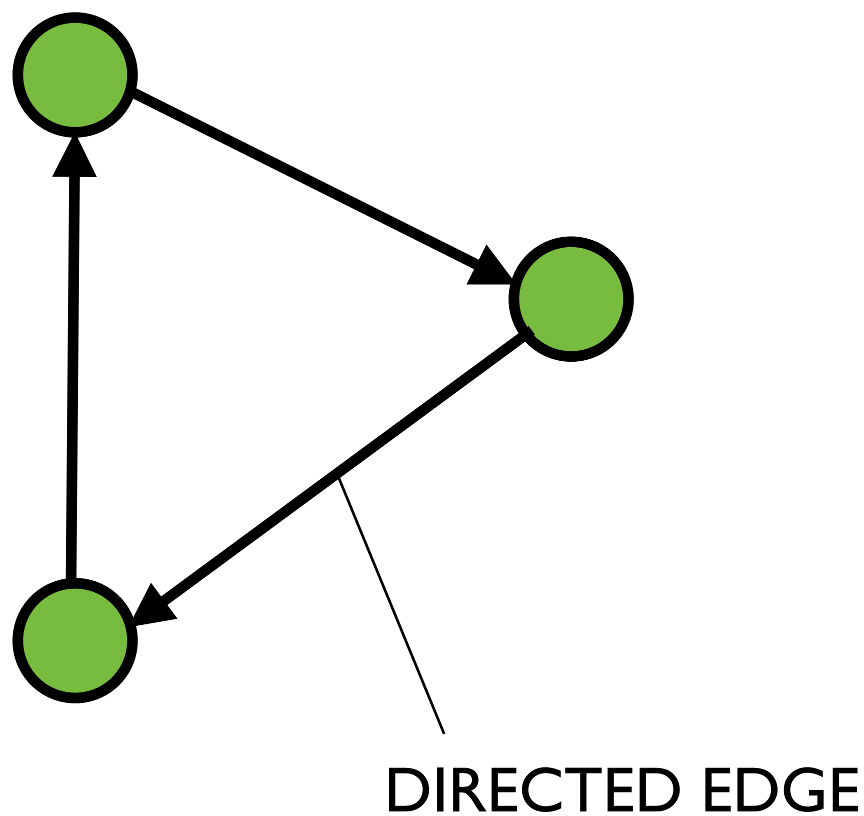
GRAPHS



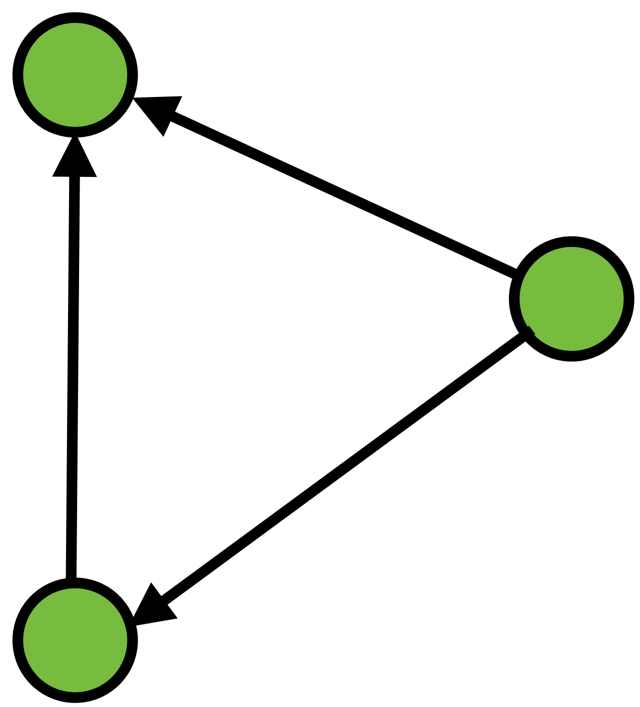
GRAPHS



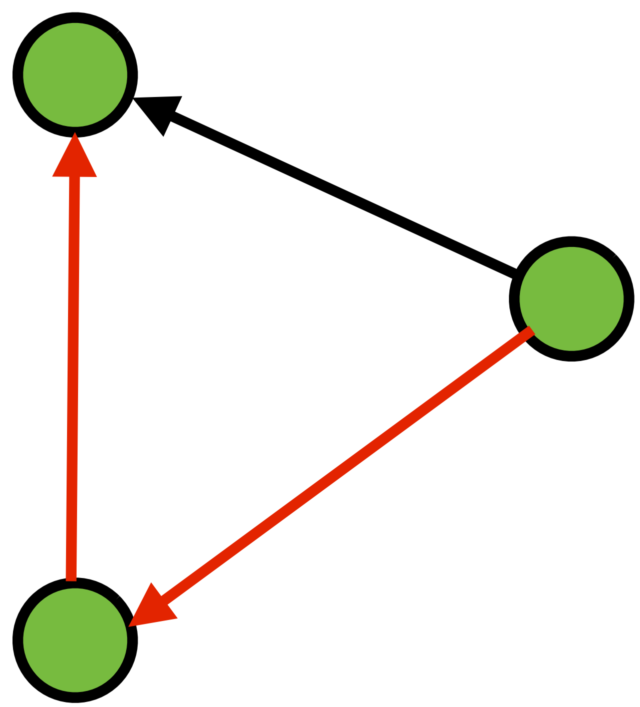
DIRECTED GRAPHS



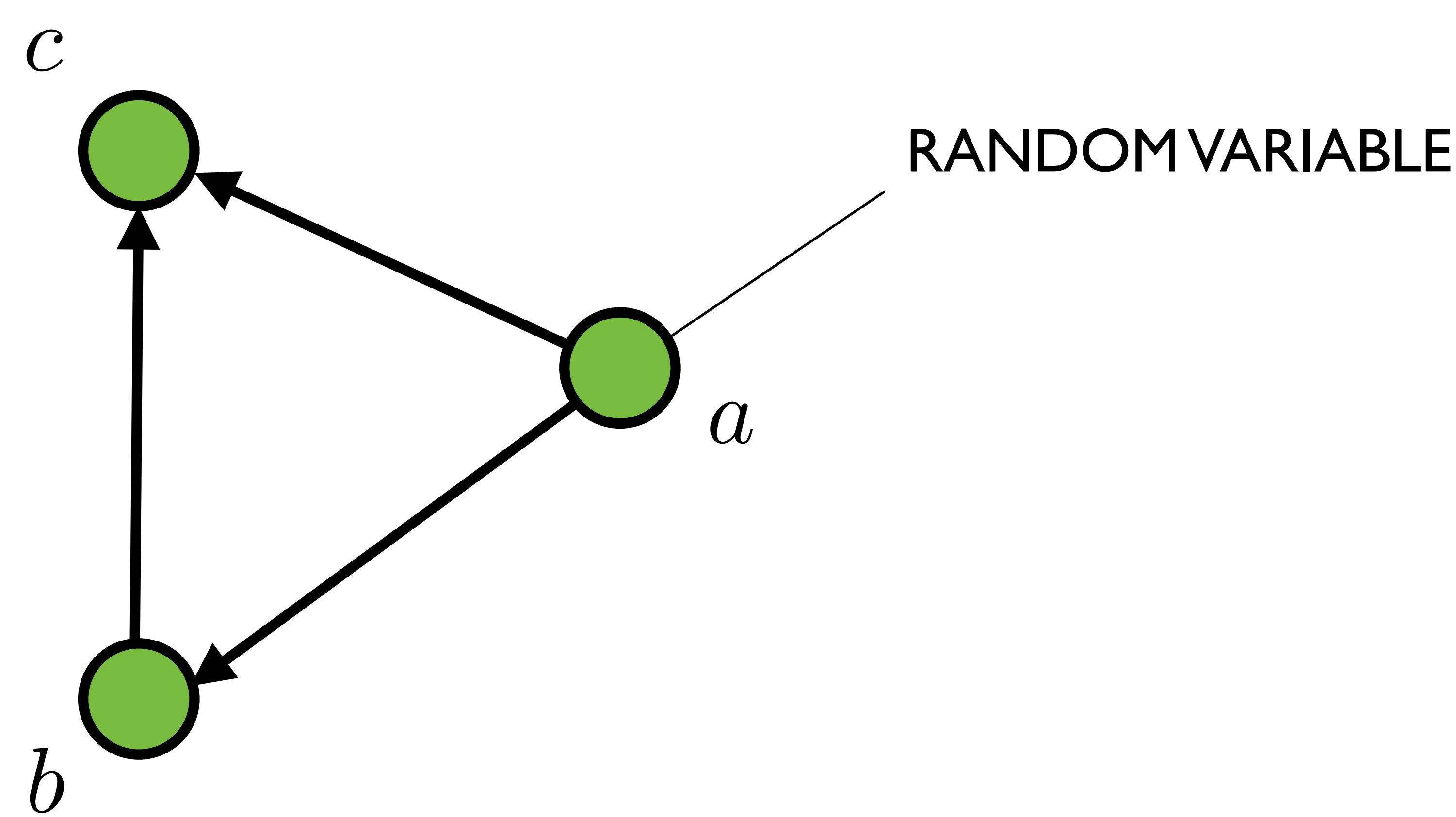
DIRECTED ACYCLIC GRAPHS (DAG)



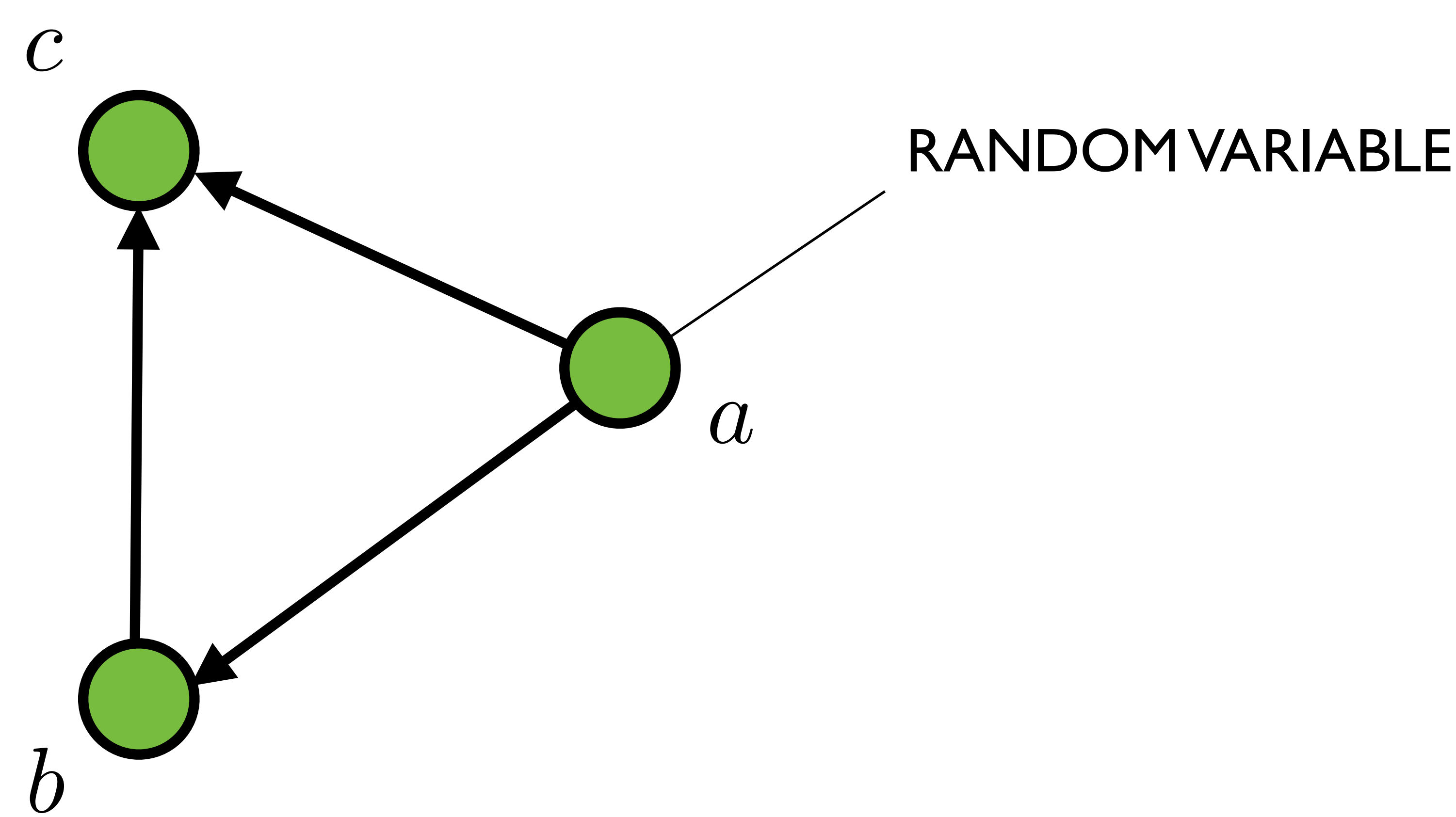
PATH



PROBABILISTIC GRAPHICAL MODELS



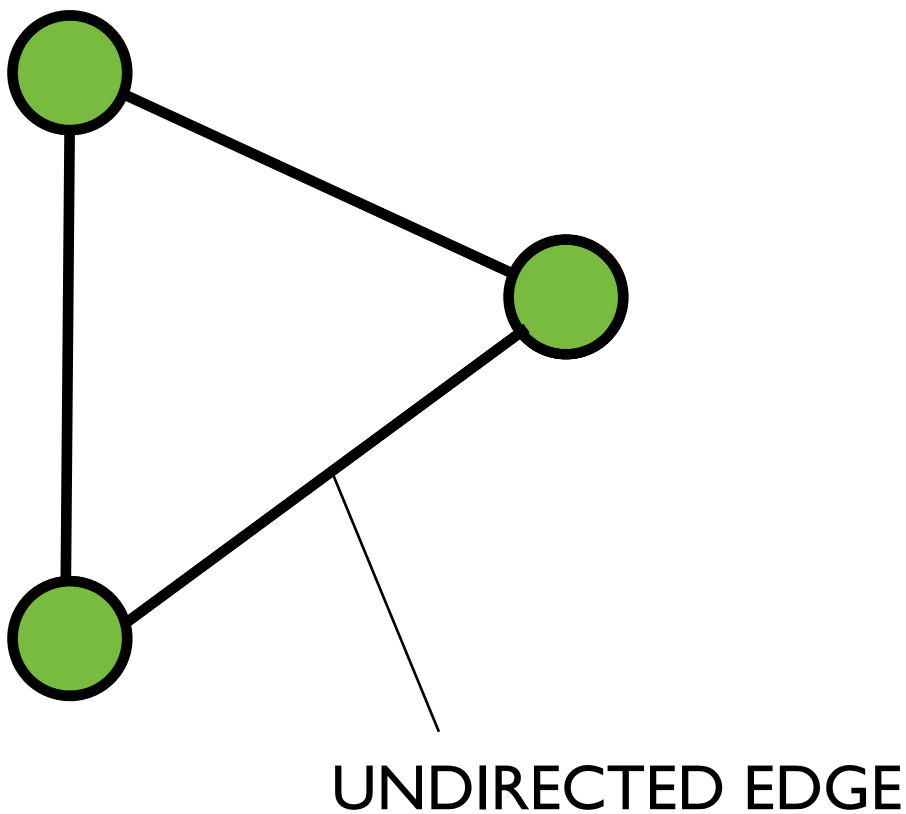
PROBABILISTIC GRAPHICAL MODELS



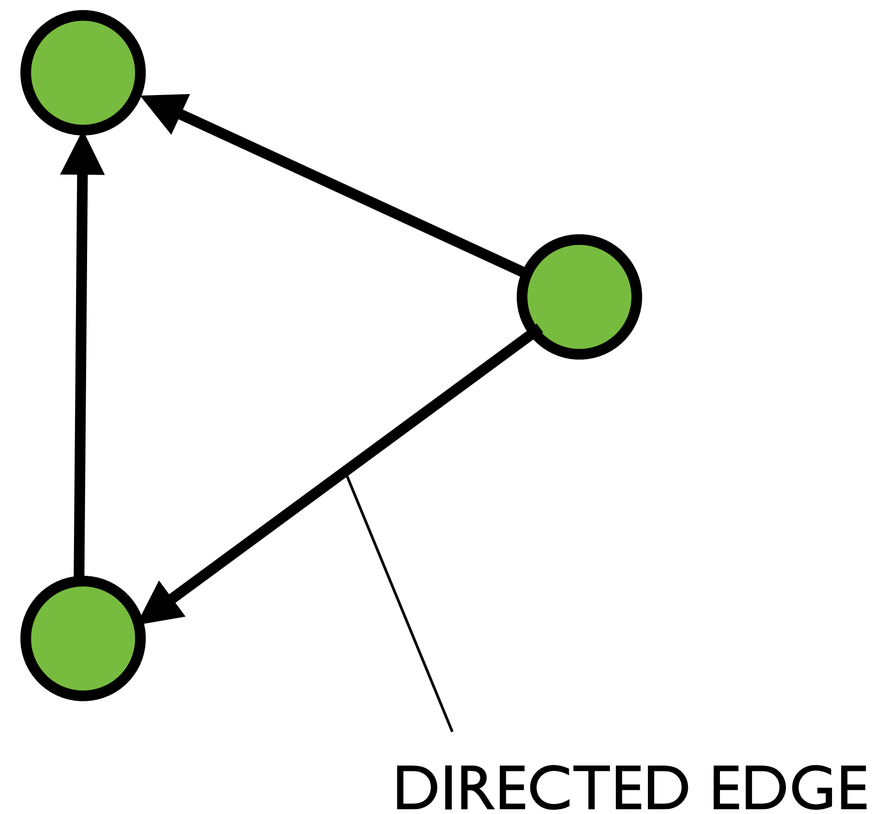
a is the **parent** of *b*

c is the **child** of *a* and *b*

PROBABILISTIC GRAPHICAL MODELS



MARKOV RANDOM FIELDS
UNDIRECTED GRAPHICAL MODELS



BAYESIAN NETWORKS
BELIEF NETWORKS
DIRECTED ACYCLIC GRAPHICAL MODELS

BAYESIAN NETWORKS

$$p(a, b, c)$$

BAYESIAN NETWORKS

$$p(a, b, c) = p(c|a, b)p(a, b)$$ **PRODUCT RULE**

BAYESIAN NETWORKS

$$p(a, b, c) = p(c|a, b)p(a, b)$$ **PRODUCT RULE**

$$= p(c|a, b)p(b|a)p(a)$$

BAYESIAN NETWORKS

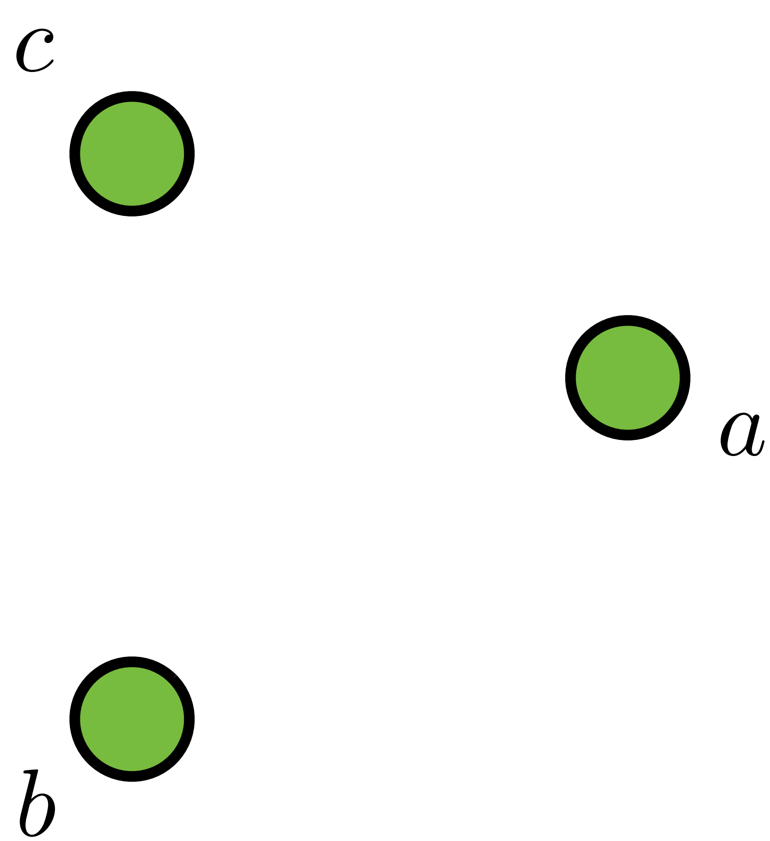
$$p(a, b, c) = p(c|a, b)p(a, b)$$ **PRODUCT RULE**

$$= \underline{p(c|a, b)}\,\underline{p(b|a)}\,\underline{p(a)}$$ Factors

BAYESIAN NETWORKS

$$p(a,b,c) = p(c|a,b)p(a,b)$$ **PRODUCT RULE**

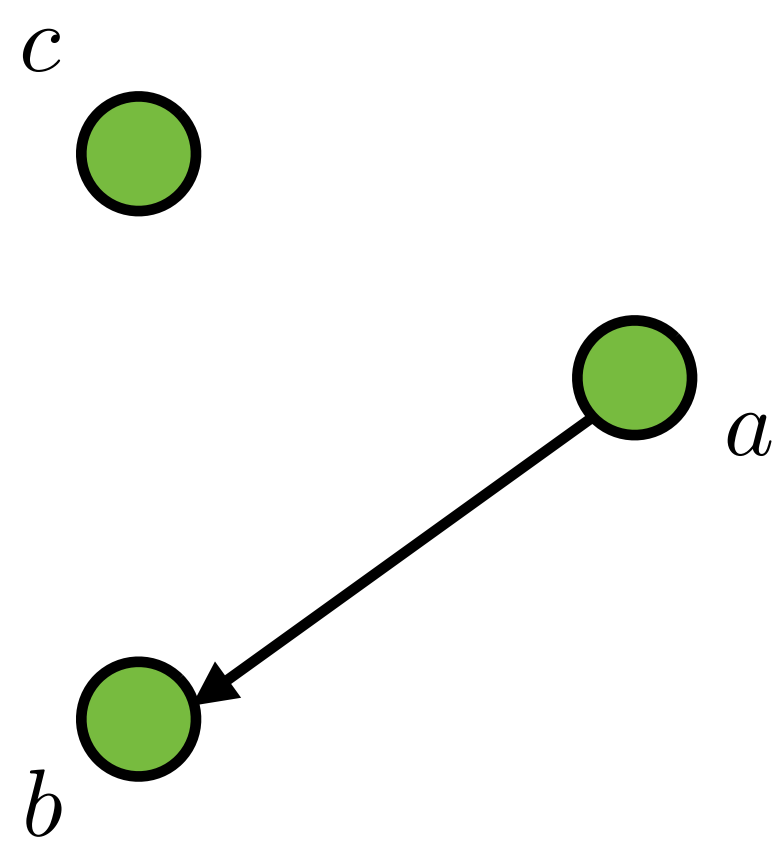
$$= \underline{p(c|a,b)}\,\underline{p(b|a)}\,\underline{p(a)}$$ Factors



BAYESIAN NETWORKS

$$p(a, b, c) = p(c|a, b)p(a, b)$$ **PRODUCT RULE**

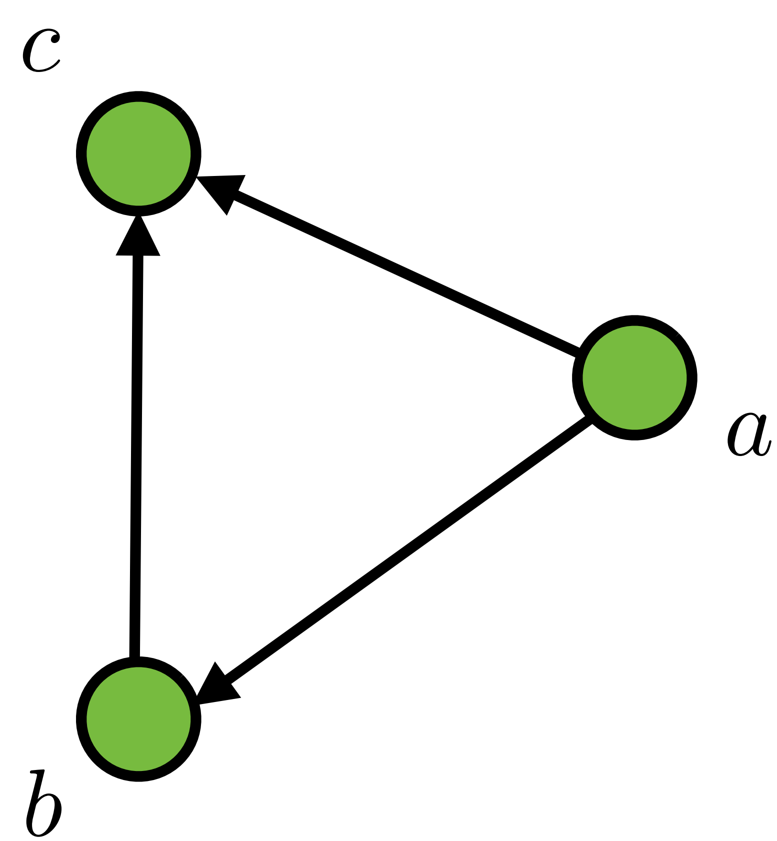
$$= \underline{p(c|a, b)}\,\underline{p(b|a)}\,\underline{p(a)}$$ Factors



BAYESIAN NETWORKS

$$p(a, b, c) = p(c|a, b)p(a, b)$$ **PRODUCT RULE**

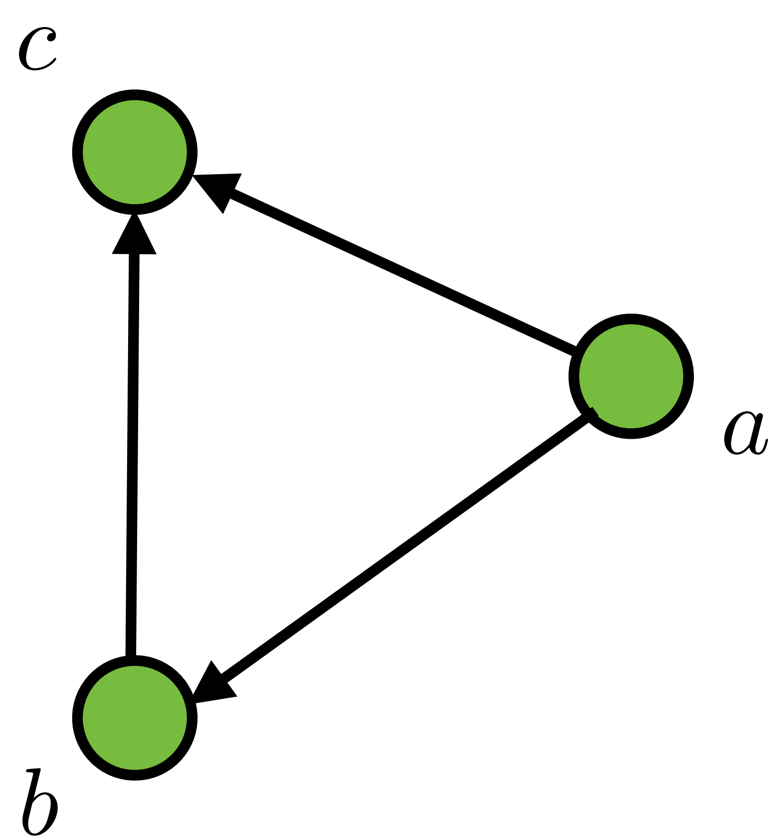
$$= \underline{p(c|a, b)}\,\underline{p(b|a)}\,\underline{p(a)}$$ Factors



BAYESIAN NETWORKS

ORDERING

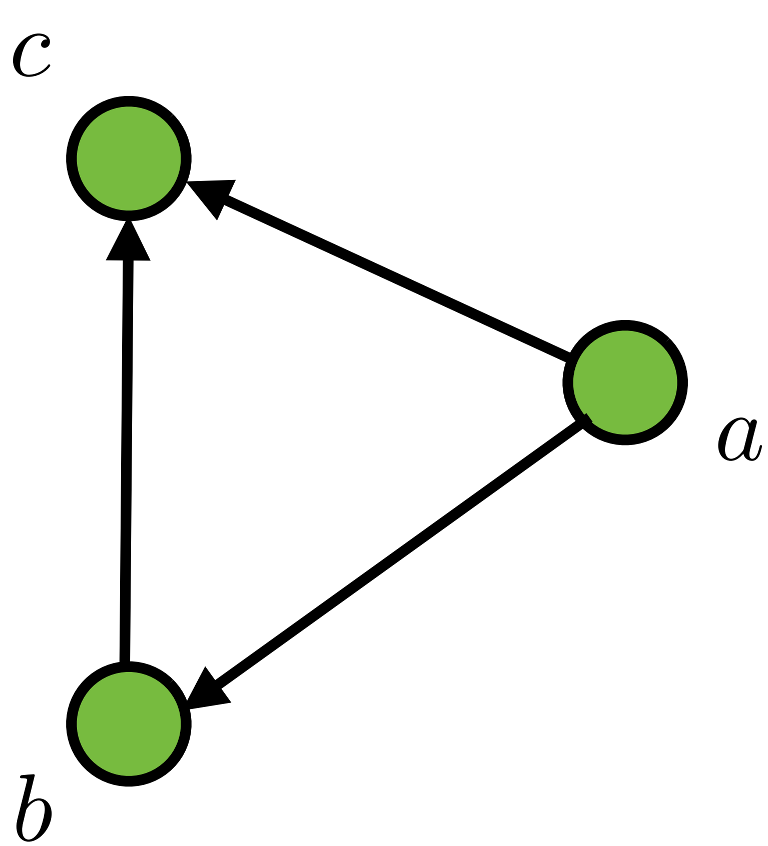
$$p(a, b, c) = p(c|a, b)p(b|a)p(a)$$



BAYESIAN NETWORKS

ORDERING

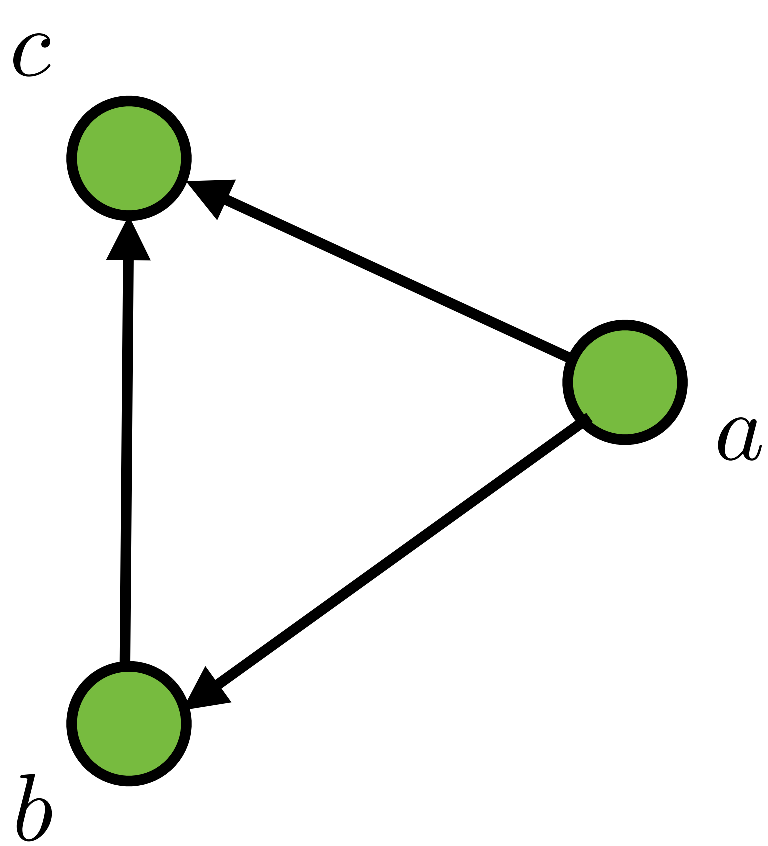
$$p(a,b,c) = p(c|a,b)p(b|a)p(a) \qquad p(a,b,c) = p(b|c,a)p(a|c)p(c)$$

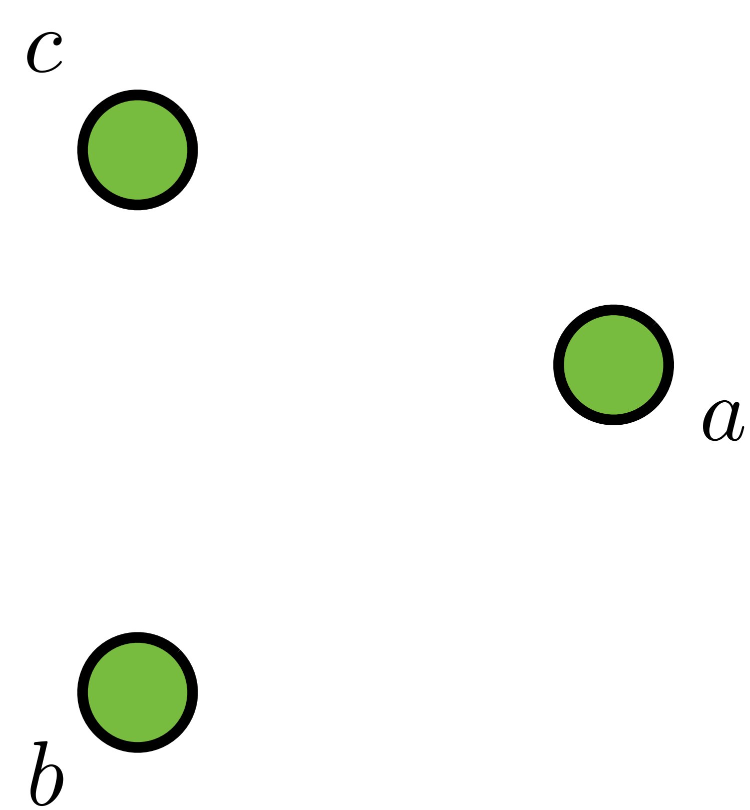


BAYESIAN NETWORKS

ORDERING

$$p(a, b, c) = p(c|a, b)p(b|a)p(a) \qquad p(a, b, c) = p(b|c, a)p(a|c)p(c)$$

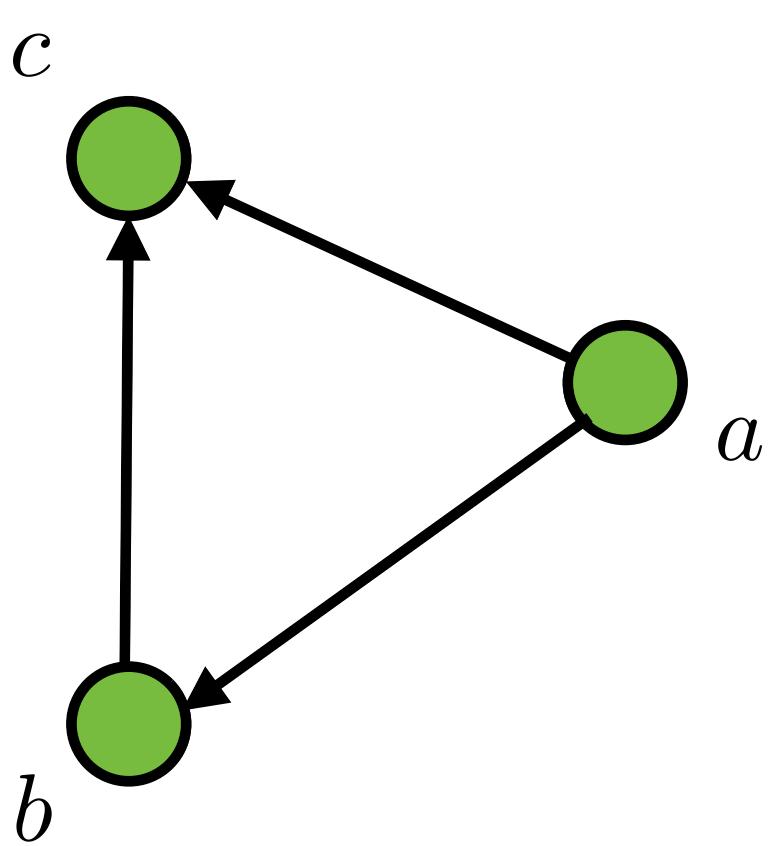




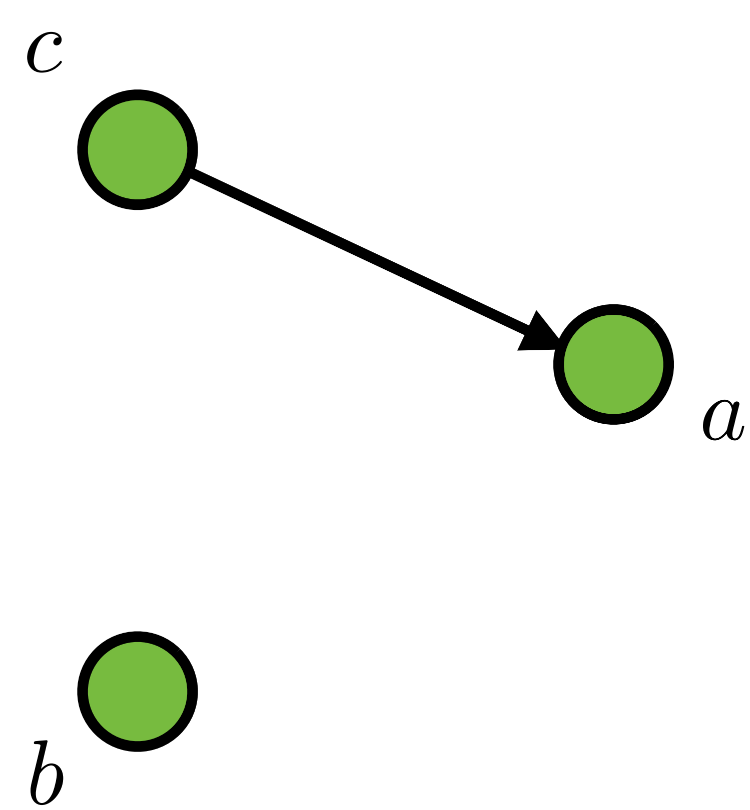
BAYESIAN NETWORKS

ORDERING

$$p(a, b, c) = p(c|a, b)p(b|a)p(a)$$



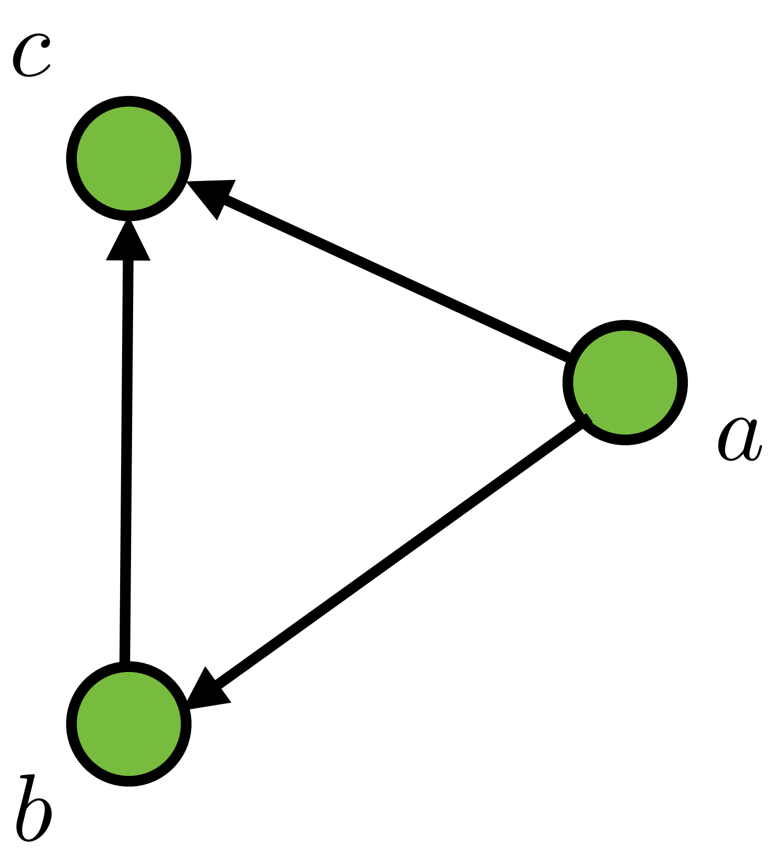
$$p(a, b, c) = p(b|c, a)p(a|c)p(c)$$



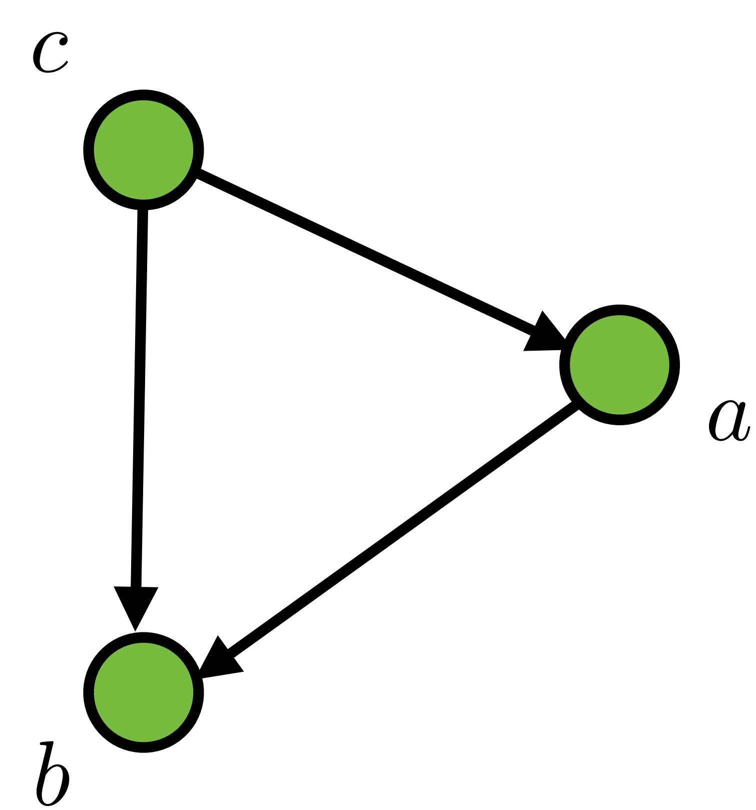
BAYESIAN NETWORKS

ORDERING

$$p(a,b,c) = p(c|a,b)p(b|a)p(a)$$

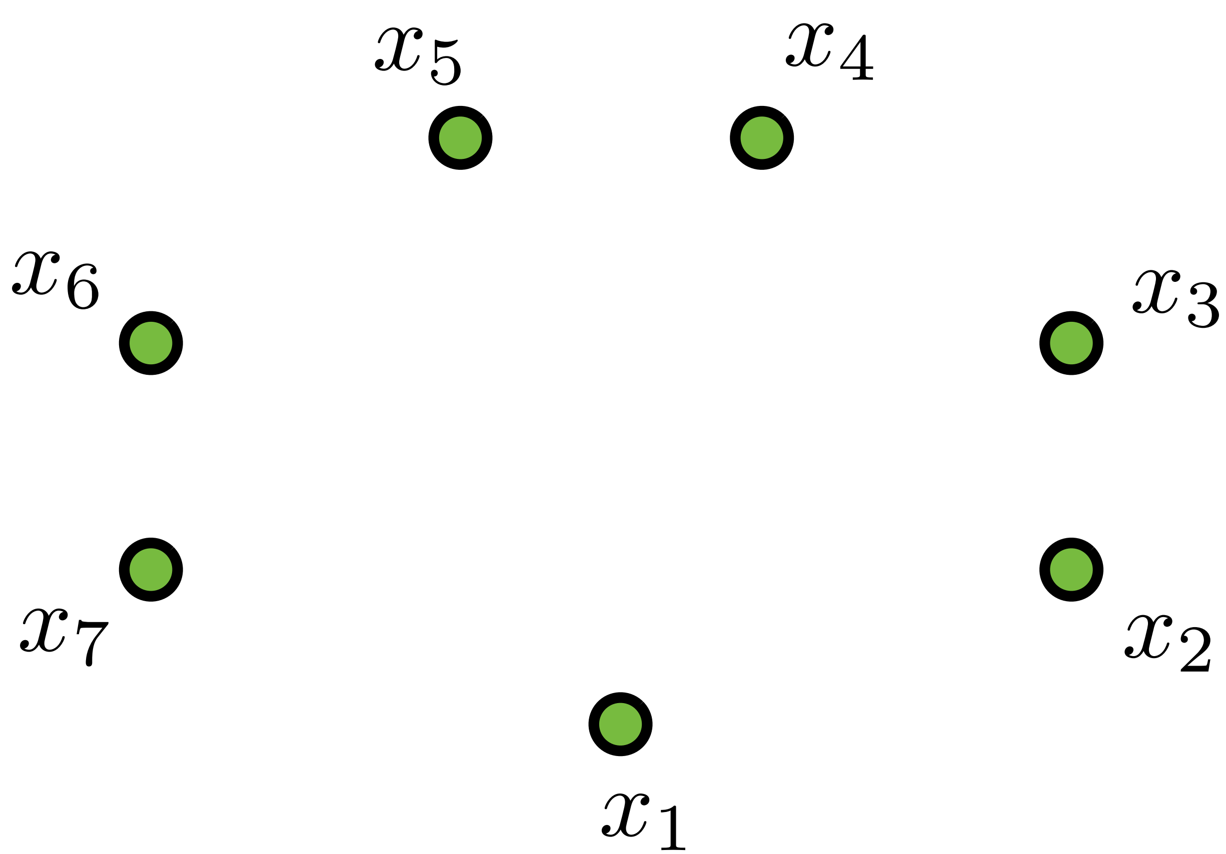


$$p(a,b,c) = p(b|c,a)p(a|c)p(c)$$



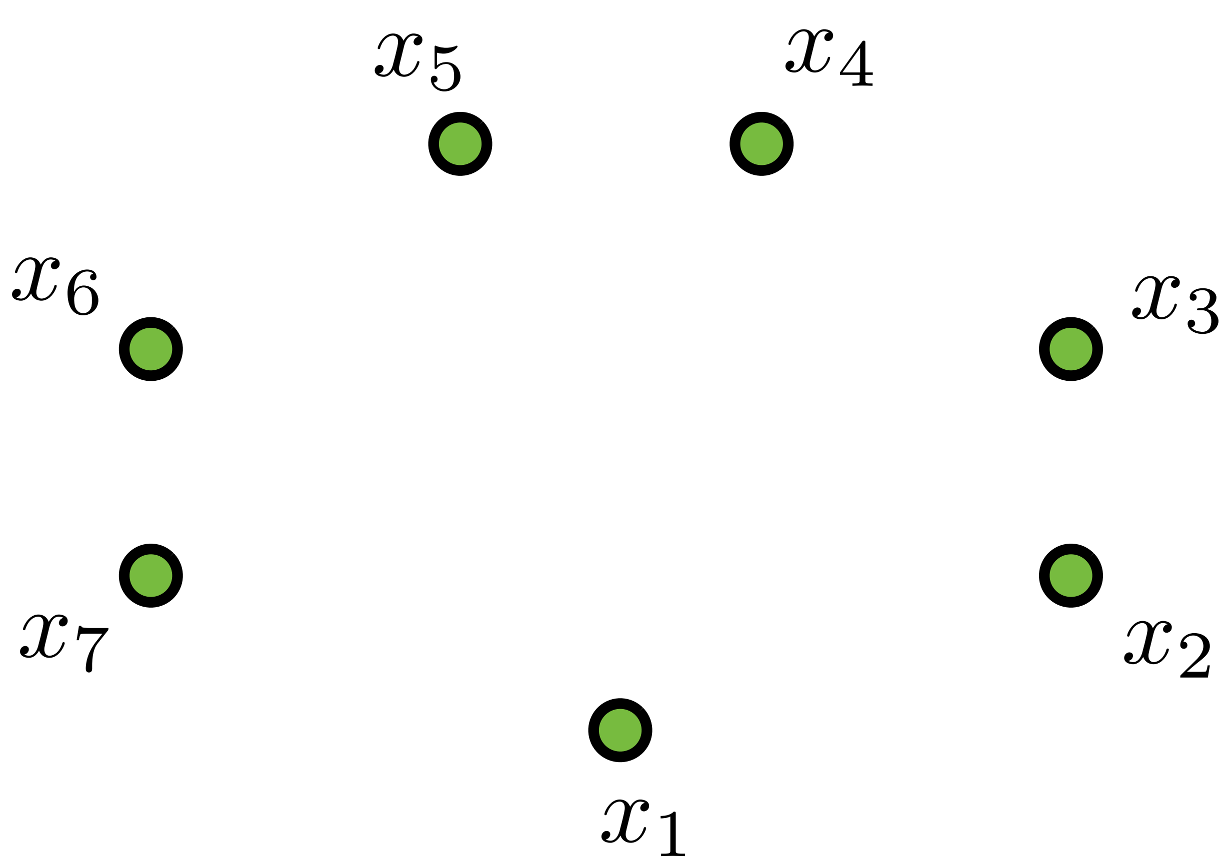
BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2, x_3, x_4, x_5, x_6, x_7)$$



BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2, x_3, x_4, x_5, x_6, x_7)$$
$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2|x_3, x_4, x_5, x_6, x_7)p(x_3, x_4, x_5, x_6, x_7)$$

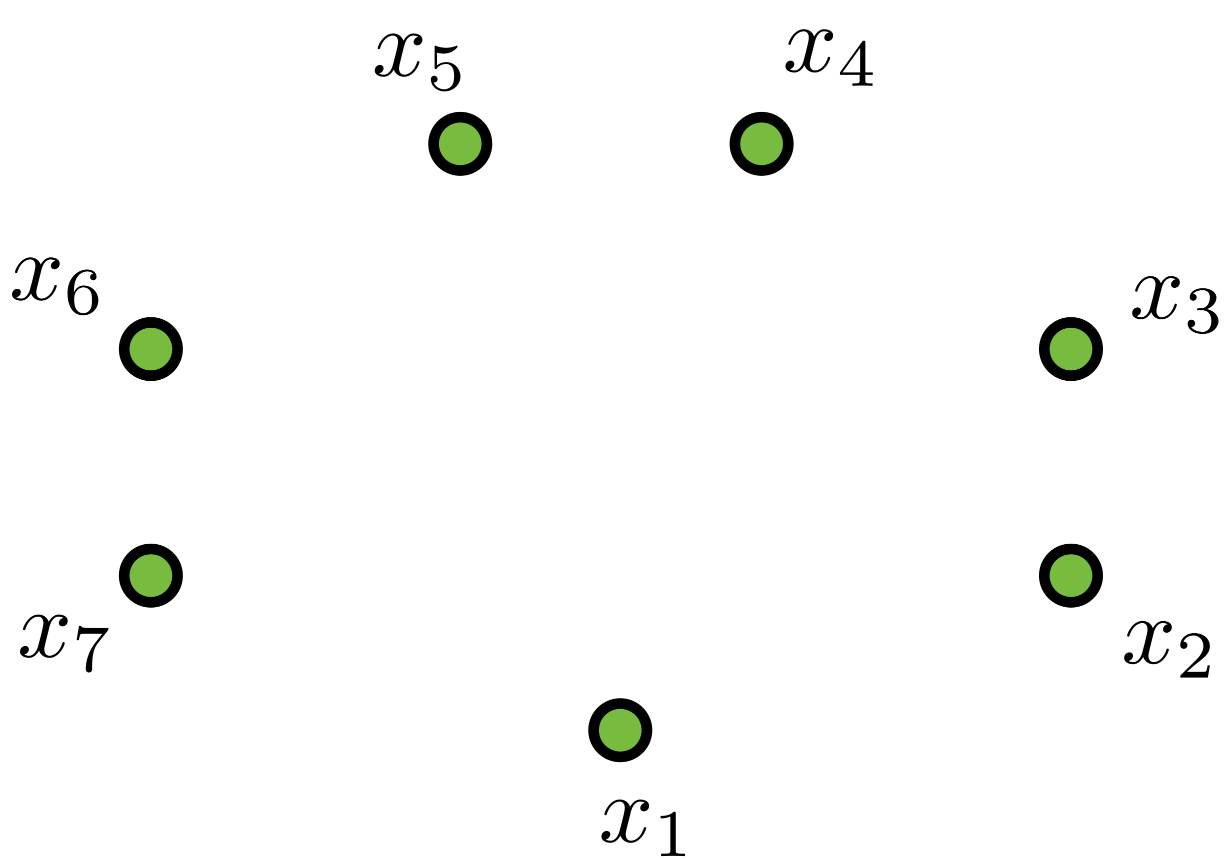


BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2, x_3, x_4, x_5, x_6, x_7)$$

$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2|x_3, x_4, x_5, x_6, x_7)p(x_3, x_4, x_5, x_6, x_7)$$

$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)$$



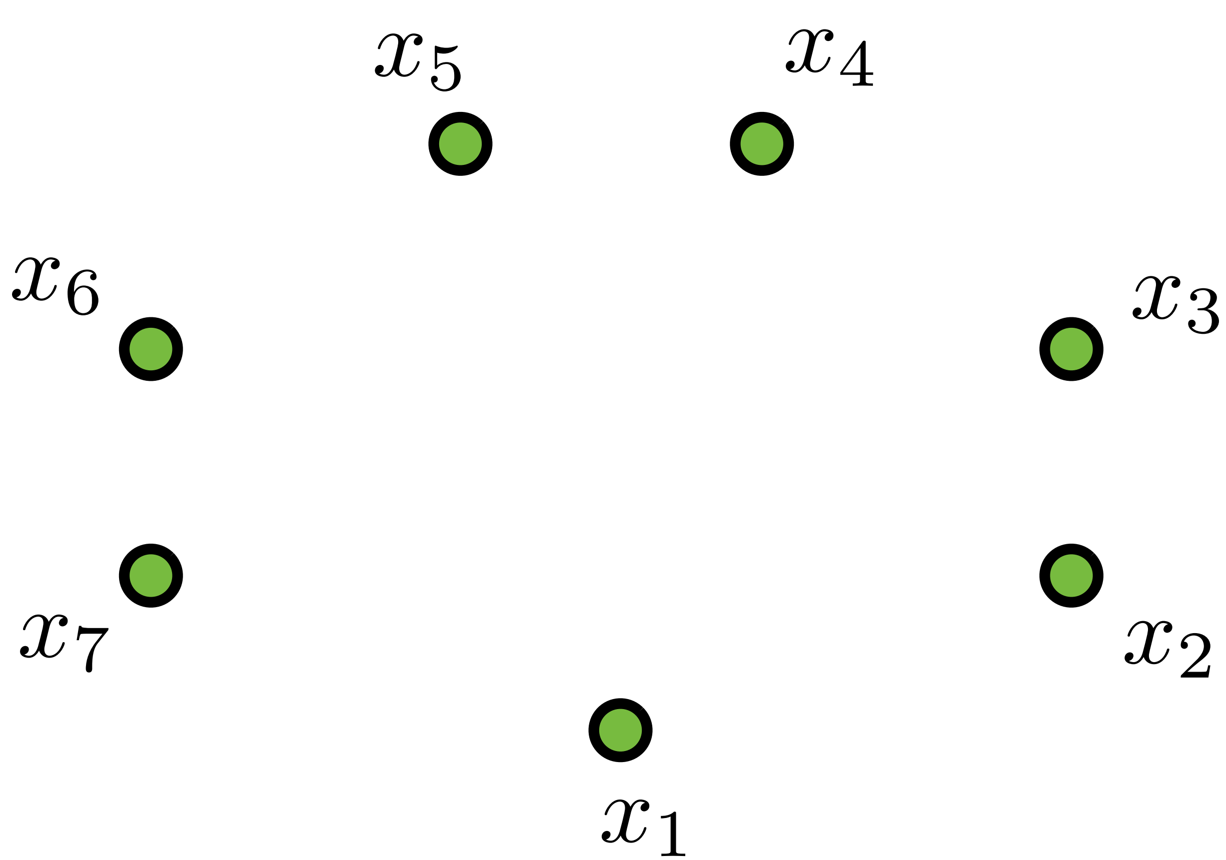
BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2, x_3, x_4, x_5, x_6, x_7)$$

$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2|x_3, x_4, x_5, x_6, x_7)p(x_3, x_4, x_5, x_6, x_7)$$

$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)$$

$$p(x_2|x_3, x_4, x_5, x_6, x_7)$$

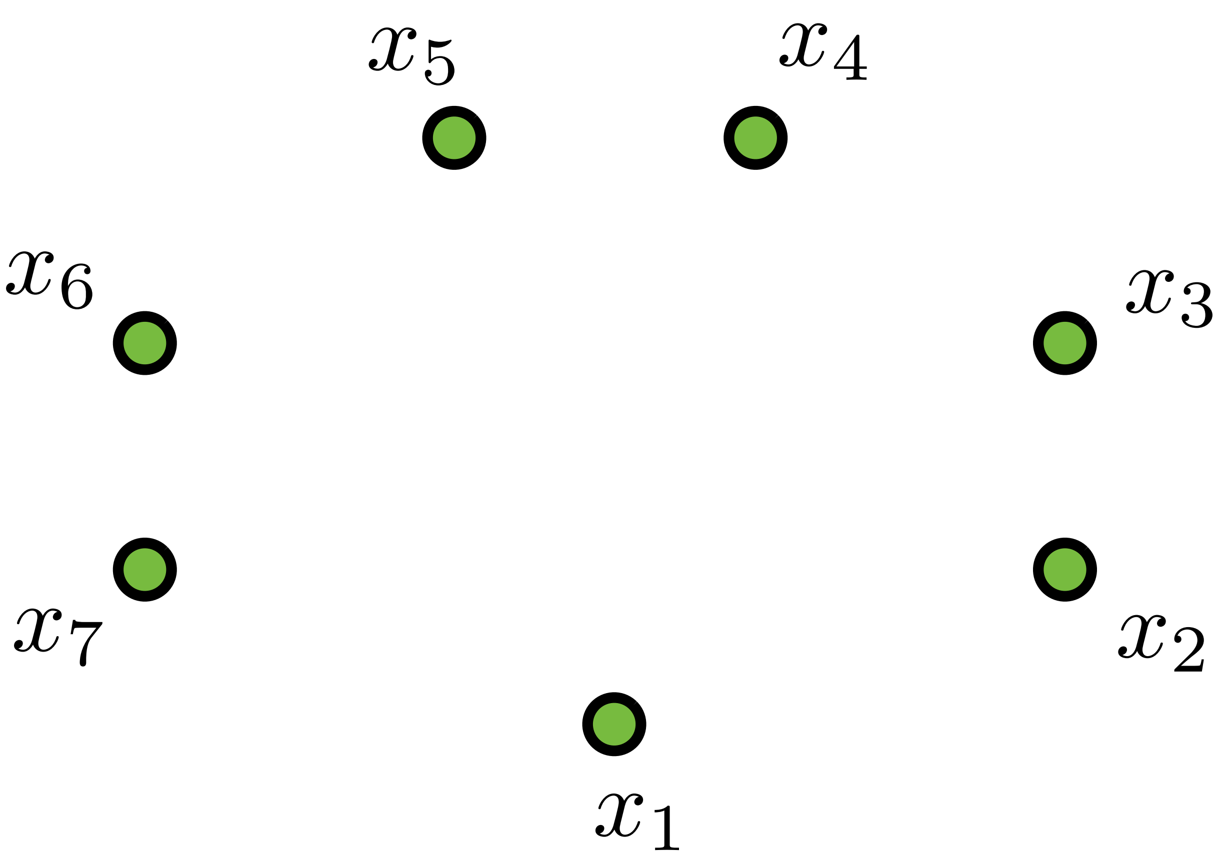


BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2, x_3, x_4, x_5, x_6, x_7)$$

$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2|x_3, x_4, x_5, x_6, x_7)p(x_3, x_4, x_5, x_6, x_7)$$

$$\begin{aligned} = \; & p(x_1|x_2, x_3, x_4, x_5, x_6, x_7) \\ & p(x_2|x_3, x_4, x_5, x_6, x_7) \\ & p(x_3|x_4, x_5, x_6, x_7) \end{aligned}$$

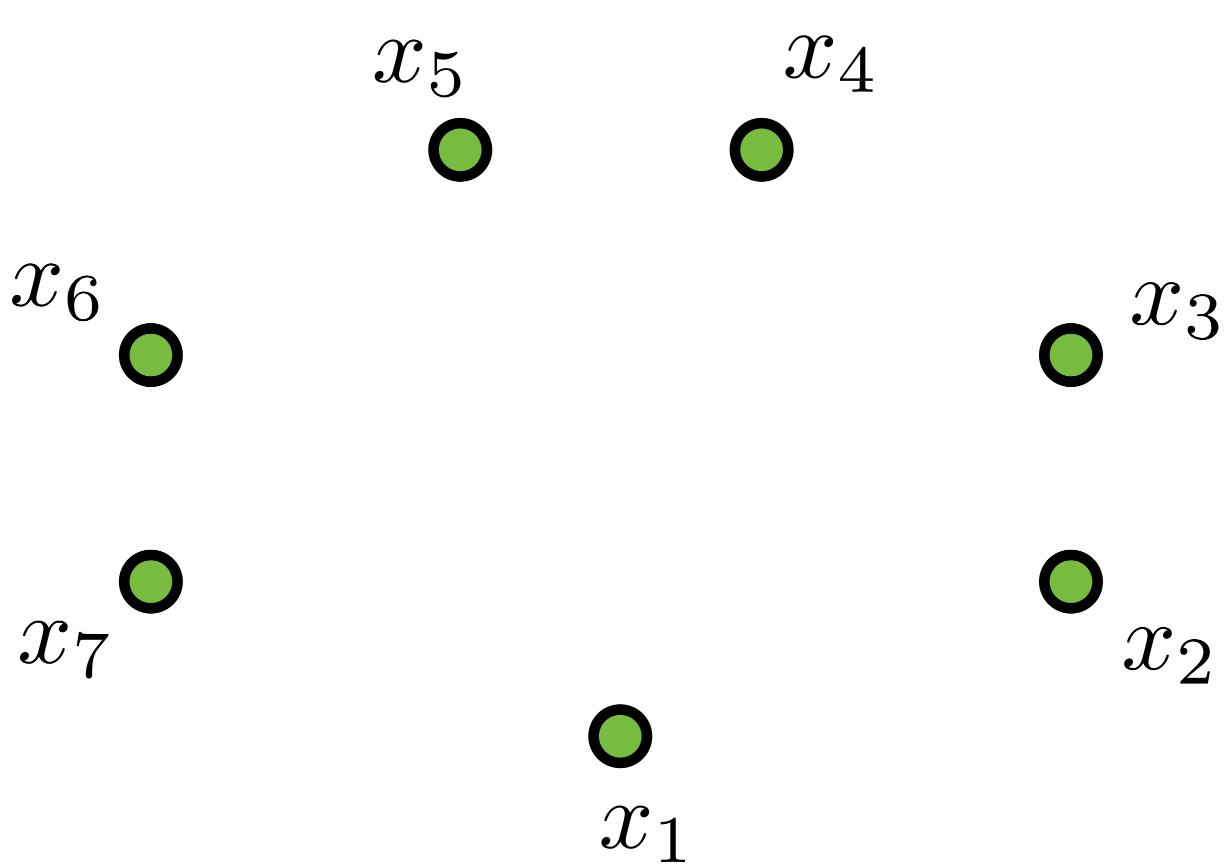


BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2, x_3, x_4, x_5, x_6, x_7)$$

$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2|x_3, x_4, x_5, x_6, x_7)p(x_3, x_4, x_5, x_6, x_7)$$

$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)$$
$$p(x_2|x_3, x_4, x_5, x_6, x_7)$$
$$p(x_3|x_4, x_5, x_6, x_7)$$
$$p(x_4|x_5, x_6, x_7)$$

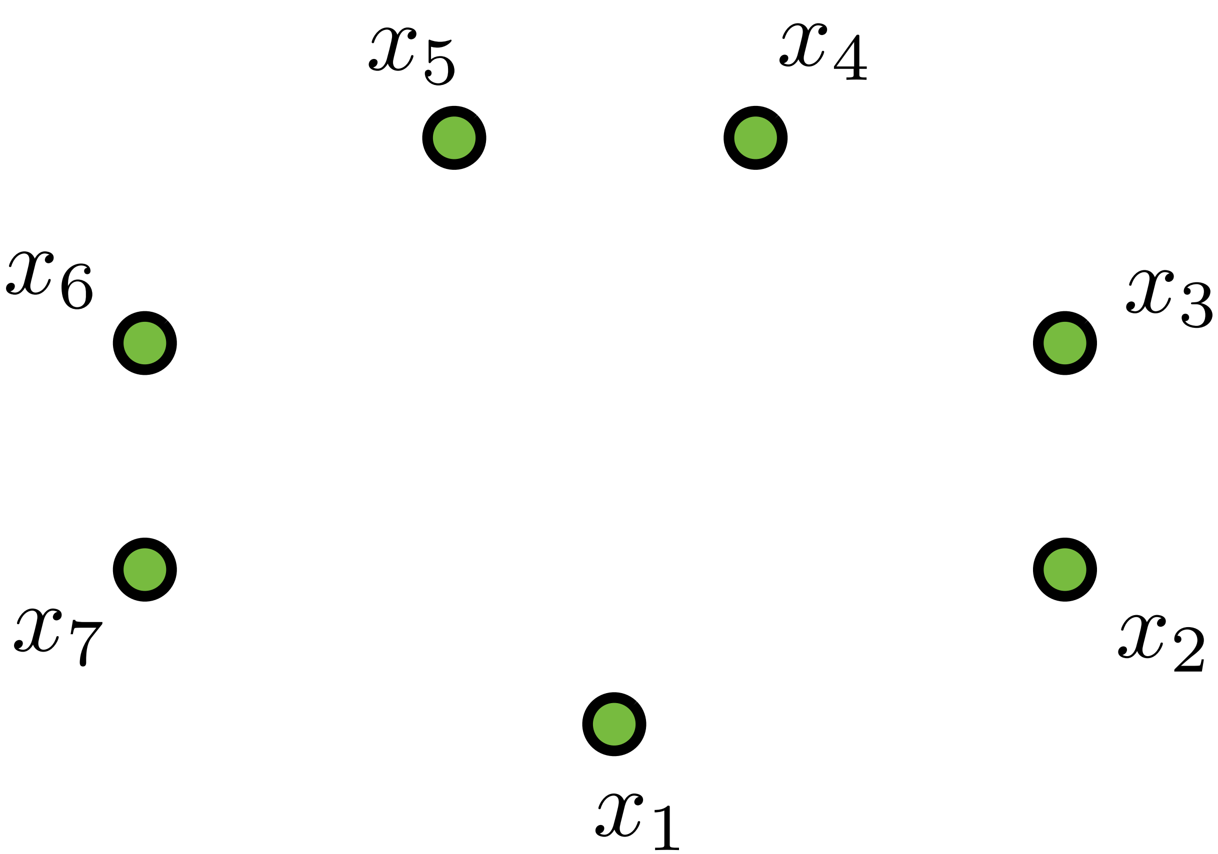


BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2, x_3, x_4, x_5, x_6, x_7)$$

$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2|x_3, x_4, x_5, x_6, x_7)p(x_3, x_4, x_5, x_6, x_7)$$

$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)$$
$$p(x_2|x_3, x_4, x_5, x_6, x_7)$$
$$p(x_3|x_4, x_5, x_6, x_7)$$
$$p(x_4|x_5, x_6, x_7)$$
$$p(x_5|x_6, x_7)$$

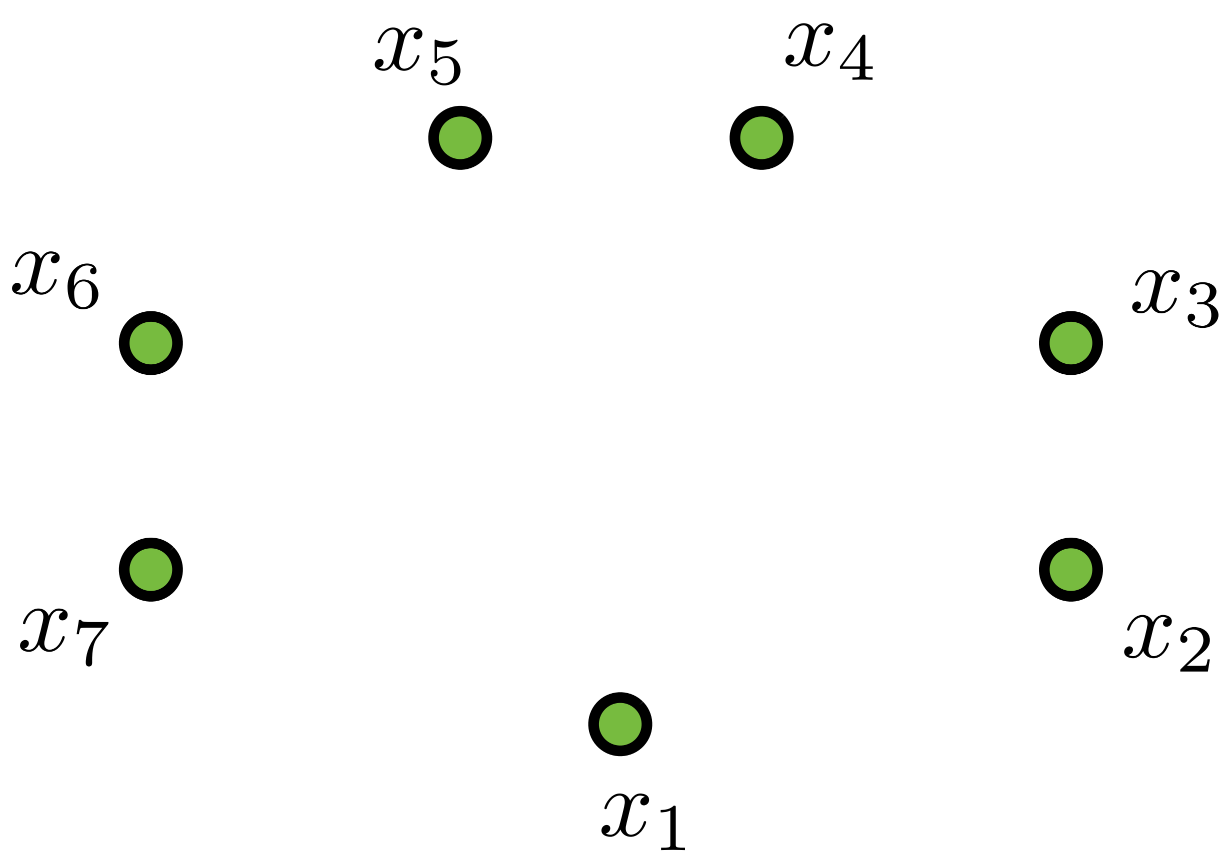


BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2, x_3, x_4, x_5, x_6, x_7)$$

$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2|x_3, x_4, x_5, x_6, x_7)p(x_3, x_4, x_5, x_6, x_7)$$

$$\begin{aligned} = \; & p(x_1|x_2, x_3, x_4, x_5, x_6, x_7) \\ & p(x_2|x_3, x_4, x_5, x_6, x_7) \\ & p(x_3|x_4, x_5, x_6, x_7) \\ & p(x_4|x_5, x_6, x_7) \\ & p(x_5|x_6, x_7) \\ & p(x_6|x_7) \end{aligned}$$

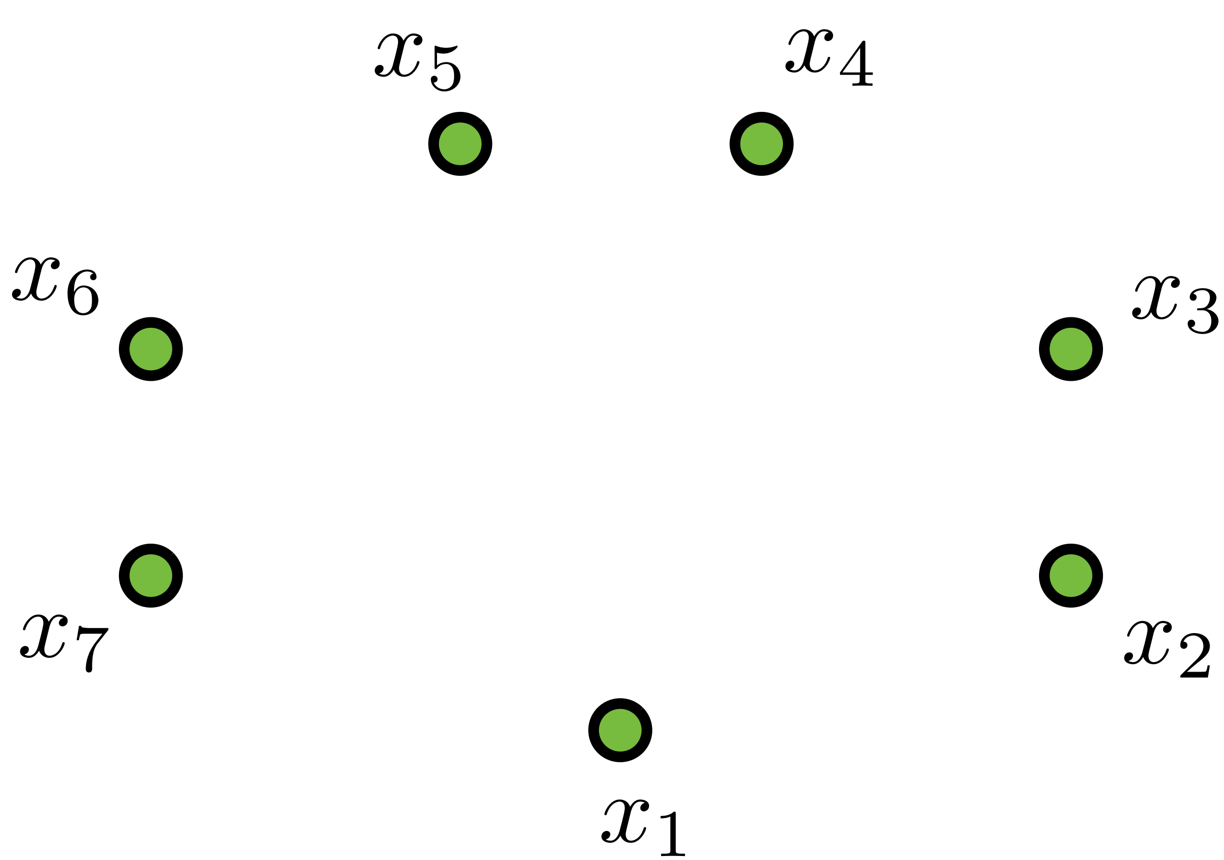


BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2, x_3, x_4, x_5, x_6, x_7)$$

$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2|x_3, x_4, x_5, x_6, x_7)p(x_3, x_4, x_5, x_6, x_7)$$

$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)$$
$$p(x_2|x_3, x_4, x_5, x_6, x_7)$$
$$p(x_3|x_4, x_5, x_6, x_7)$$
$$p(x_4|x_5, x_6, x_7)$$
$$p(x_5|x_6, x_7)$$
$$p(x_6|x_7)$$
$$p(x_7)$$

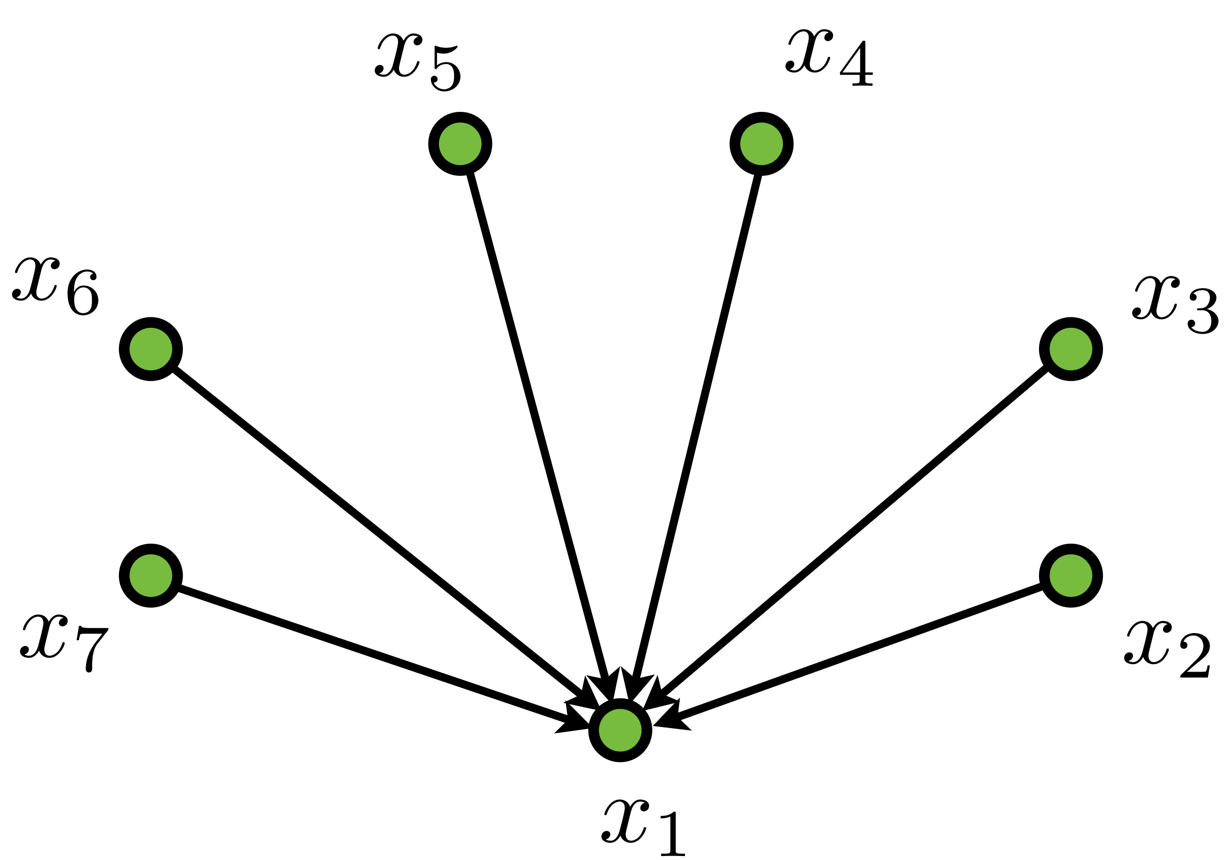


BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2, x_3, x_4, x_5, x_6, x_7)$$

$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2|x_3, x_4, x_5, x_6, x_7)p(x_3, x_4, x_5, x_6, x_7)$$

$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)$$
$$p(x_2|x_3, x_4, x_5, x_6, x_7)$$
$$p(x_3|x_4, x_5, x_6, x_7)$$
$$p(x_4|x_5, x_6, x_7)$$
$$p(x_5|x_6, x_7)$$
$$p(x_6|x_7)$$
$$p(x_7)$$



BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2, x_3, x_4, x_5, x_6, x_7)$$

$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2|x_3, x_4, x_5, x_6, x_7)p(x_3, x_4, x_5, x_6, x_7)$$

$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)$$

$$p(x_2|x_3, x_4, x_5, x_6, x_7)$$

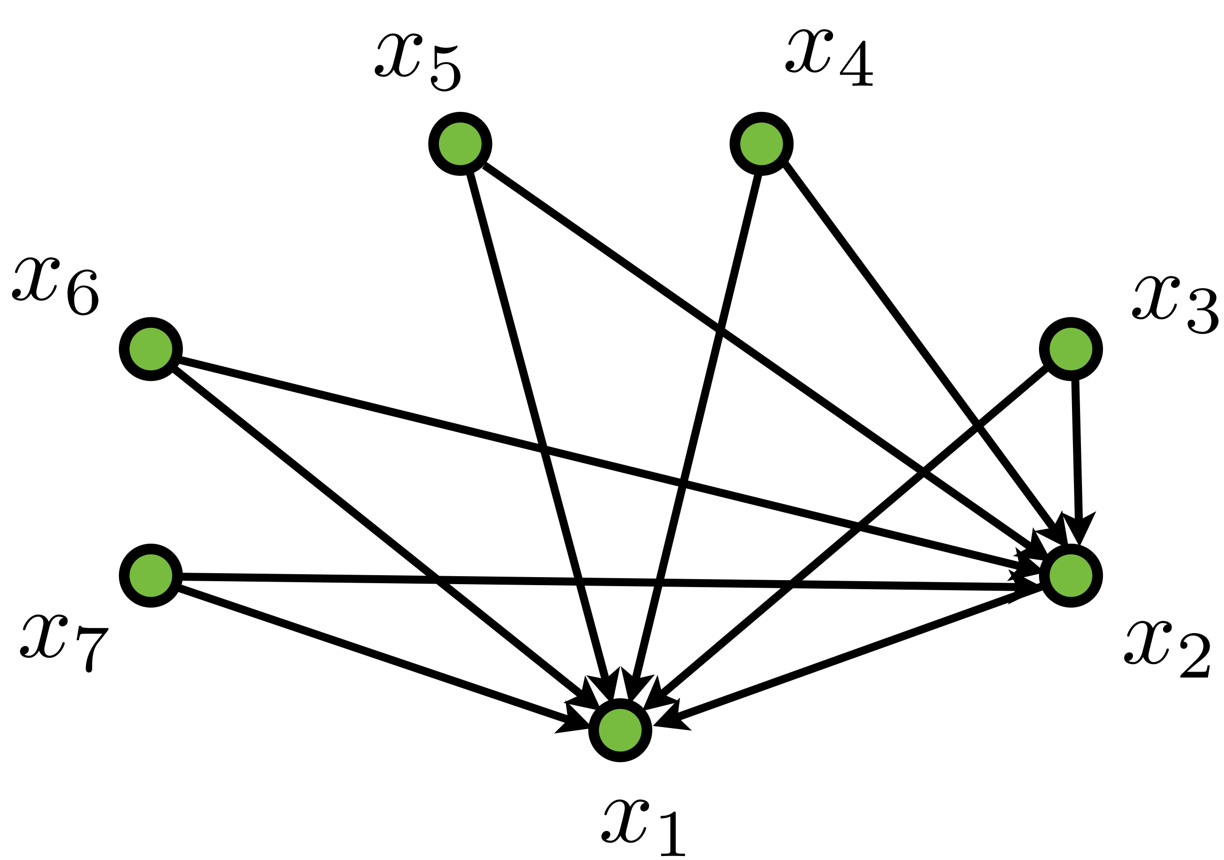
$$p(x_3|x_4, x_5, x_6, x_7)$$

$$p(x_4|x_5, x_6, x_7)$$

$$p(x_5|x_6, x_7)$$

$$p(x_6|x_7)$$

$$p(x_7)$$

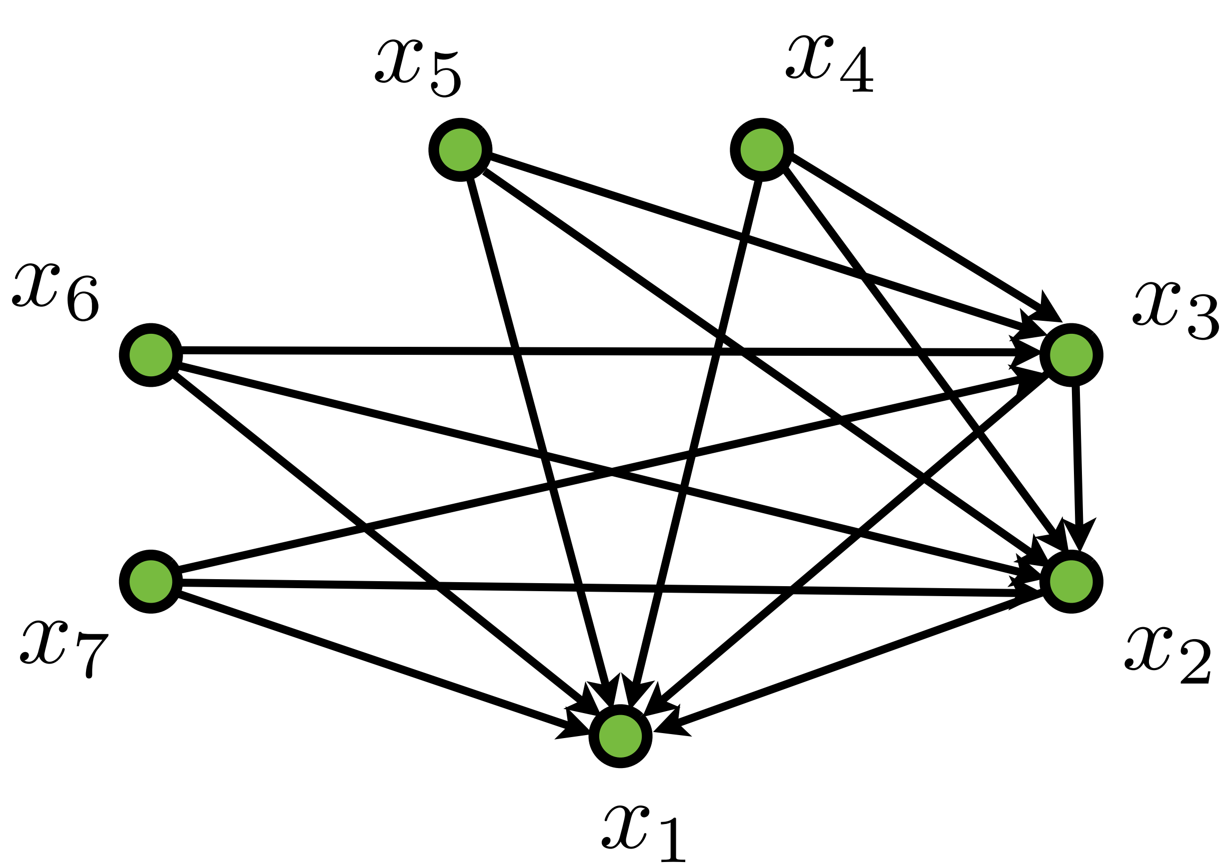


BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2, x_3, x_4, x_5, x_6, x_7)$$

$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2|x_3, x_4, x_5, x_6, x_7)p(x_3, x_4, x_5, x_6, x_7)$$

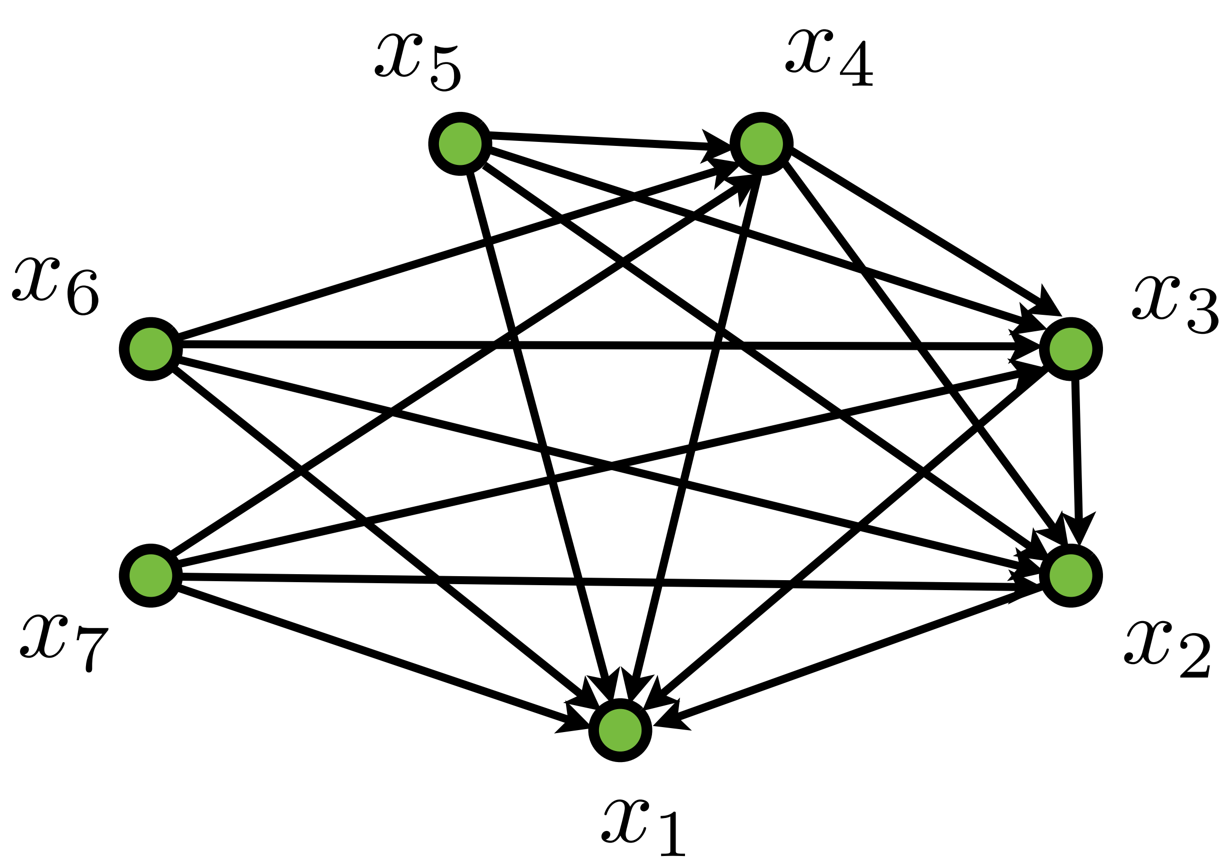
$$\begin{aligned} = \; & p(x_1|x_2, x_3, x_4, x_5, x_6, x_7) \\ & p(x_2|x_3, x_4, x_5, x_6, x_7) \\ & p(x_3|x_4, x_5, x_6, x_7) \\ & p(x_4|x_5, x_6, x_7) \\ & p(x_5|x_6, x_7) \\ & p(x_6|x_7) \\ & p(x_7) \end{aligned}$$



BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2, x_3, x_4, x_5, x_6, x_7)$$

$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2|x_3, x_4, x_5, x_6, x_7)p(x_3, x_4, x_5, x_6, x_7)$$



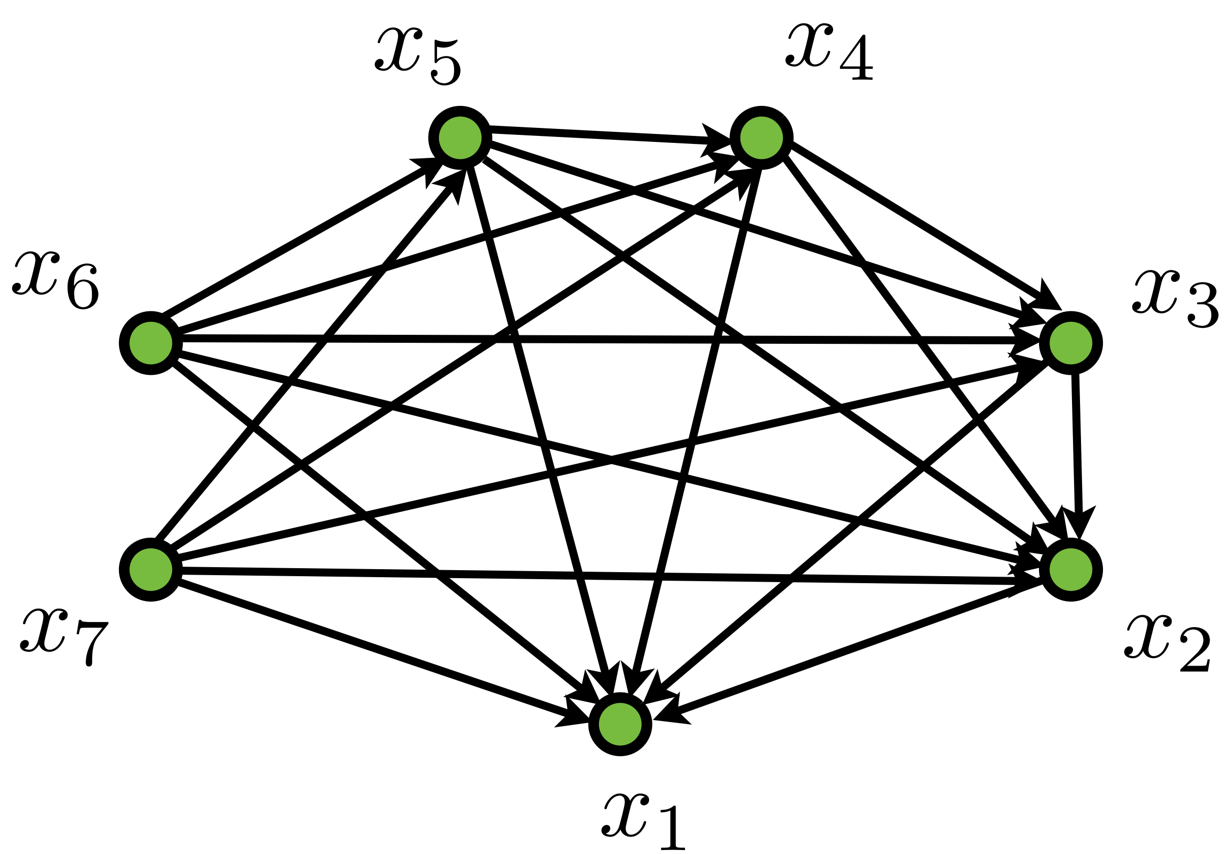
$$\begin{aligned} = \ & p(x_1|x_2, x_3, x_4, x_5, x_6, x_7) \\ & p(x_2|x_3, x_4, x_5, x_6, x_7) \\ & p(x_3|x_4, x_5, x_6, x_7) \\ & p(x_4|x_5, x_6, x_7) \\ & p(x_5|x_6, x_7) \\ & p(x_6|x_7) \\ & p(x_7) \end{aligned}$$

BAYESIAN NETWORKS

$$p(x_1,x_2,x_3,x_4,x_5,x_6,x_7) = p(x_1|x_2,x_3,x_4,x_5,x_6,x_7)p(x_2,x_3,x_4,x_5,x_6,x_7)$$

$$= p(x_1|x_2,x_3,x_4,x_5,x_6,x_7)p(x_2|x_3,x_4,x_5,x_6,x_7)p(x_3,x_4,x_5,x_6,x_7)$$

$$= p(x_1|x_2,x_3,x_4,x_5,x_6,x_7)$$
$$p(x_2|x_3,x_4,x_5,x_6,x_7)$$
$$p(x_3|x_4,x_5,x_6,x_7)$$
$$p(x_4|x_5,x_6,x_7)$$
$$p(x_5|x_6,x_7)$$
$$p(x_6|x_7)$$
$$p(x_7)$$

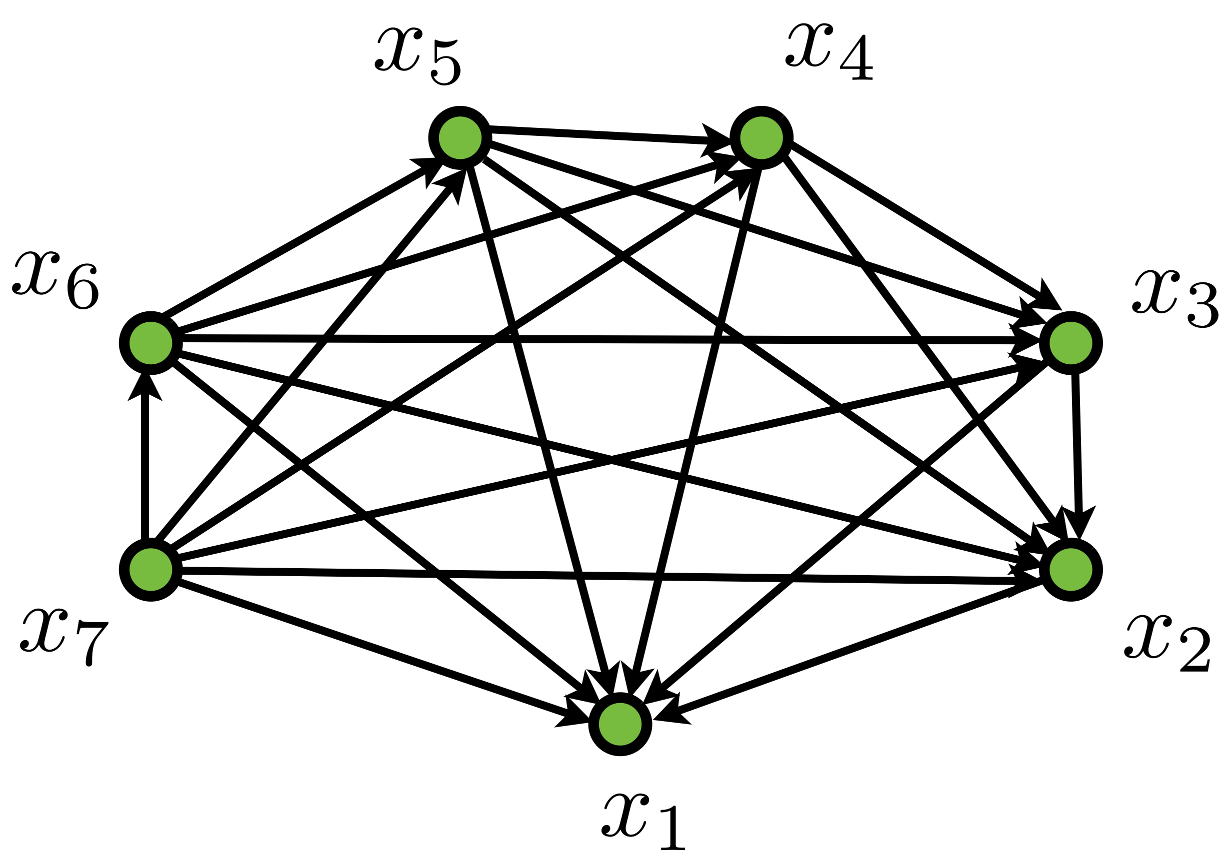


BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2, x_3, x_4, x_5, x_6, x_7)$$

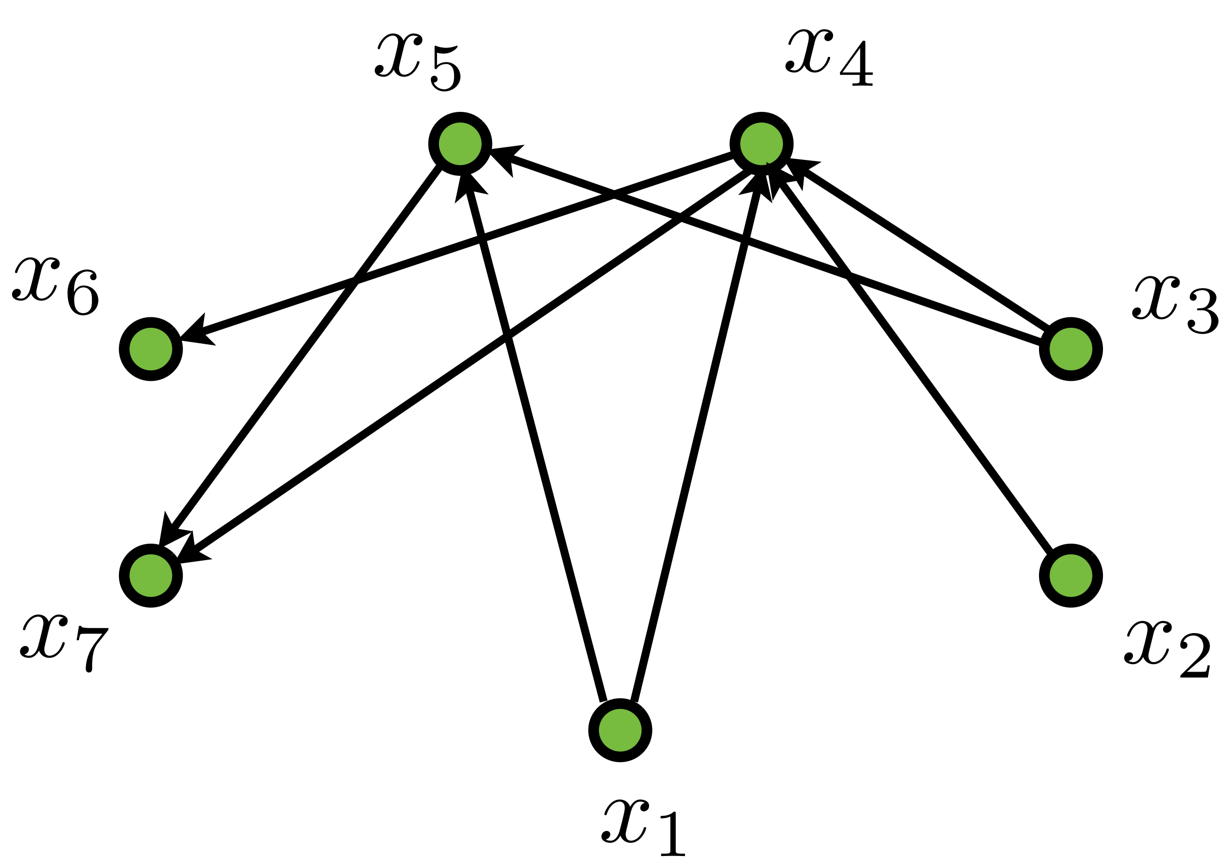
$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)p(x_2|x_3, x_4, x_5, x_6, x_7)p(x_3, x_4, x_5, x_6, x_7)$$

$$= p(x_1|x_2, x_3, x_4, x_5, x_6, x_7)$$
$$p(x_2|x_3, x_4, x_5, x_6, x_7)$$
$$p(x_3|x_4, x_5, x_6, x_7)$$
$$p(x_4|x_5, x_6, x_7)$$
$$p(x_5|x_6, x_7)$$
$$p(x_6|x_7)$$
$$p(x_7)$$



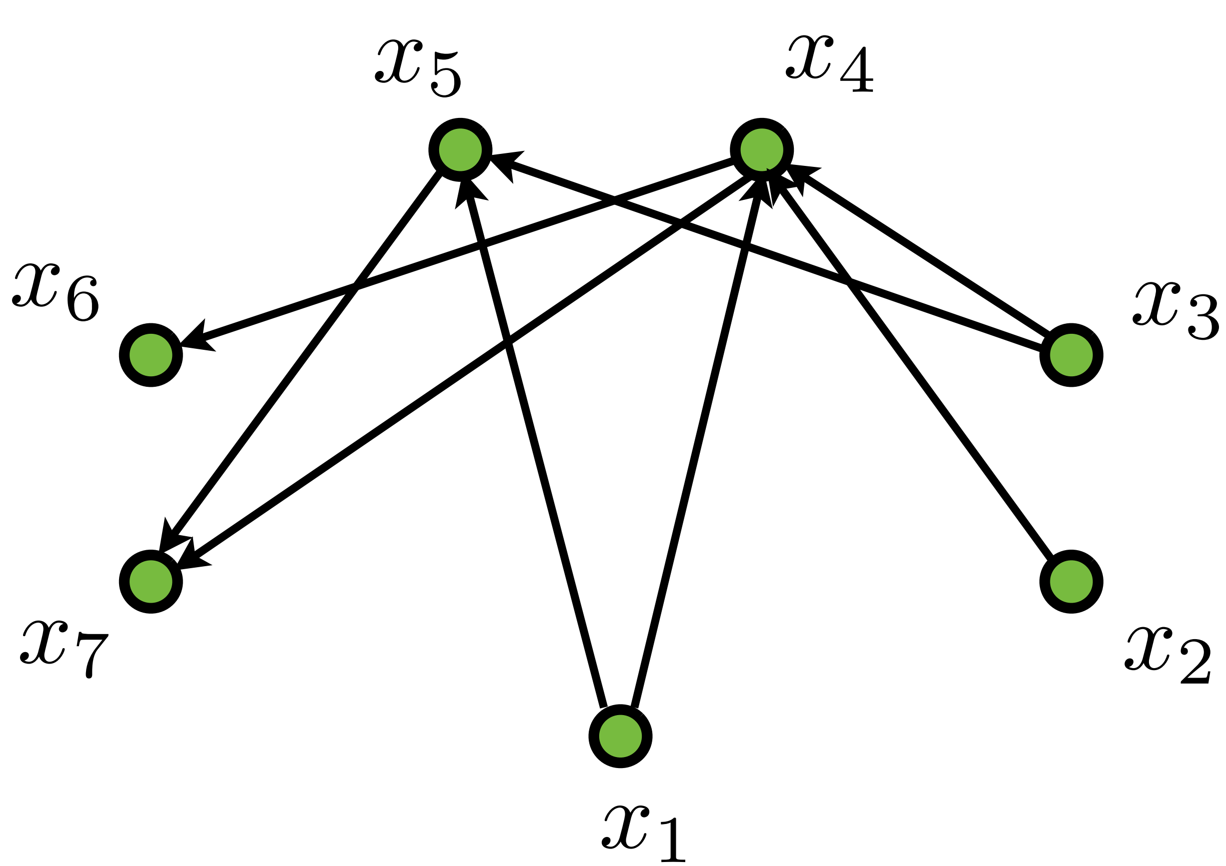
BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) =$$



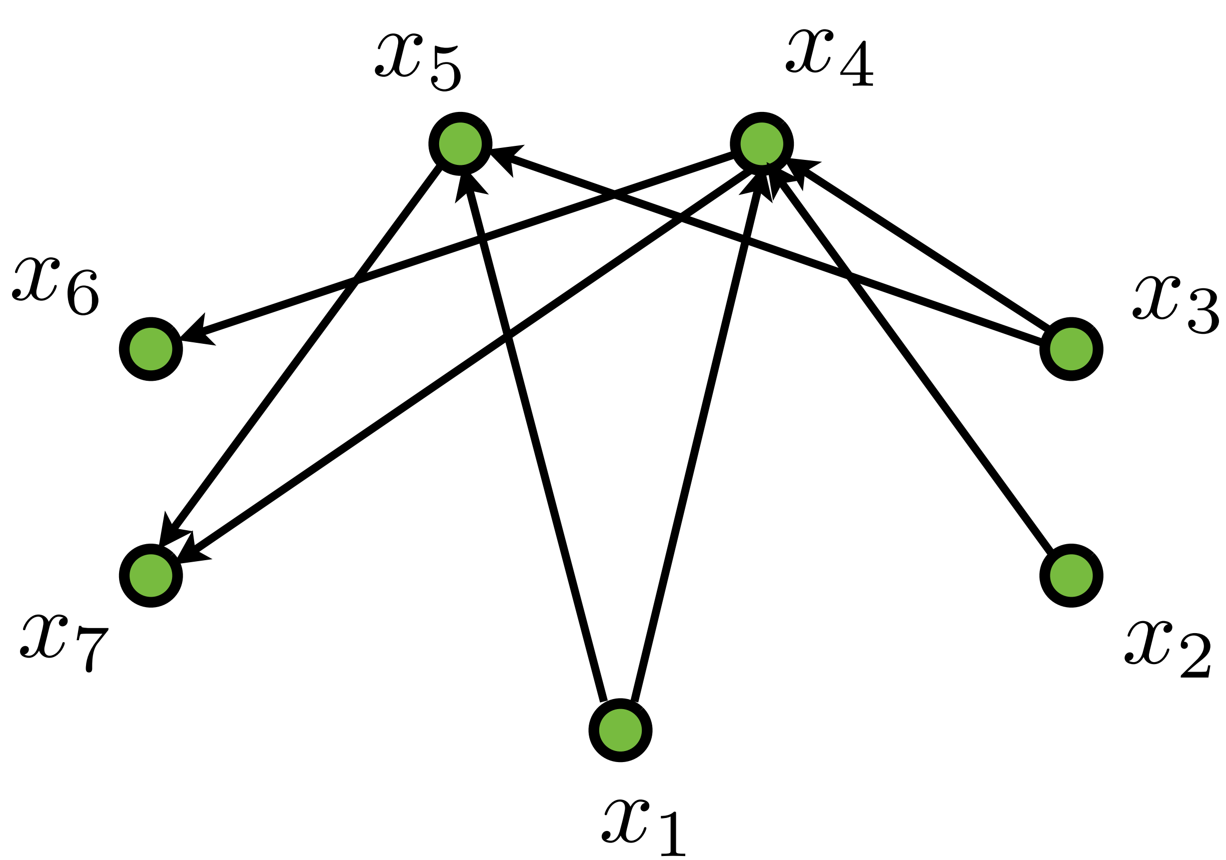
BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = \; p(x_1)$$



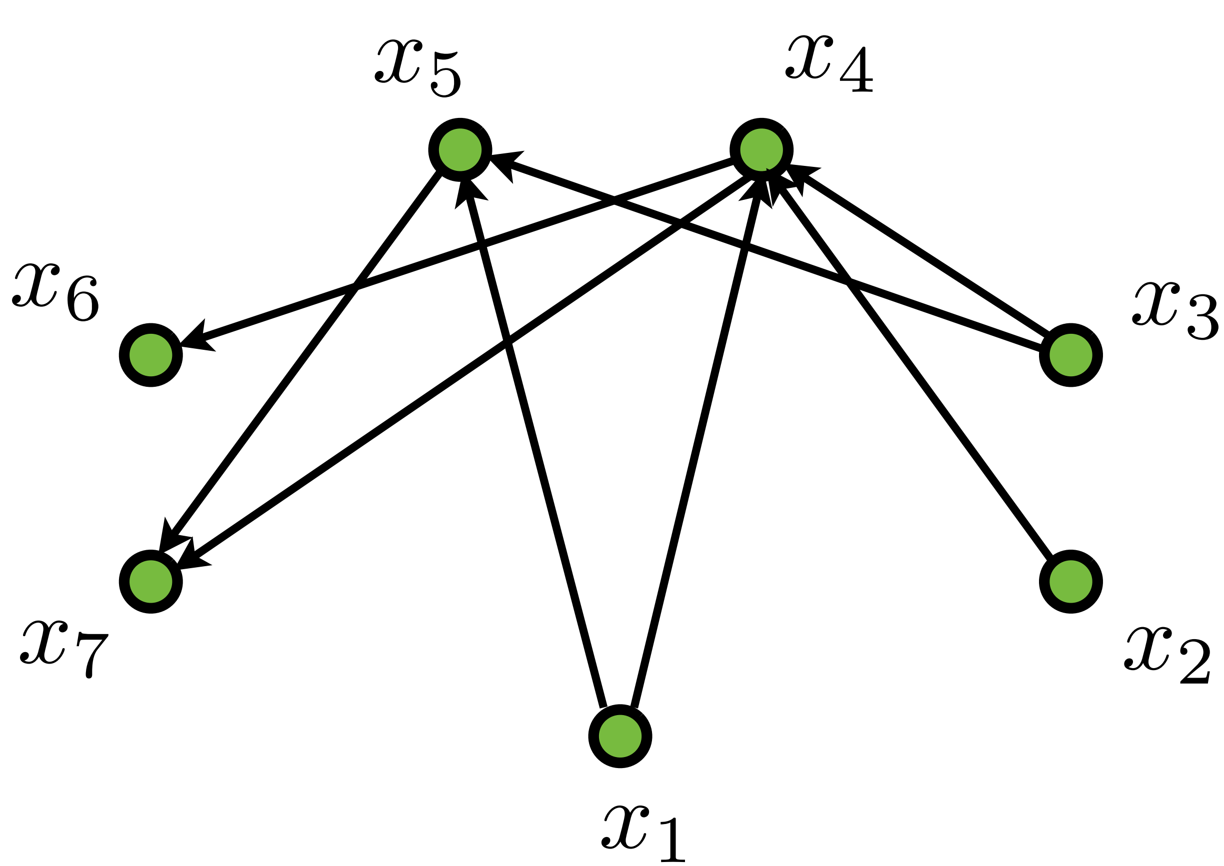
BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = \begin{array}{r} p(x_1) \\ p(x_2) \end{array}$$



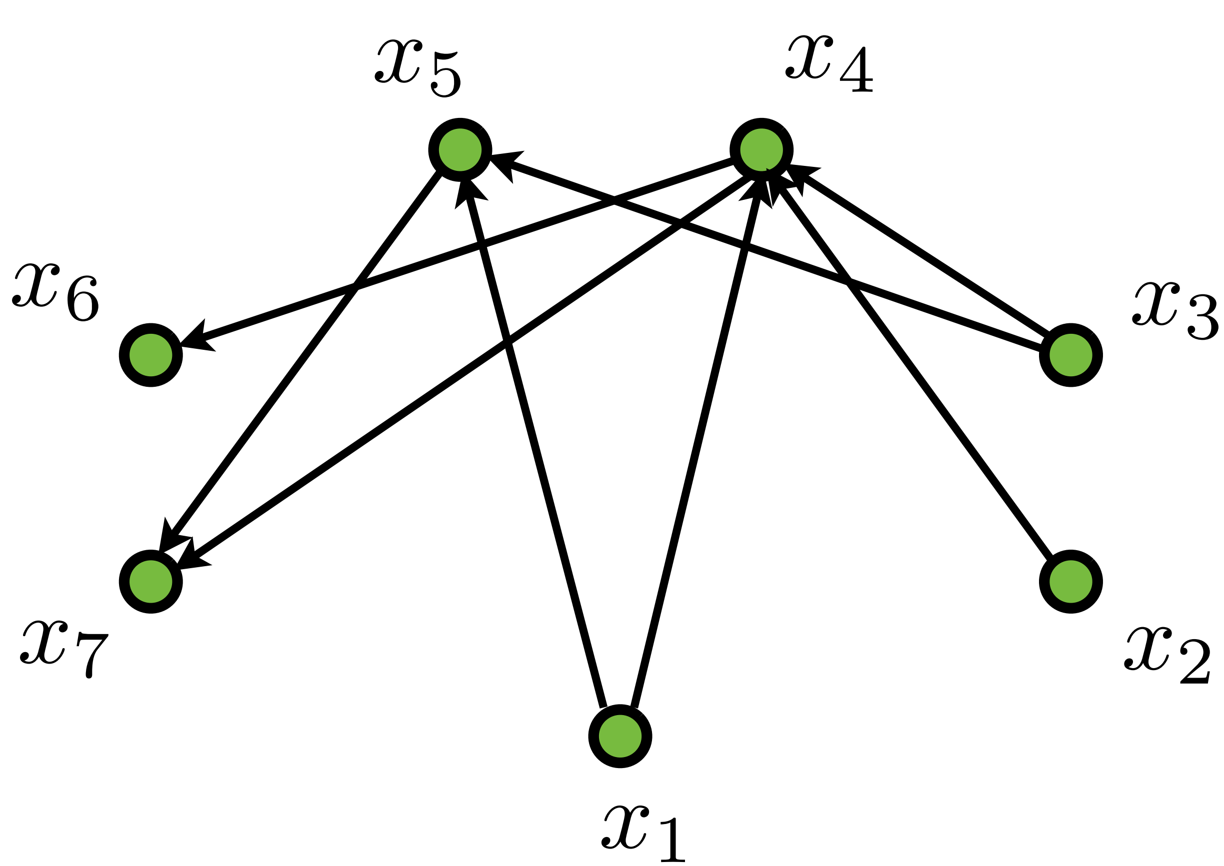
BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = \begin{array}{l} p(x_1) \\ p(x_2) \\ p(x_3) \end{array}$$



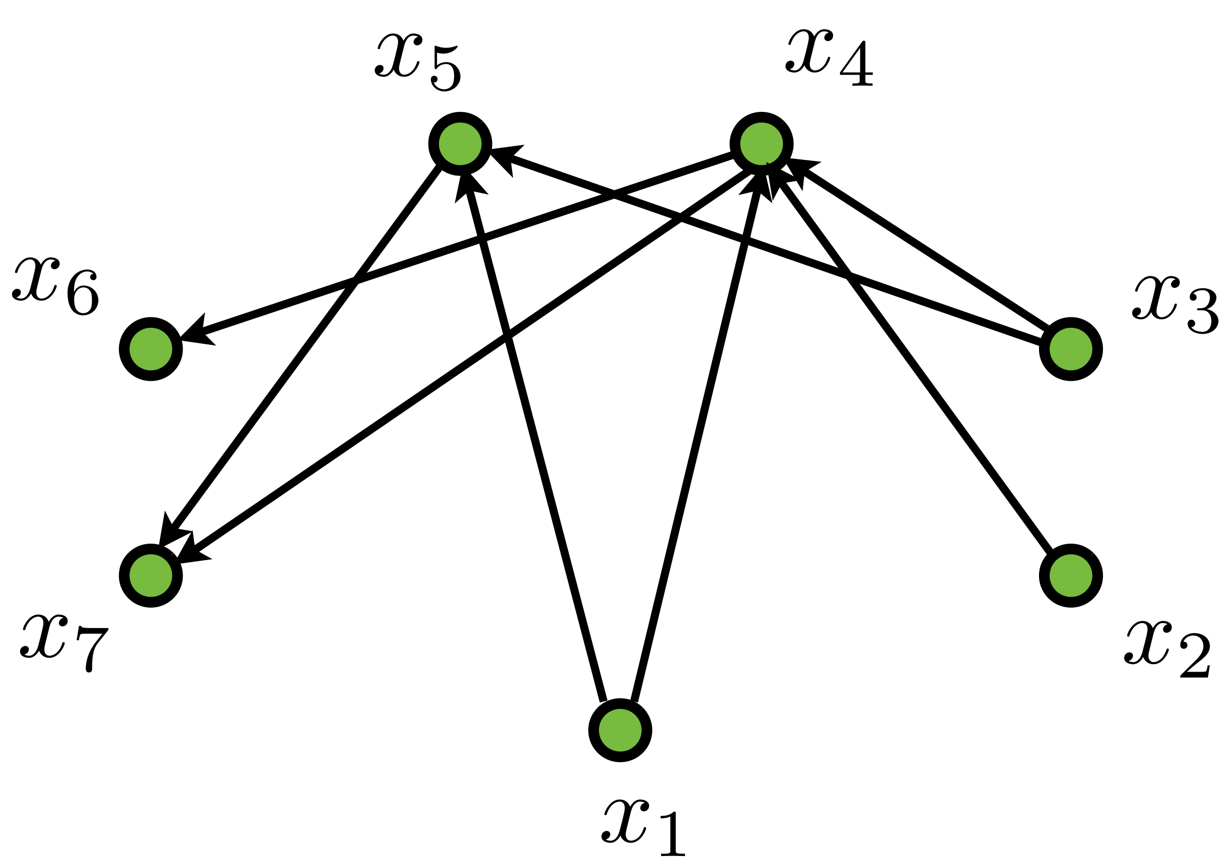
BAYESIAN NETWORKS

$$
\begin{aligned}
p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = \ & p(x_1) \\
& p(x_2) \\
& p(x_3) \\
& p(x_4 | x_1, x_2, x_3)
\end{aligned}
$$



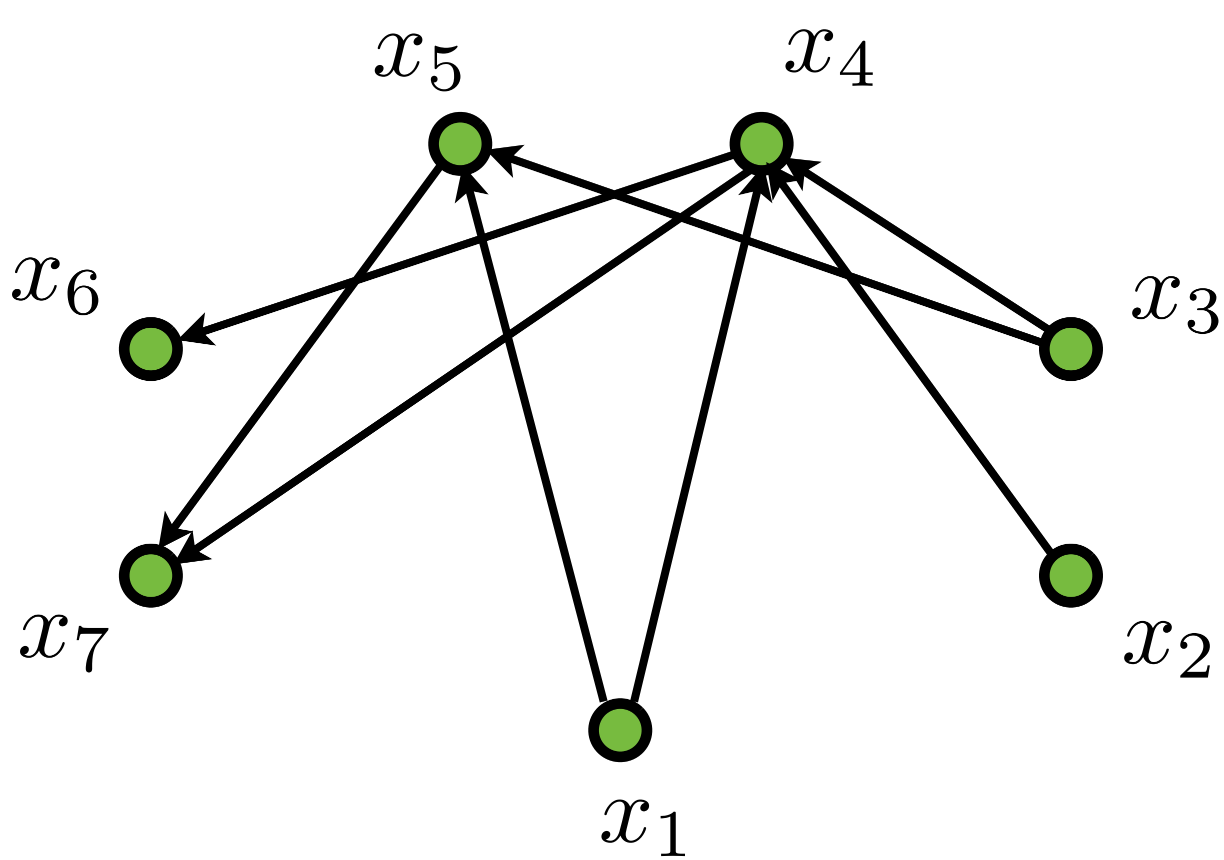
BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = \; p(x_1)$$
$$p(x_2)$$
$$p(x_3)$$
$$p(x_4|x_1, x_2, x_3)$$
$$p(x_5|x_1, x_3)$$



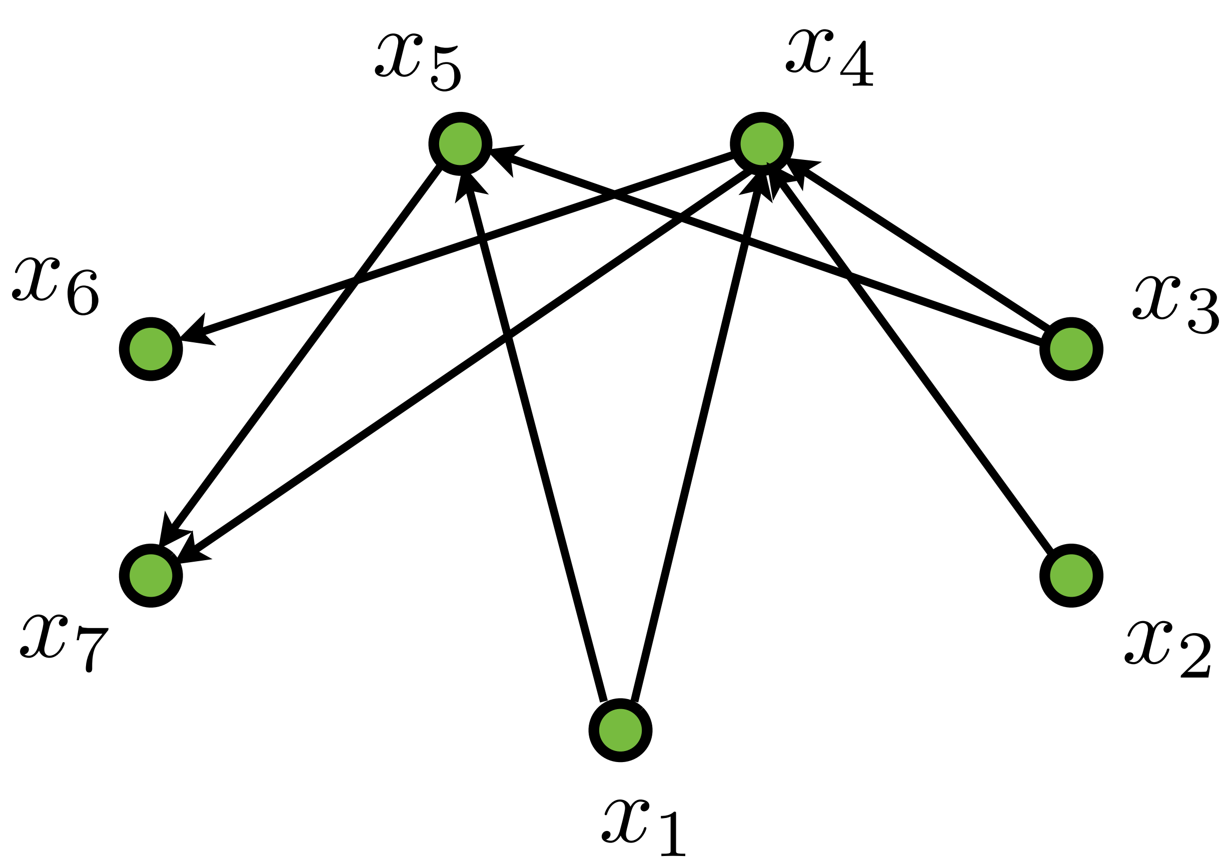
BAYESIAN NETWORKS

$$
\begin{aligned}
p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = \ & p(x_1) \\
& p(x_2) \\
& p(x_3) \\
& p(x_4 | x_1, x_2, x_3) \\
& p(x_5 | x_1, x_3) \\
& p(x_6 | x_4)
\end{aligned}
$$



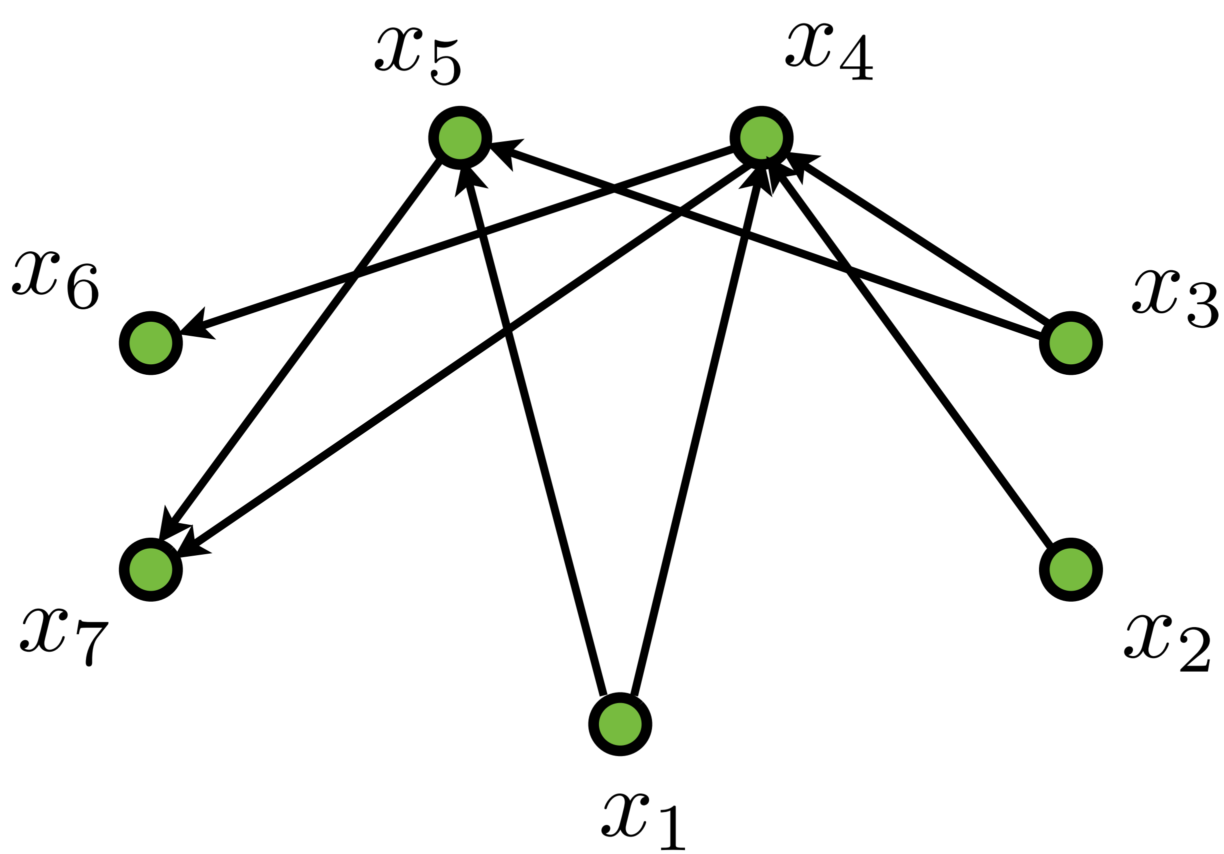
BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = \begin{array}{r} p(x_1) \\ p(x_2) \\ p(x_3) \\ p(x_4|x_1, x_2, x_3) \\ p(x_5|x_1, x_3) \\ p(x_6|x_4) \\ p(x_7|x_4, x_5) \end{array}$$



BAYESIAN NETWORKS

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = p(x_1)\, p(x_2)\, p(x_3)\, p(x_4|x_1, x_2, x_3)\, p(x_5|x_1, x_3)\, p(x_6|x_4)\, p(x_7|x_4, x_5)$$



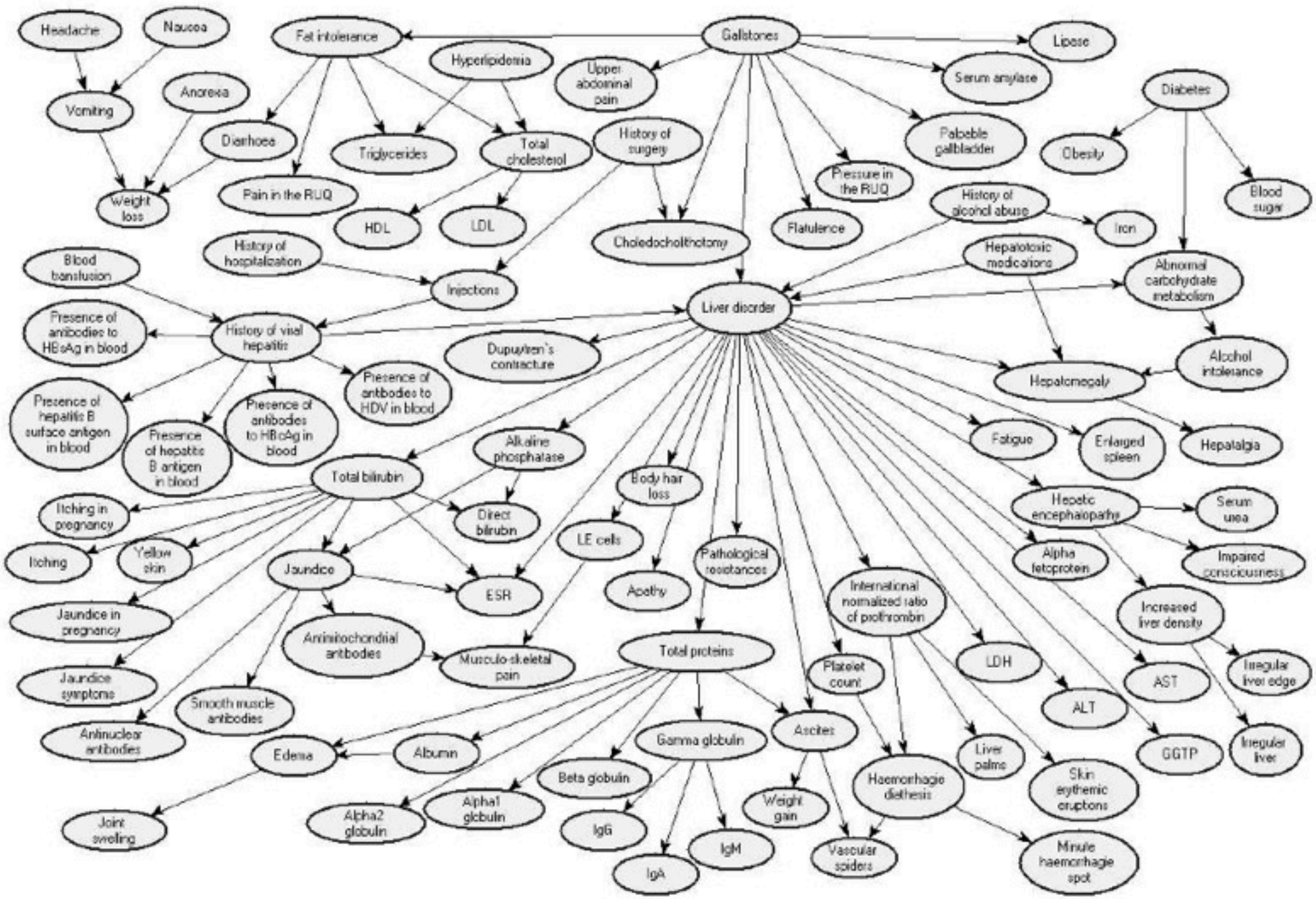
parents of *x*

$$p(\mathbf{x}) = \prod_{k=1}^{K} p(x_k|\mathrm{pa}_k)$$

BAYESIAN NETWORKS

- BAYESIAN NETWORKS ARE...
  - A REPRESENTATION OF JOINT PROBABILITY
  - A COLLECTION OF CONDITIONAL INDEPENDENCE STATEMENTS
  - EXAMPLES?

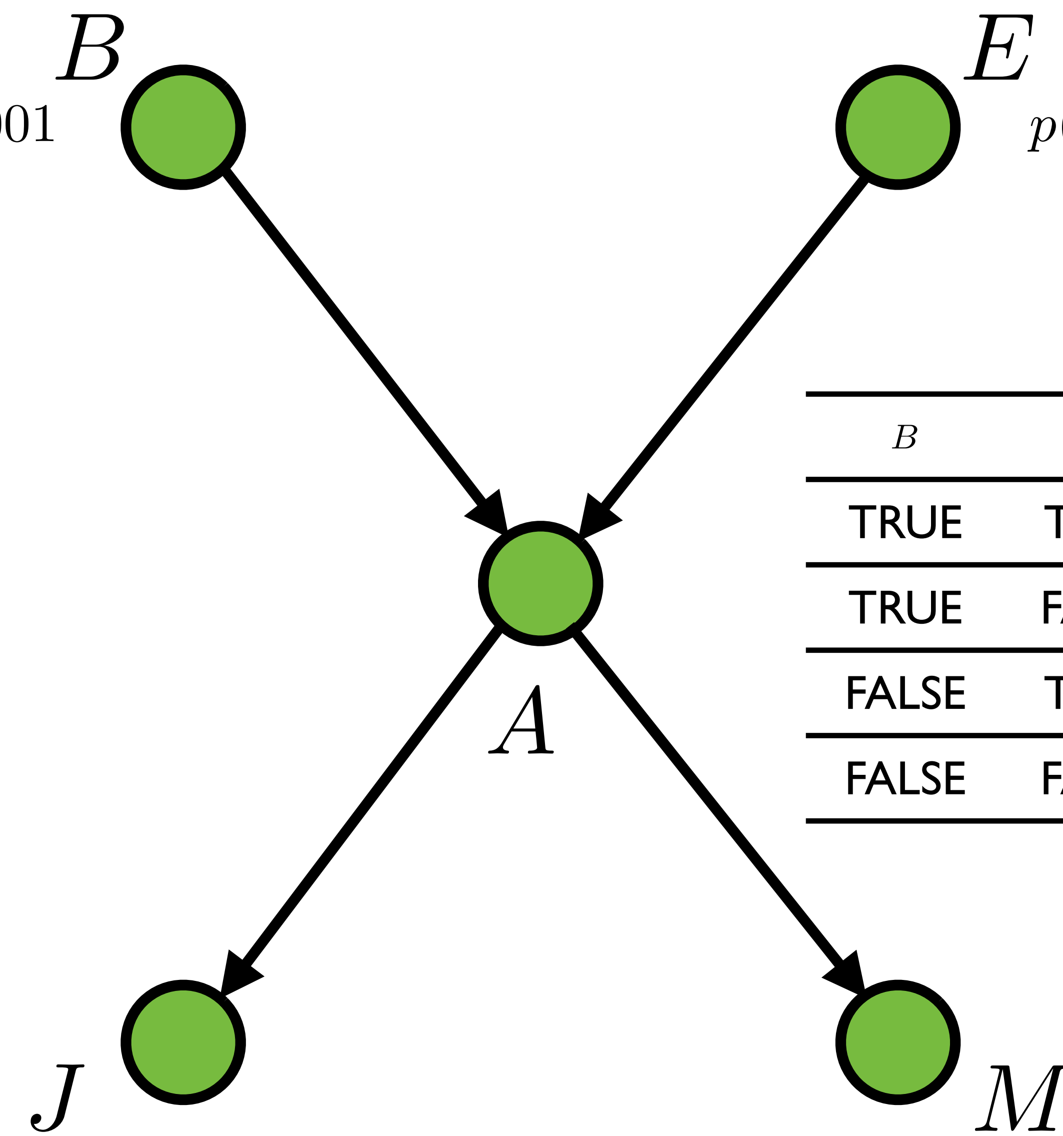
Diagnosis of Liver Disorders



Agnieszka Onisko, Marek J. Druzdzel and Hanna Wasyluk. [A Bayesian network model for diagnosis of liver disorders](). In *Proceedings of the Eleventh Conference on Biocybernetics and Biomedical Engineering*, pages 842-846, Warsaw, Poland, December 2-4, 1999.

$p(B = \texttt{TRUE}) = 0.001$

$p(E = \texttt{TRUE}) = 0.002$

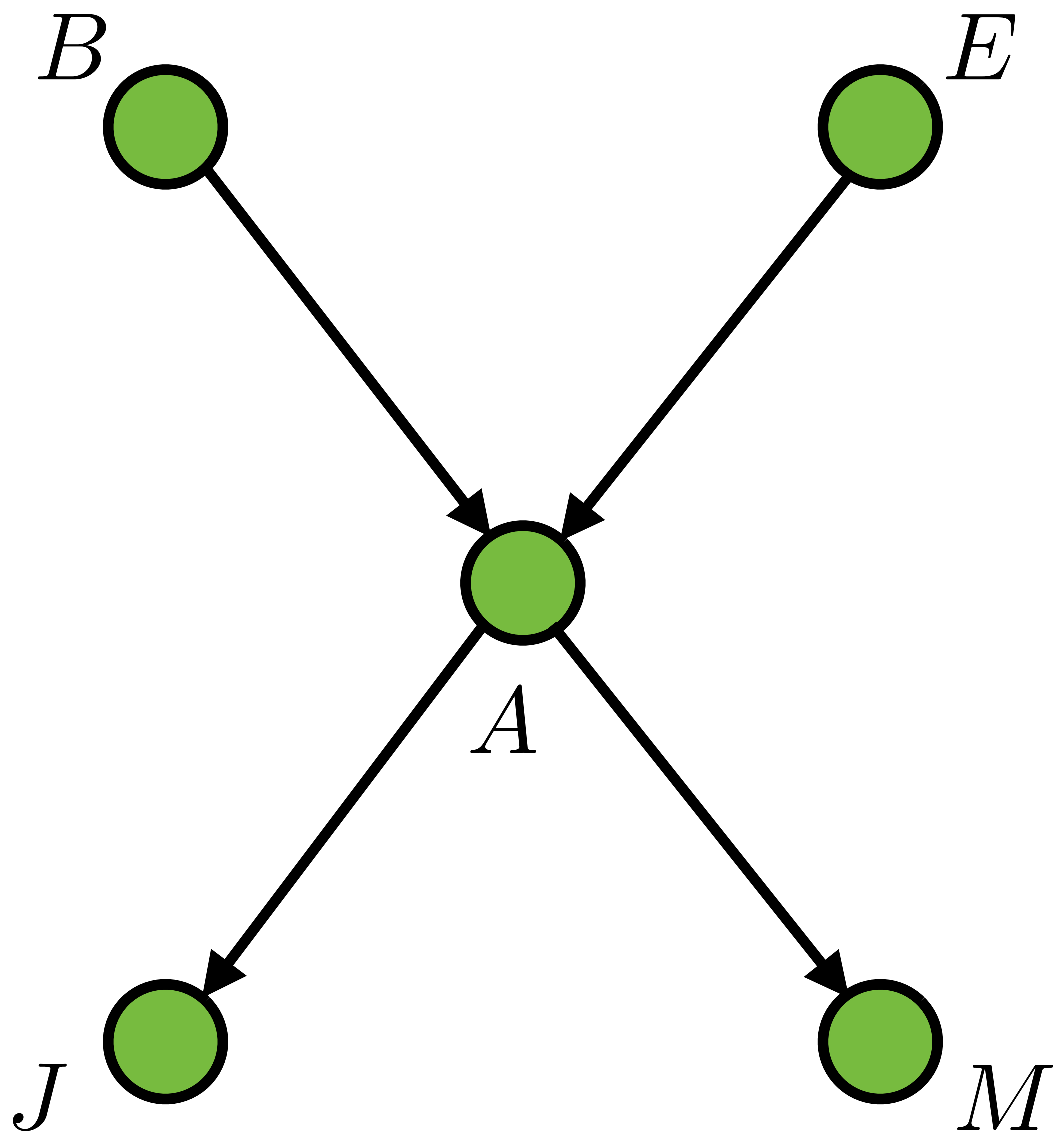


B	E	$p(A\|B,E)$
TRUE	TRUE	0.95
TRUE	FALSE	0.95
FALSE	TRUE	0.29
FALSE	FALSE	0.001

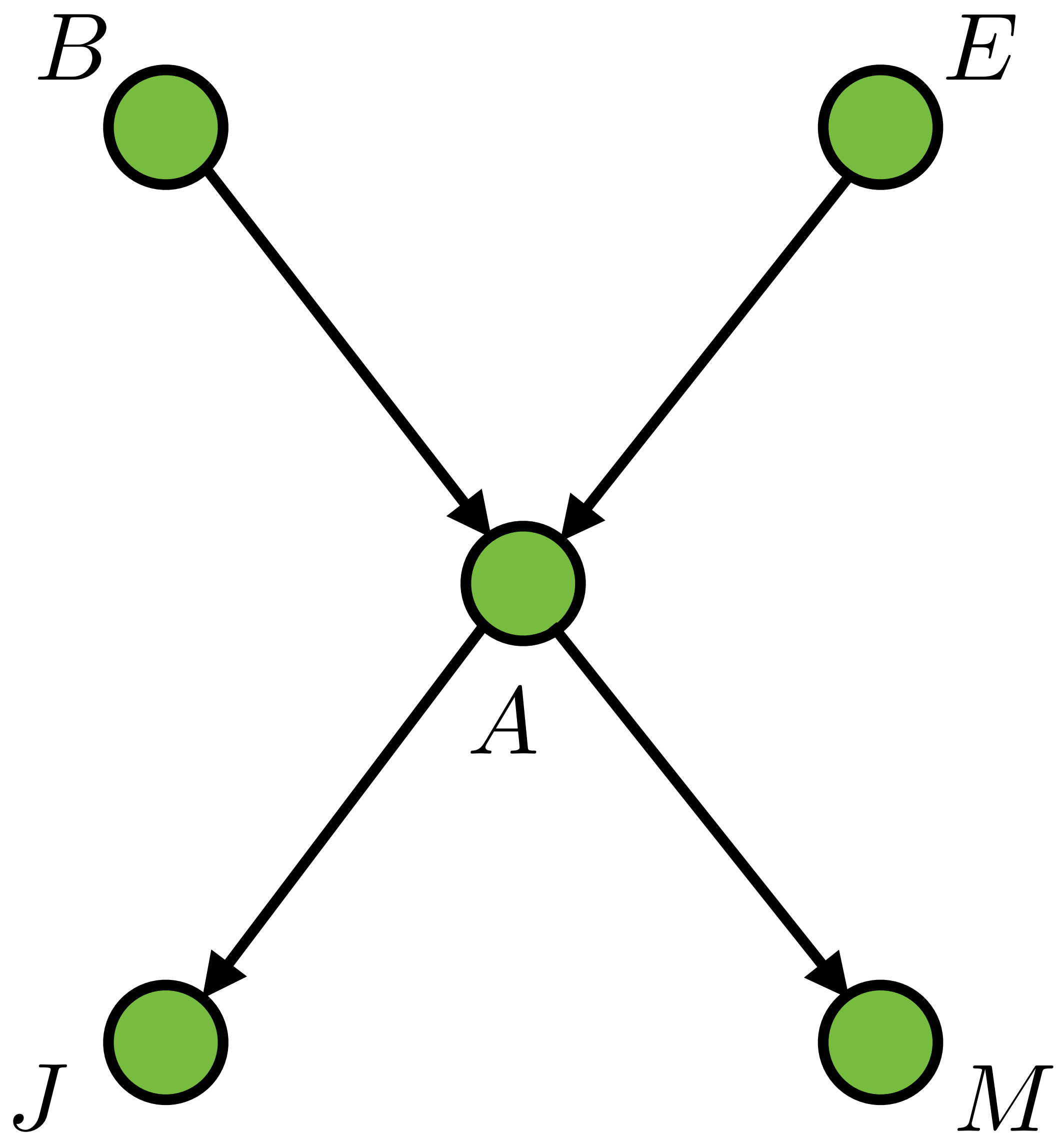
A	$p(J\|A)$
TRUE	0.90
FALSE	0.05

A	$p(M\|A)$
TRUE	0.70
FALSE	0.01

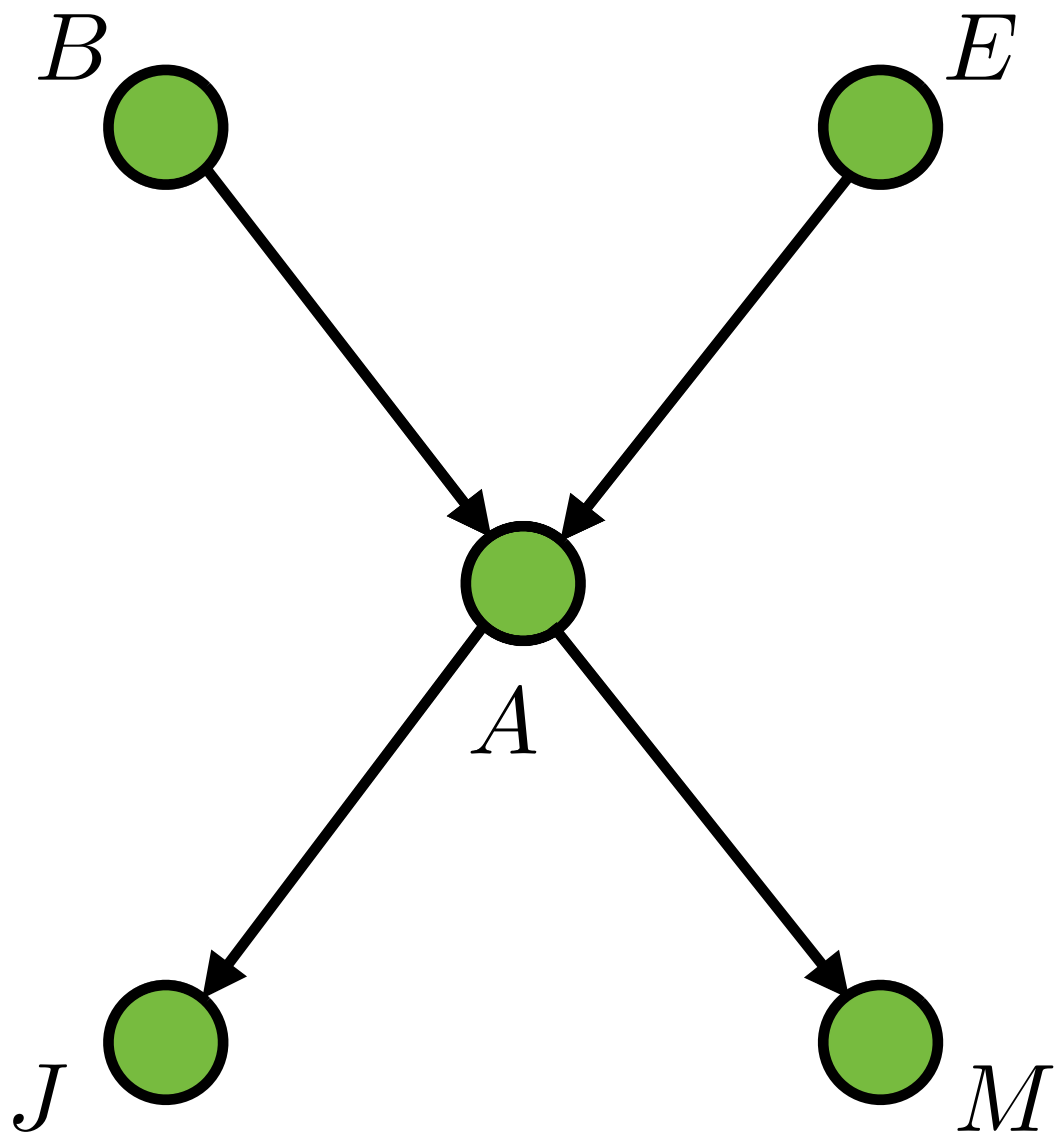
$$p(B, E, A, J, M) = p(J|A)$$



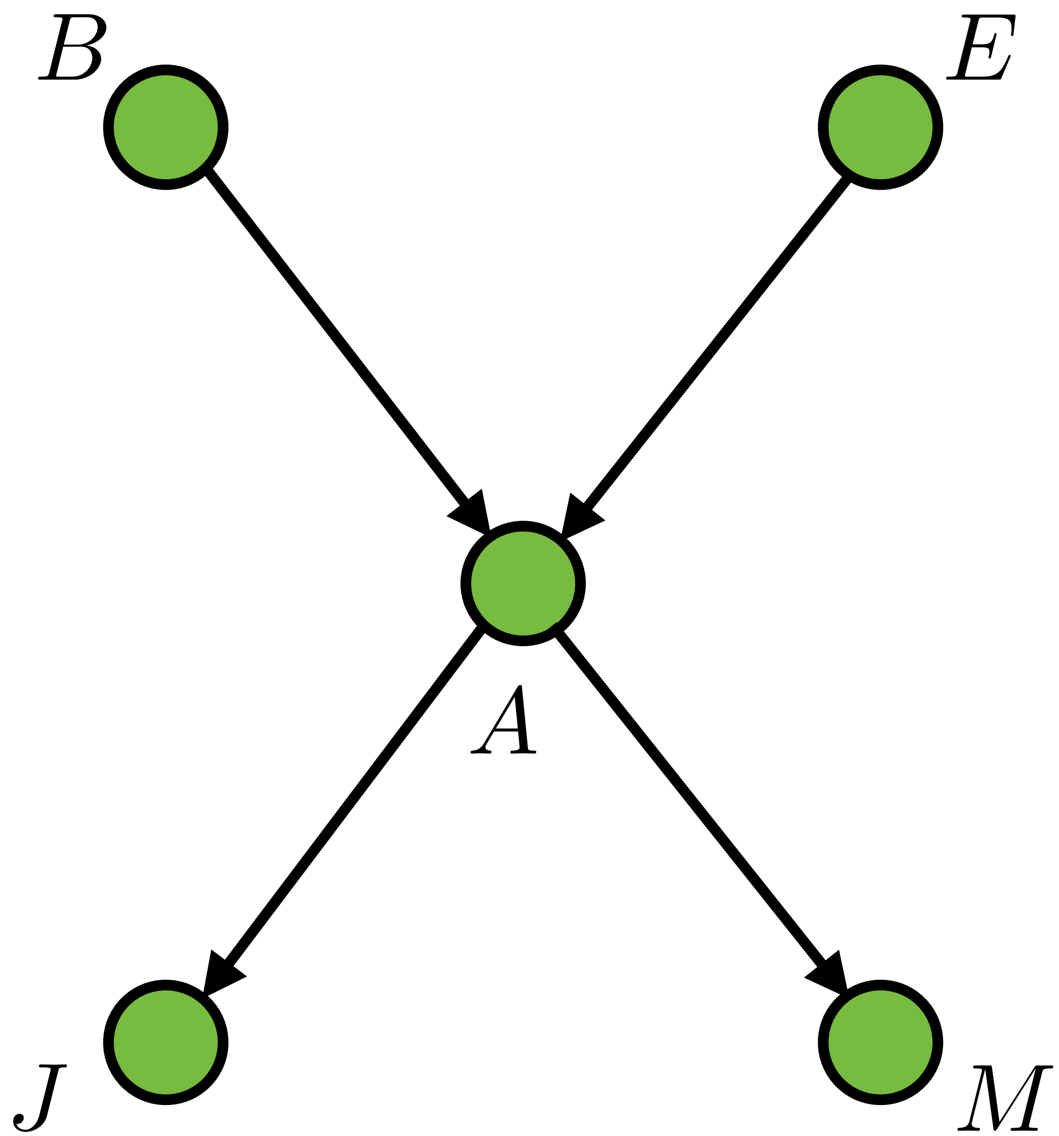
$$p(B, E, A, J, M) = p(J|A)p(M|A)$$



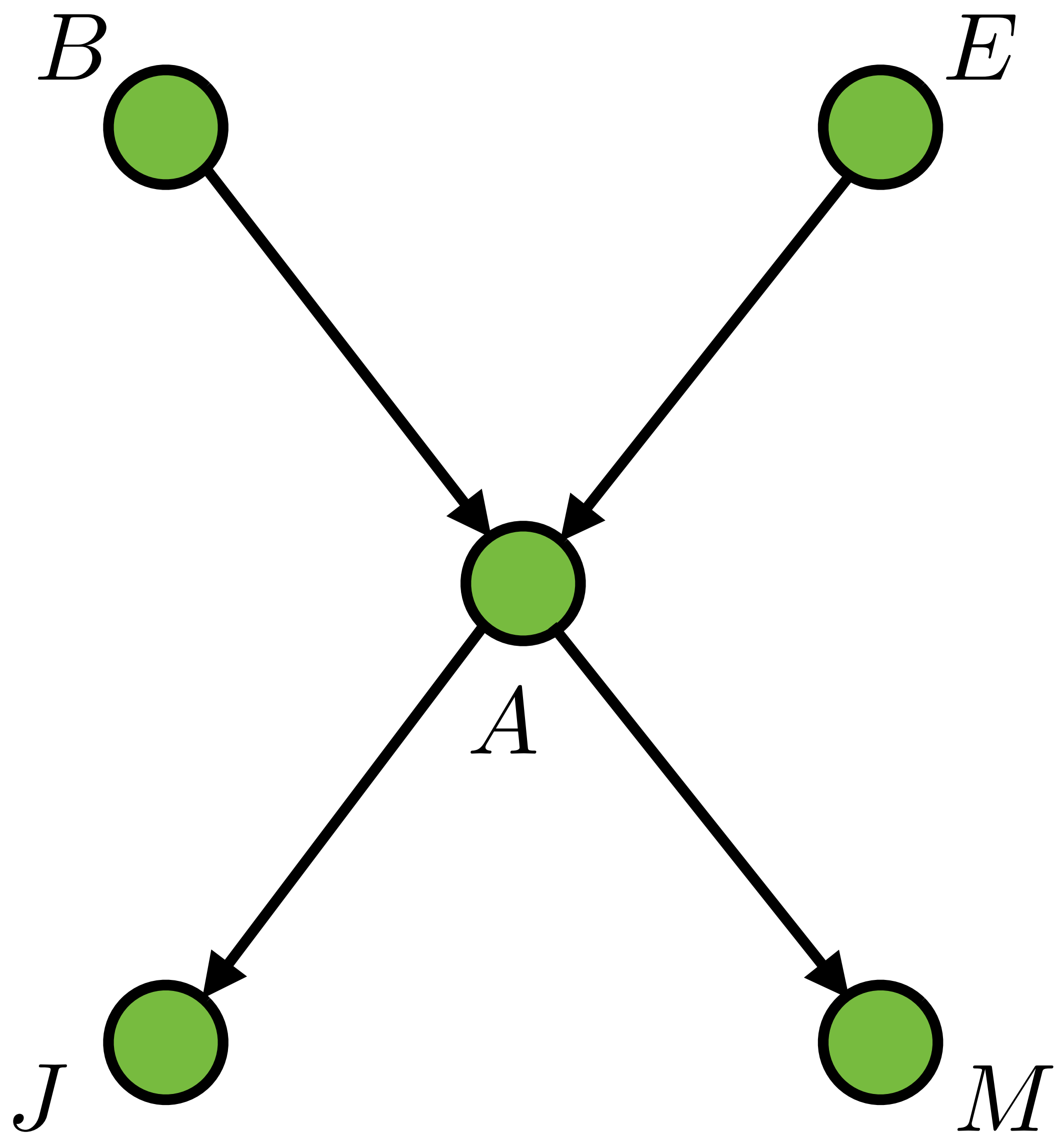
$$p(B, E, A, J, M) = p(J|A)p(M|A)p(A|B, E)$$



$$p(B, E, A, J, M) = p(J|A)p(M|A)p(A|B, E)p(B)$$



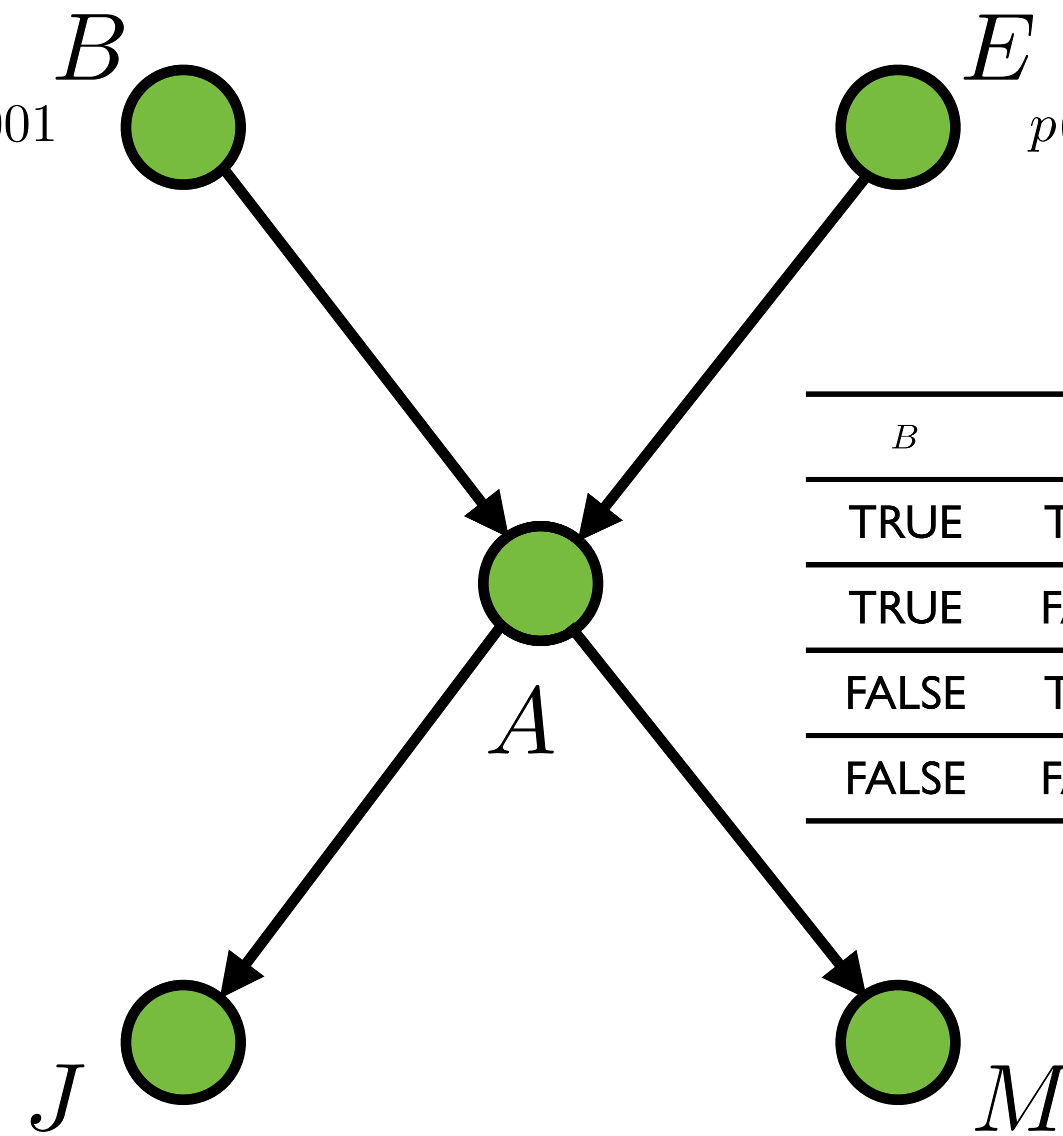
$$p(B, E, A, J, M) = p(J|A)p(M|A)p(A|B, E)p(B)p(E)$$



$$p(B, E, A, J, M) = p(J|A)p(M|A)p(A|B, E)p(B)p(E)$$

$p(B = \texttt{TRUE}) = 0.001$

$p(E = \texttt{TRUE}) = 0.002$



B	E	$p(A\|B,E)$
TRUE	TRUE	0.95
TRUE	FALSE	0.95
FALSE	TRUE	0.29
FALSE	FALSE	0.001

A	$p(J\|A)$
TRUE	0.90
FALSE	0.05

A	$p(M\|A)$
TRUE	0.70
FALSE	0.01

$$p(B, E, A, J, M) = p(J|A)p(M|A)p(A|B, E)p(B)p(E)$$

$$p(B, E, A, J, M) = p(J|A)p(M|A)p(A|B, E)p(B)p(E)$$

- WHAT IS THE PROBABILITY THAT:
  - ALARM SOUNDED AND
  - NO BURGLARY AND
  - NO EARTHQUAKE AND
  - MARY CALLS AND
  - JOHN CALLS?

$$p(B, E, A, J, M) = p(J|A)p(M|A)p(A|B, E)p(B)p(E)$$

- WHAT IS THE PROBABILITY THAT:
  - ALARM SOUNDED AND
  - NO BURGLARY AND
  - NO EARTHQUAKE AND
  - MARY CALLS AND
  - JOHN CALLS?

$$p(B = \texttt{FALSE}, E = \texttt{FALSE}, A = \texttt{TRUE}, J = \texttt{TRUE}, M = \texttt{TRUE})$$

$$p(B, E, A, J, M) = p(J|A)p(M|A)p(A|B, E)p(B)p(E)$$

- WHAT IS THE PROBABILITY THAT:
  - ALARM SOUNDED AND
  - NO BURGLARY AND
  - NO EARTHQUAKE AND
  - MARY CALLS AND
  - JOHN CALLS?

$$\begin{aligned} p(B = \texttt{FALSE}, E = \texttt{FALSE}, A = \texttt{TRUE}, J = \texttt{TRUE}, M = \texttt{TRUE}) \\ = \; p(J = \texttt{TRUE}|A = \texttt{TRUE}) \end{aligned}$$

$$p(B, E, A, J, M) = p(J|A)p(M|A)p(A|B, E)p(B)p(E)$$

- WHAT IS THE PROBABILITY THAT:
  - ALARM SOUNDED AND
  - NO BURGLARY AND
  - NO EARTHQUAKE AND
  - MARY CALLS AND
  - JOHN CALLS?

$$\begin{aligned} p(B = \texttt{FALSE}, E = \texttt{FALSE}, A = \texttt{TRUE}, J = \texttt{TRUE}, M = \texttt{TRUE}) \\ = \ p(J = \texttt{TRUE}|A = \texttt{TRUE}) \\ p(M = \texttt{TRUE}|A = \texttt{TRUE}) \end{aligned}$$

$$p(B,E,A,J,M) = p(J|A)p(M|A)p(A|B,E)p(B)p(E)$$

- WHAT IS THE PROBABILITY THAT:
  - ALARM SOUNDED AND
  - NO BURGLARY AND
  - NO EARTHQUAKE AND
  - MARY CALLS AND
  - JOHN CALLS?

$$\begin{aligned} p(B = \texttt{FALSE}, E = \texttt{FALSE}, A = \texttt{TRUE}, J = \texttt{TRUE}, M = \texttt{TRUE}) \\ = p(J = \texttt{TRUE}|A = \texttt{TRUE}) \\ p(M = \texttt{TRUE}|A = \texttt{TRUE}) \\ p(A = \texttt{TRUE}|B = \texttt{FALSE}, E = \texttt{FALSE}) \end{aligned}$$

$$p(B,E,A,J,M) = p(J|A)p(M|A)p(A|B,E)p(B)p(E)$$

- WHAT IS THE PROBABILITY THAT:
  - ALARM SOUNDED AND
  - NO BURGLARY AND
  - NO EARTHQUAKE AND
  - MARY CALLS AND
  - JOHN CALLS?

$$\begin{aligned}
p(B = \texttt{FALSE}, E = \texttt{FALSE}, A = \texttt{TRUE}, J = \texttt{TRUE}, M = \texttt{TRUE}) \\
= \; p(J = \texttt{TRUE}|A = \texttt{TRUE}) \\
p(M = \texttt{TRUE}|A = \texttt{TRUE}) \\
p(A = \texttt{TRUE}|B = \texttt{FALSE}, E = \texttt{FALSE}) \\
p(B = \texttt{FALSE})
\end{aligned}$$

$$p(B, E, A, J, M) = p(J|A)p(M|A)p(A|B, E)p(B)p(E)$$

- WHAT IS THE PROBABILITY THAT:
  - ALARM SOUNDED AND
  - NO BURGLARY AND
  - NO EARTHQUAKE AND
  - MARY CALLS AND
  - JOHN CALLS?

$$\begin{aligned} p(B = \texttt{FALSE}, E = \texttt{FALSE}, A = \texttt{TRUE}, J = \texttt{TRUE}, M = \texttt{TRUE}) \\ = p(J = \texttt{TRUE}|A = \texttt{TRUE}) \\ p(M = \texttt{TRUE}|A = \texttt{TRUE}) \\ p(A = \texttt{TRUE}|B = \texttt{FALSE}, E = \texttt{FALSE}) \\ p(B = \texttt{FALSE}) \\ p(E = \texttt{FALSE}) \end{aligned}$$

$$p(B, E, A, J, M) = p(J|A)p(M|A)p(A|B, E)p(B)p(E)$$

- WHAT IS THE PROBABILITY THAT:
  - ALARM SOUNDED AND
  - NO BURGLARY AND
  - NO EARTHQUAKE AND
  - MARY CALLS AND
  - JOHN CALLS?

$$\begin{aligned} p(B = \texttt{FALSE}, E = \texttt{FALSE}, A = \texttt{TRUE}, J = \texttt{TRUE}, M = \texttt{TRUE}) & \\ = \; p(J = \texttt{TRUE}|A = \texttt{TRUE}) & \qquad 0.90 \\ p(M = \texttt{TRUE}|A = \texttt{TRUE}) & \qquad 0.70 \\ p(A = \texttt{TRUE}|B = \texttt{FALSE}, E = \texttt{FALSE}) & \qquad 0.001 \\ p(B = \texttt{FALSE}) & \qquad 0.999 \\ p(E = \texttt{FALSE}) & \qquad 0.998 \end{aligned}$$

$$p(B, E, A, J, M) = p(J|A)p(M|A)p(A|B, E)p(B)p(E)$$

- WHAT IS THE PROBABILITY THAT:
  - ALARM SOUNDED AND
  - NO BURGLARY AND
  - NO EARTHQUAKE AND
  - MARY CALLS AND
  - JOHN CALLS?

$$\begin{aligned} p(B = \texttt{FALSE}, E = \texttt{FALSE}, A = \texttt{TRUE}, J = \texttt{TRUE}, M = \texttt{TRUE}) & \\ = p(J = \texttt{TRUE}|A = \texttt{TRUE}) & \quad 0.90 \\ p(M = \texttt{TRUE}|A = \texttt{TRUE}) & \quad 0.70 \\ p(A = \texttt{TRUE}|B = \texttt{FALSE}, E = \texttt{FALSE}) & \quad 0.001 \\ p(B = \texttt{FALSE}) & \quad 0.999 \\ p(E = \texttt{FALSE}) & \quad \times\ 0.998 \\ & \quad \overline{0.00062} \end{aligned}$$

WILL I GET A CALL IF THERE IS A BURGLARY?

NETWORK CONSTRUCTION

- CHOOSE RANDOM VARIABLES X_i THAT DESCRIBE THE DOMAIN

NETWORK CONSTRUCTION

- CHOOSE RANDOM VARIABLES X_i THAT DESCRIBE THE DOMAIN
- CHOOSE AN ORDERING FOR THE VARIABLES

NETWORK CONSTRUCTION

- CHOOSE RANDOM VARIABLES X_i THAT DESCRIBE THE DOMAIN
- CHOOSE AN ORDERING FOR THE VARIABLES
- WHILE THERE ARE VARIABLES:

NETWORK CONSTRUCTION

- CHOOSE RANDOM VARIABLES X_i THAT DESCRIBE THE DOMAIN
- CHOOSE AN ORDERING FOR THE VARIABLES
- WHILE THERE ARE VARIABLES:
    - PICK A VARIABLE X_i AND ADD A NODE TO THE NETWORK

NETWORK CONSTRUCTION

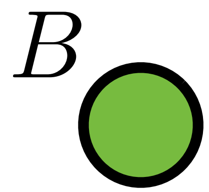
- CHOOSE RANDOM VARIABLES X_i THAT DESCRIBE THE DOMAIN
- CHOOSE AN ORDERING FOR THE VARIABLES
- WHILE THERE ARE VARIABLES:
  - PICK A VARIABLE X_i AND ADD A NODE TO THE NETWORK
  - SET PARENTS(X_i) TO SOME SET OF EXISTING NODES

NETWORK CONSTRUCTION

- CHOOSE RANDOM VARIABLES X_i THAT DESCRIBE THE DOMAIN
- CHOOSE AN ORDERING FOR THE VARIABLES
- WHILE THERE ARE VARIABLES:
    - PICK A VARIABLE X_i AND ADD A NODE TO THE NETWORK
    - SET PARENTS(X_i) TO SOME SET OF EXISTING NODES
    - DEFINE A CONDITIONAL PROBABILITY TABLE FOR X_i

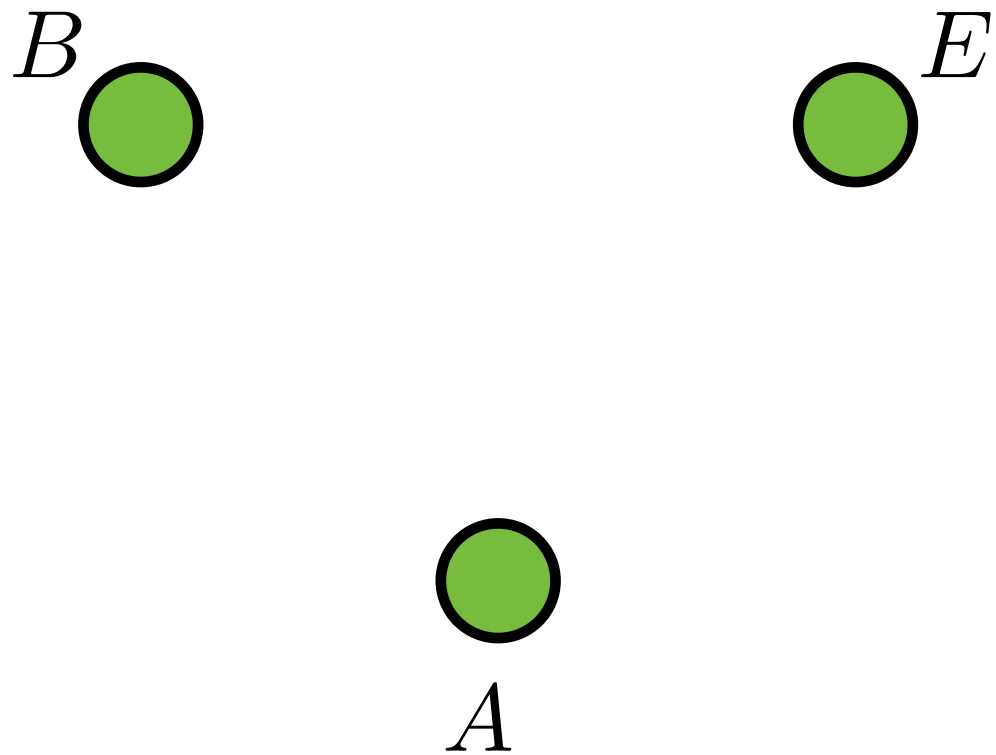
$$\{B, E, A, J, M\}$$

$$\{B, E, A, J, M\}$$

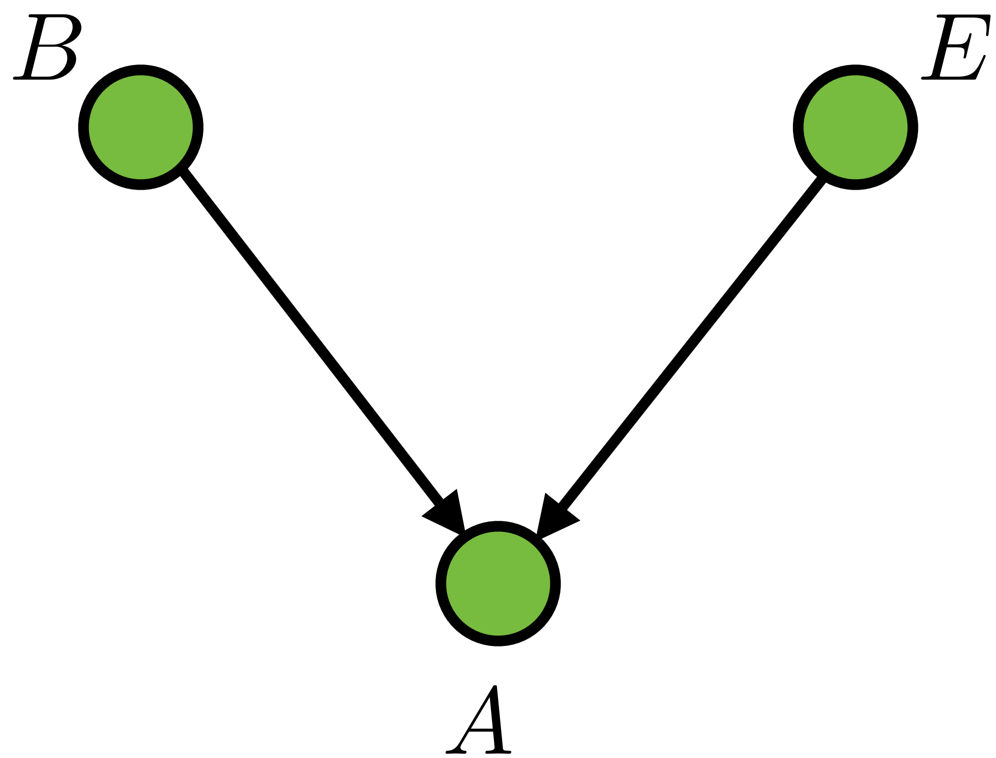


$$\{B, E, A, J, M\}$$

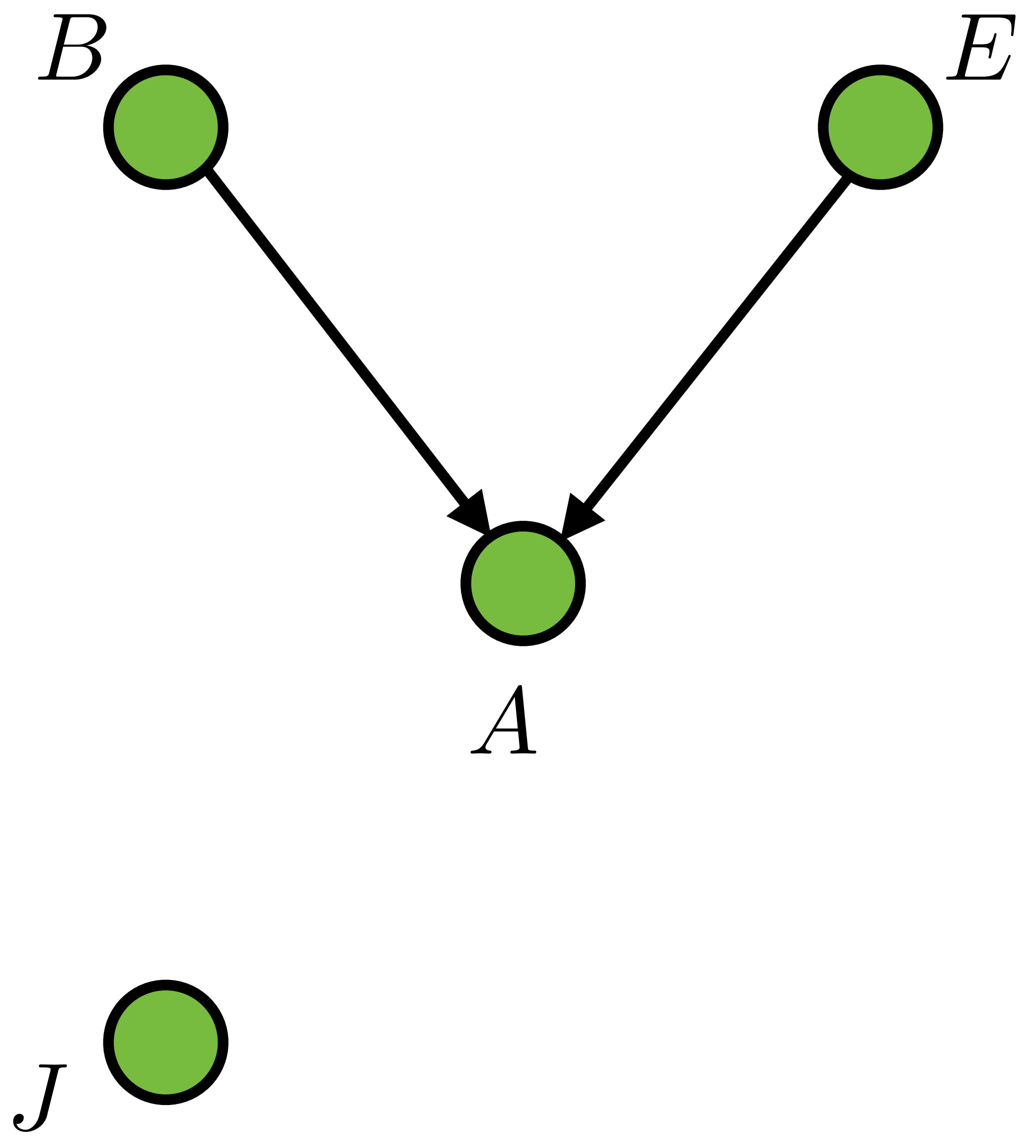
$$\{B, E, A, J, M\}$$



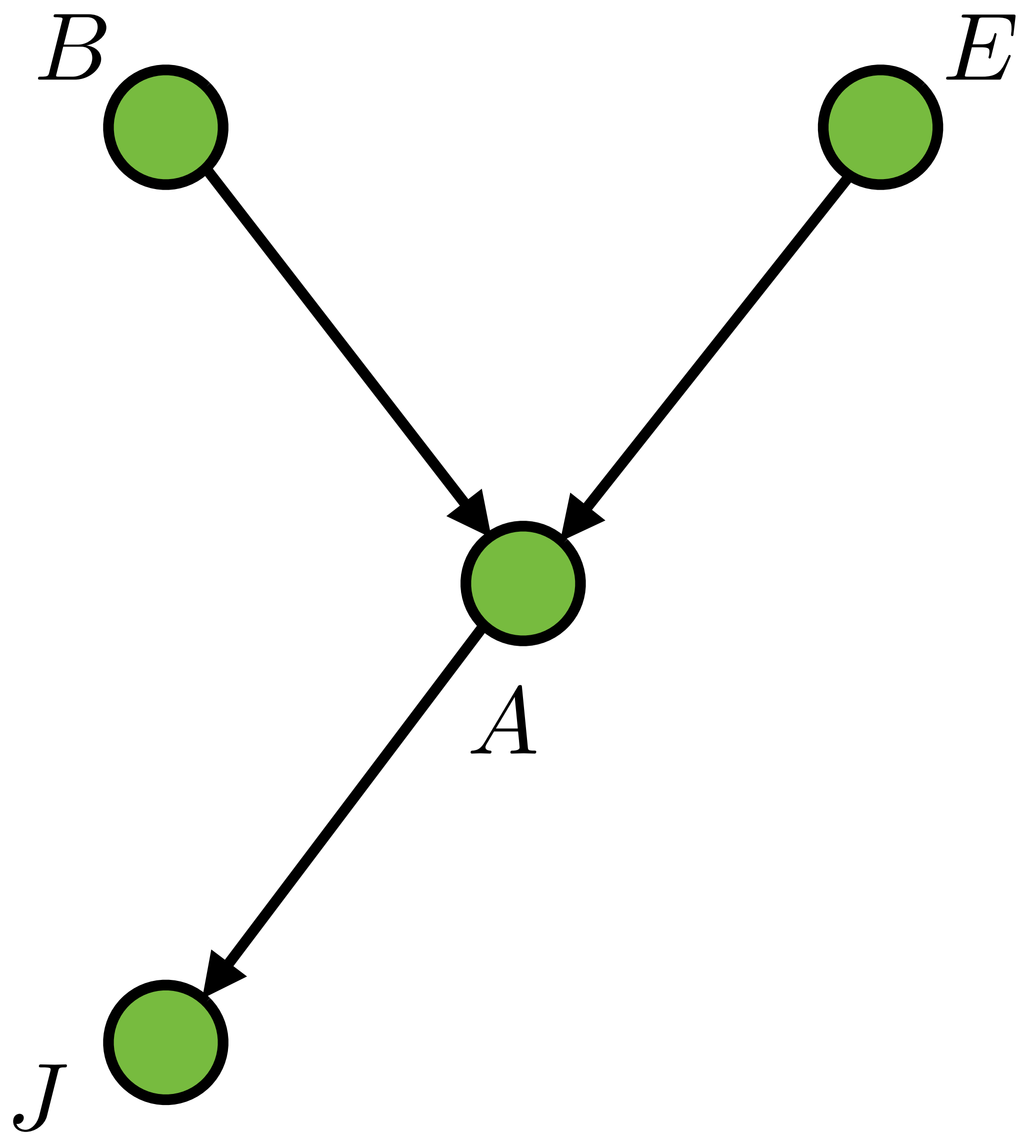
$$\{B, E, A, J, M\}$$



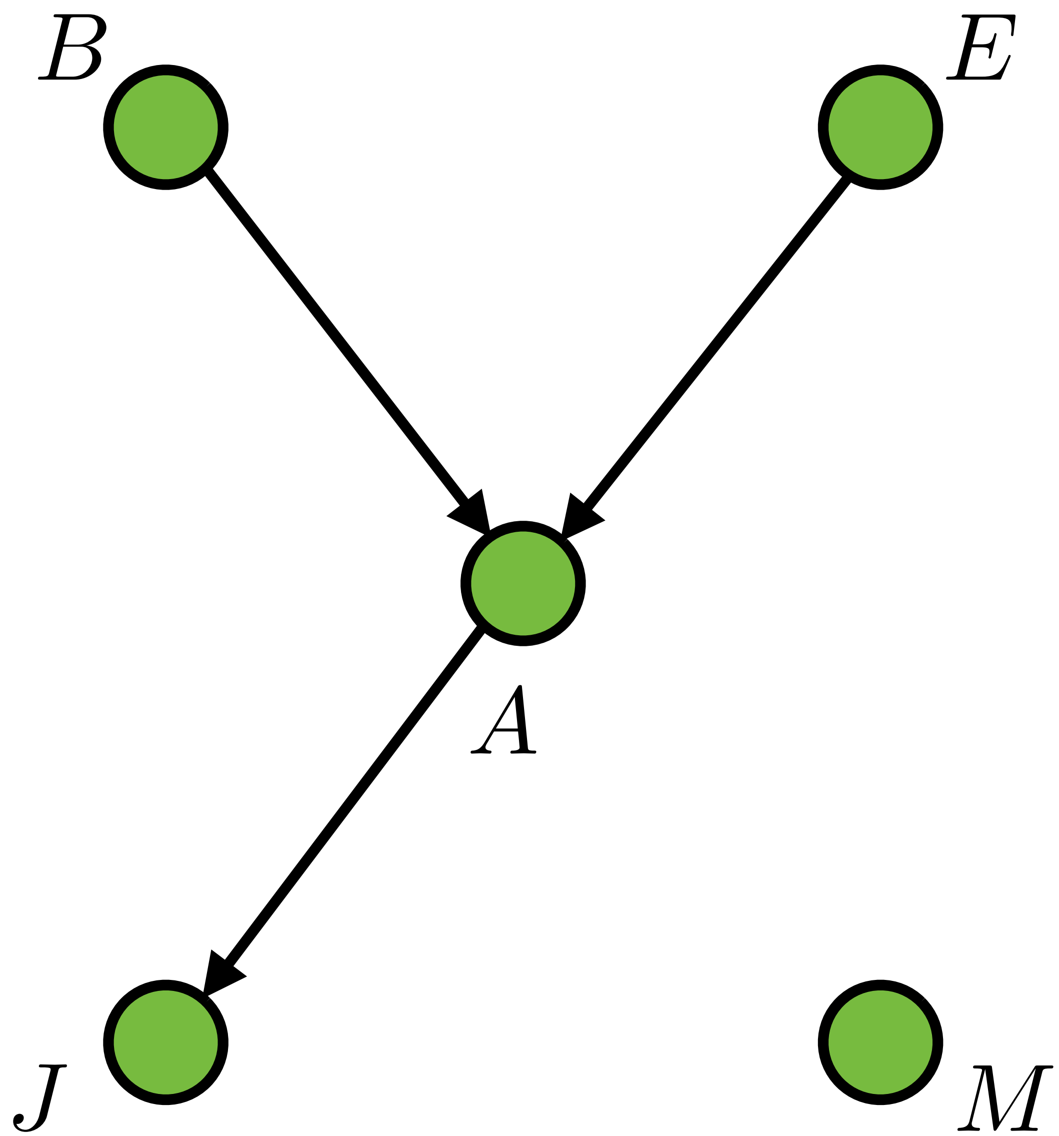
$$\{B, E, A, J, M\}$$



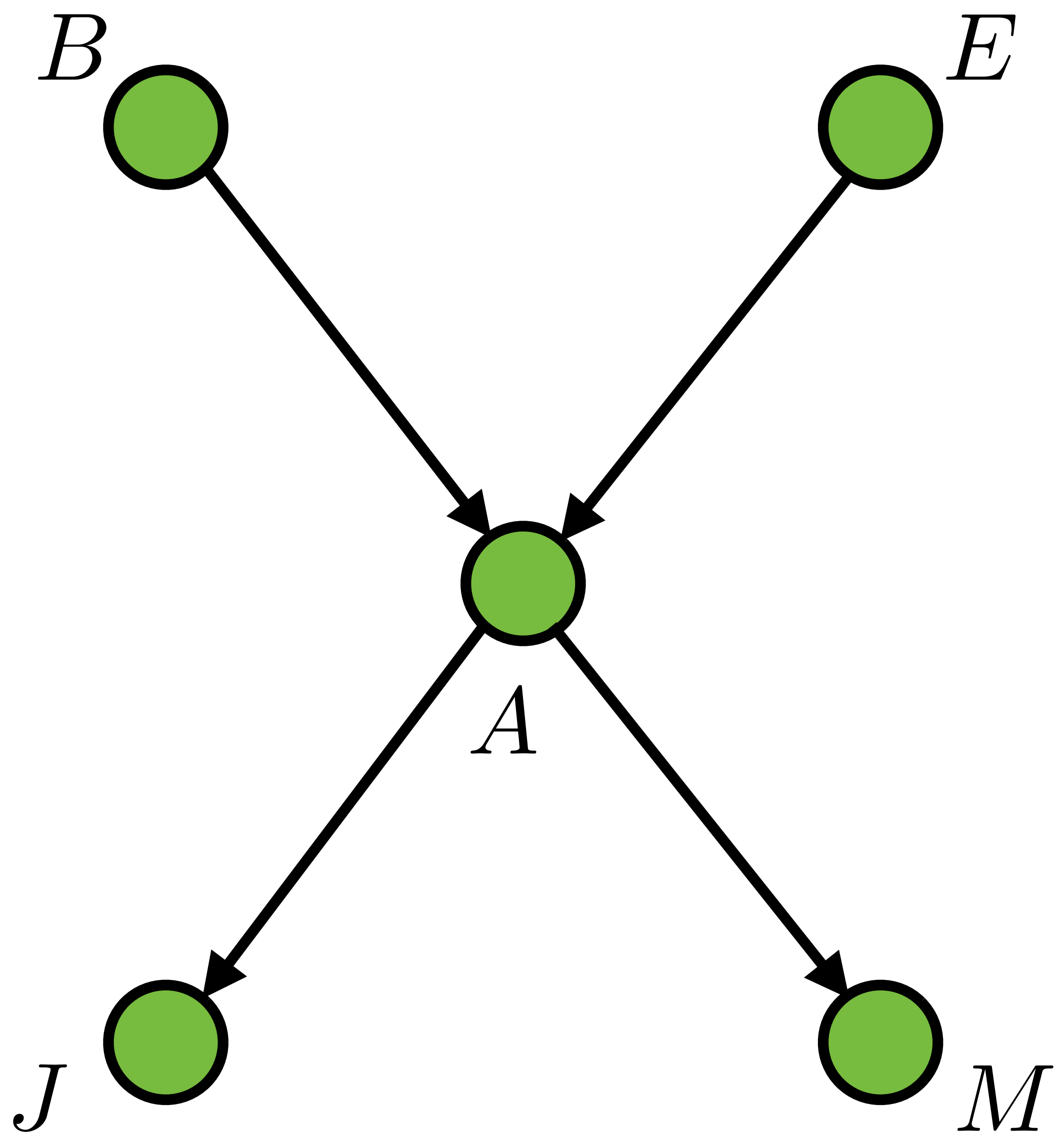
$$\{B, E, A, J, M\}$$



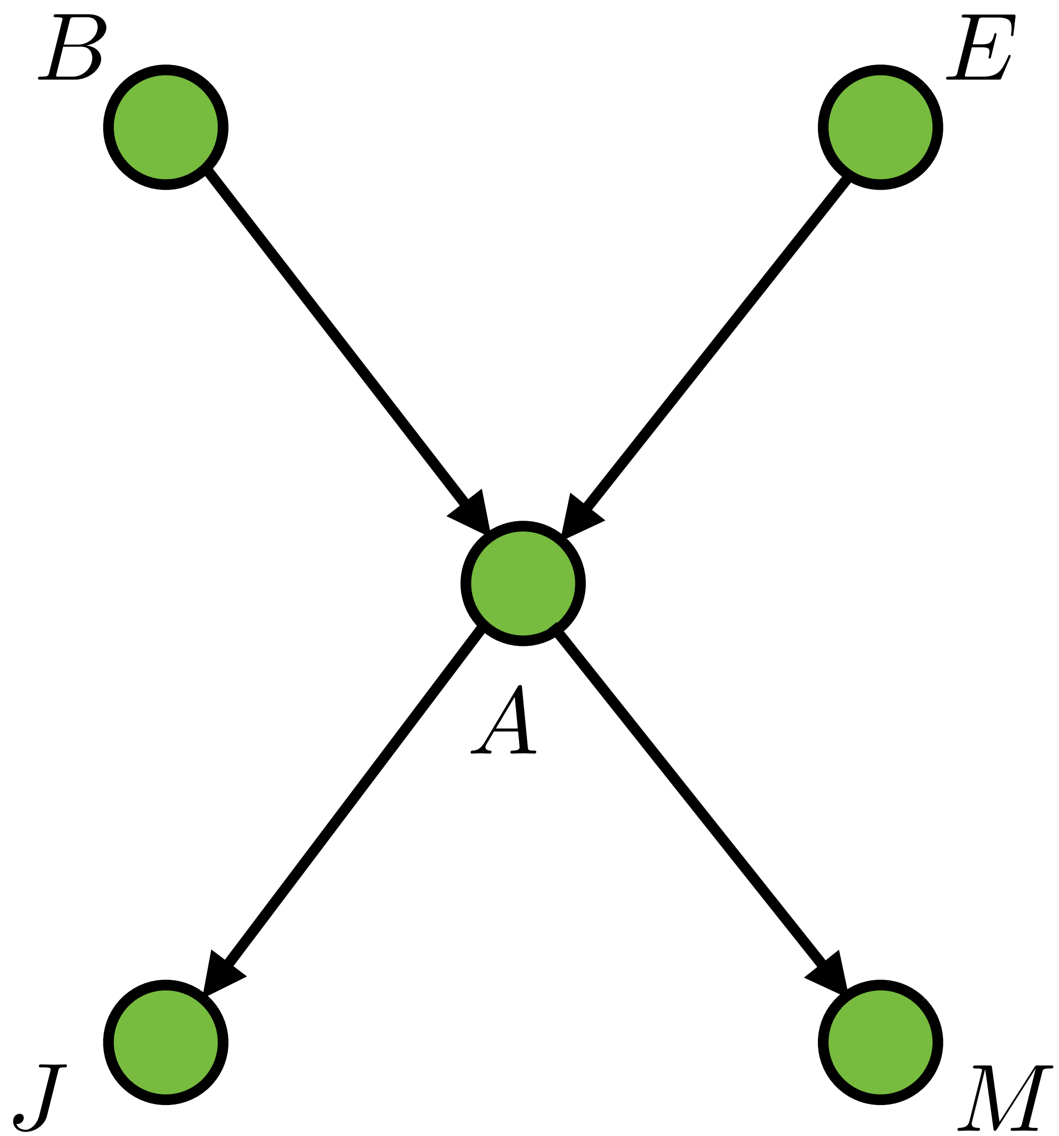
$$\{B, E, A, J, M\}$$



$$\{B, E, A, J, M\}$$



$$\{B, E, A, J, M\}$$

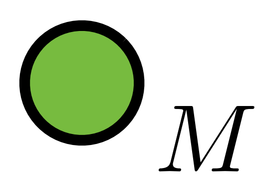


RULE OF THUMB: RESPECT CAUSALITY

GOOD ORDERING BECAUSE "ROOT CAUSES" ADDED FIRST

$$\{M, J, A, B, E\}$$

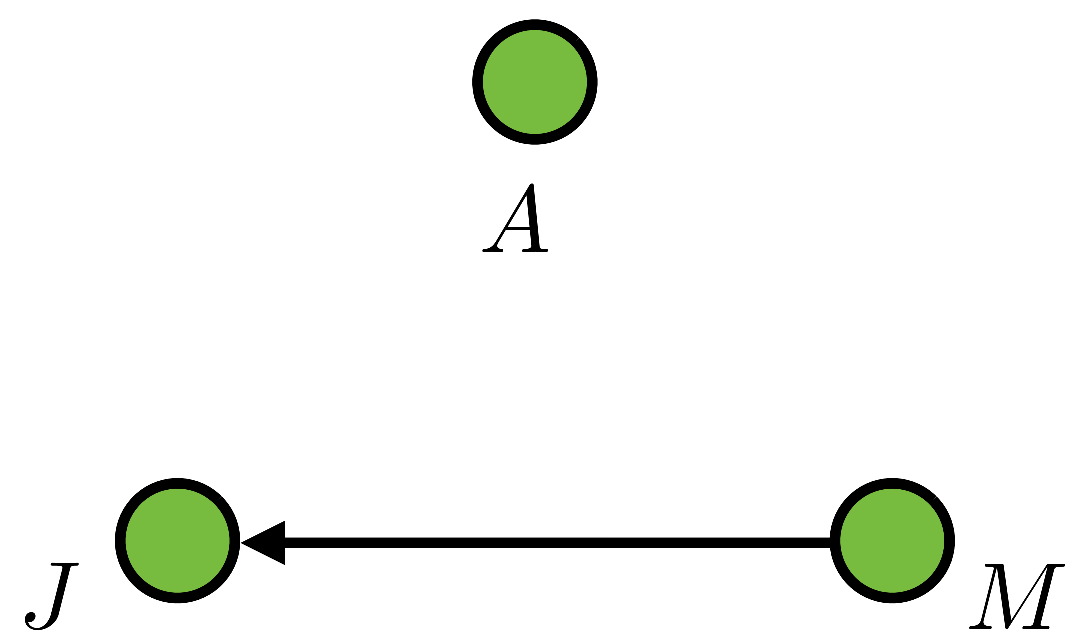
$$\{M, J, A, B, E\}$$



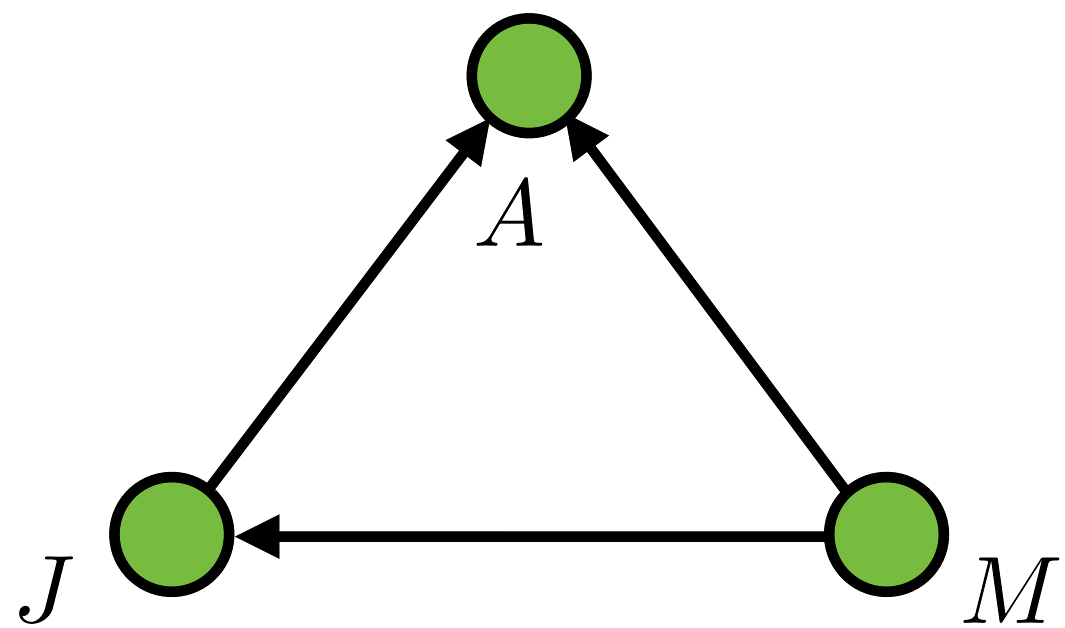
$$\{M, J, A, B, E\}$$

$$\{M, J, A, B, E\}$$

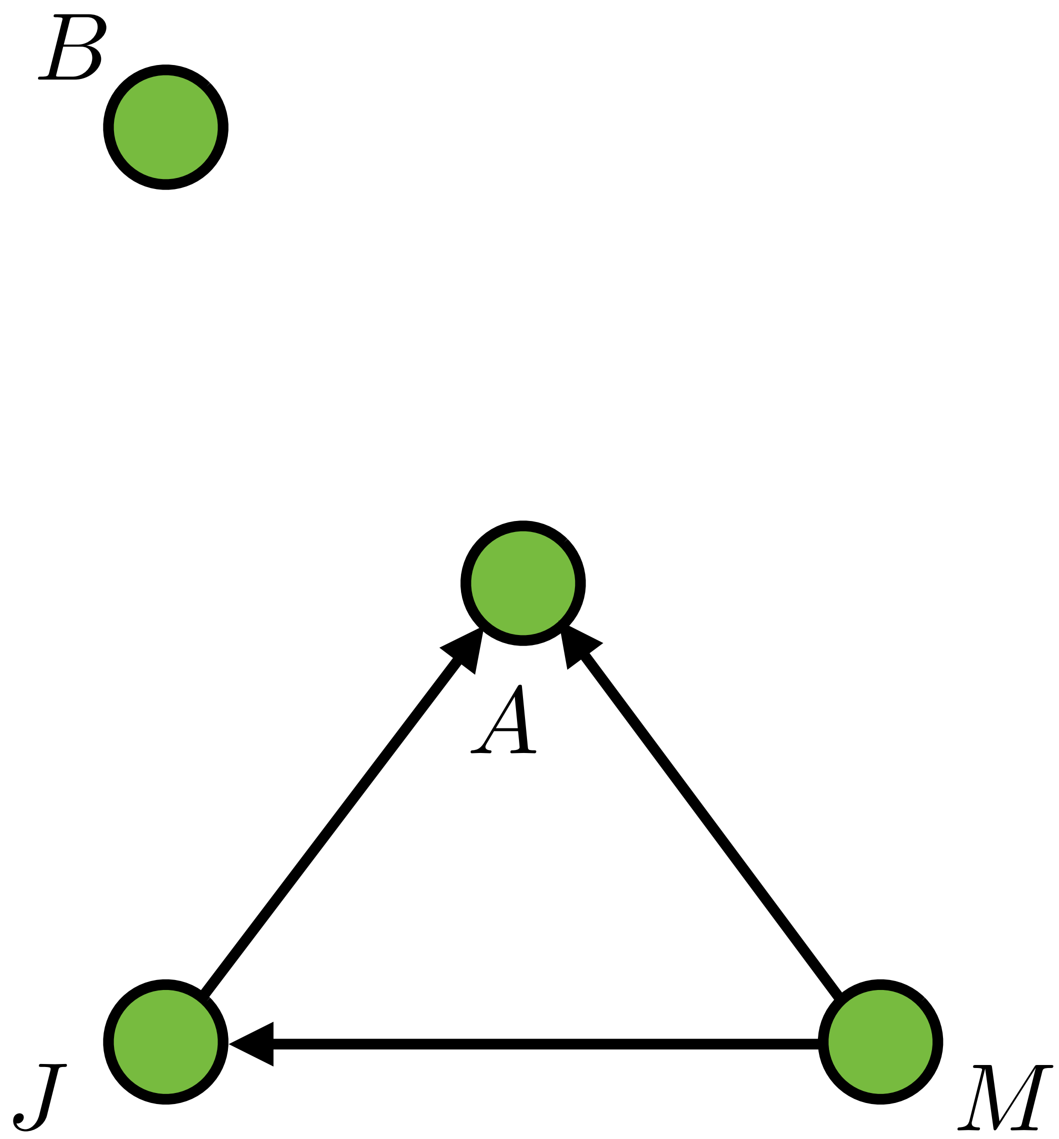
$$\{M, J, A, B, E\}$$



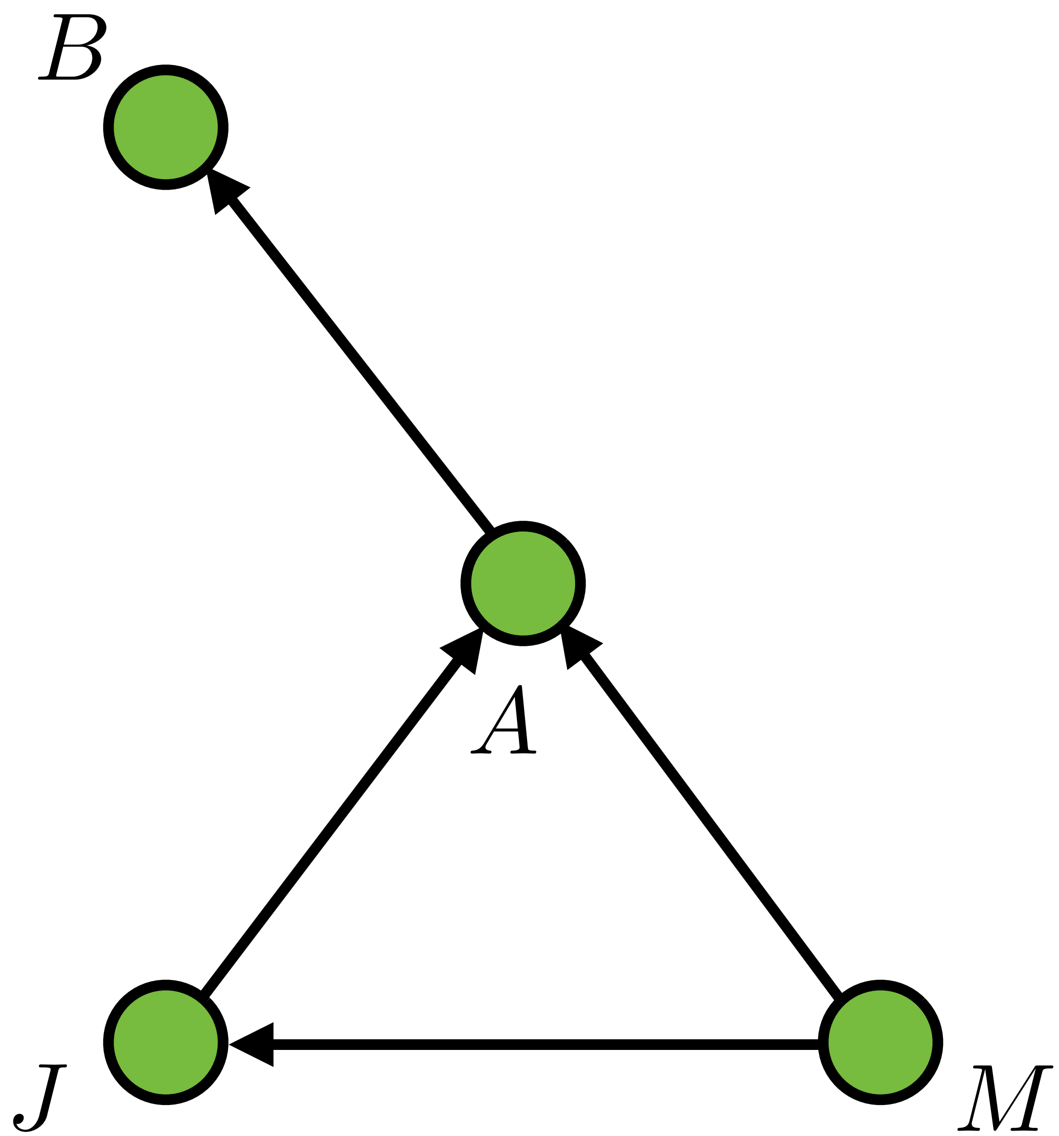
$$\{M, J, A, B, E\}$$



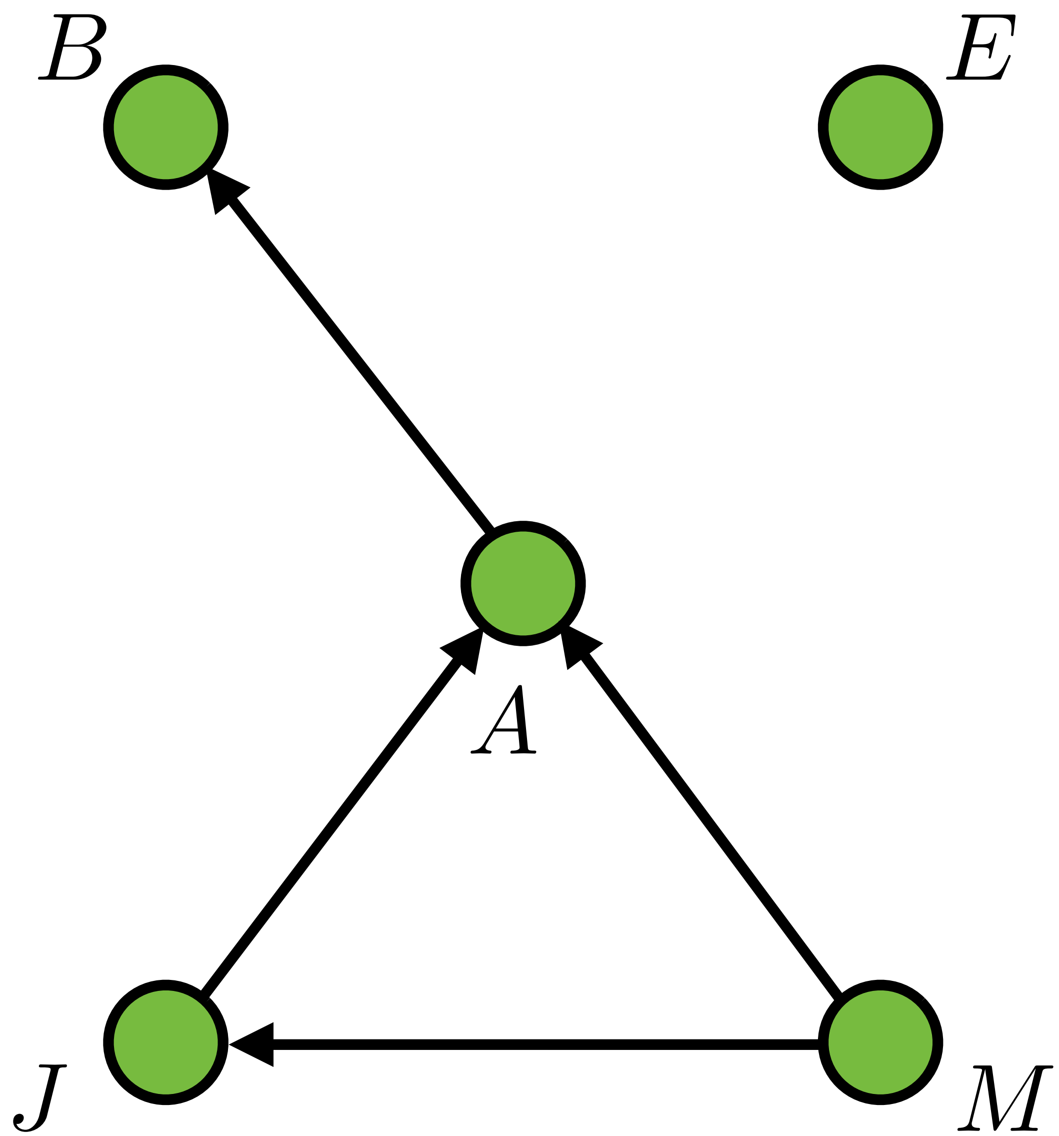
$$\{M, J, A, B, E\}$$



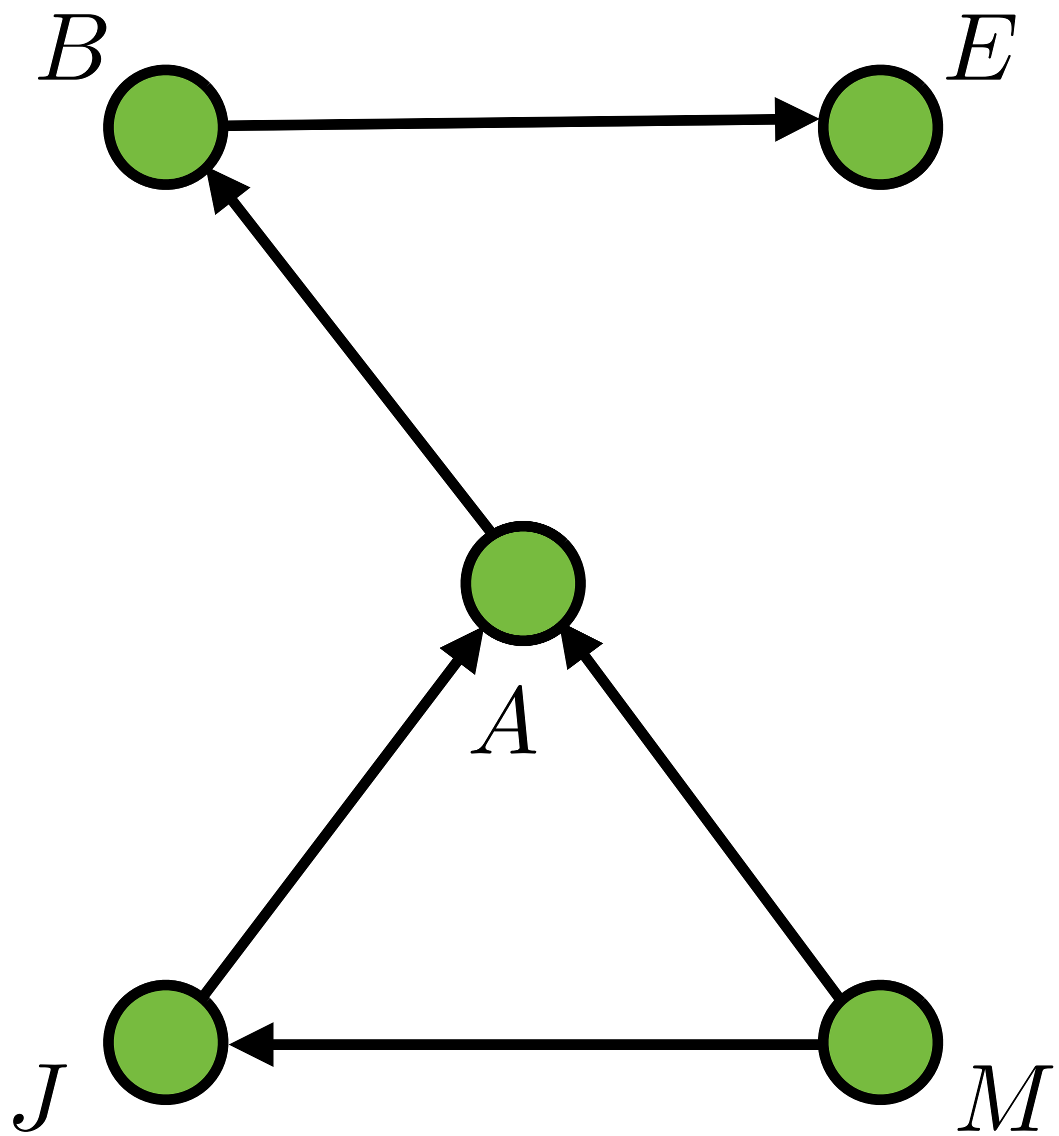
$$\{M, J, A, B, E\}$$



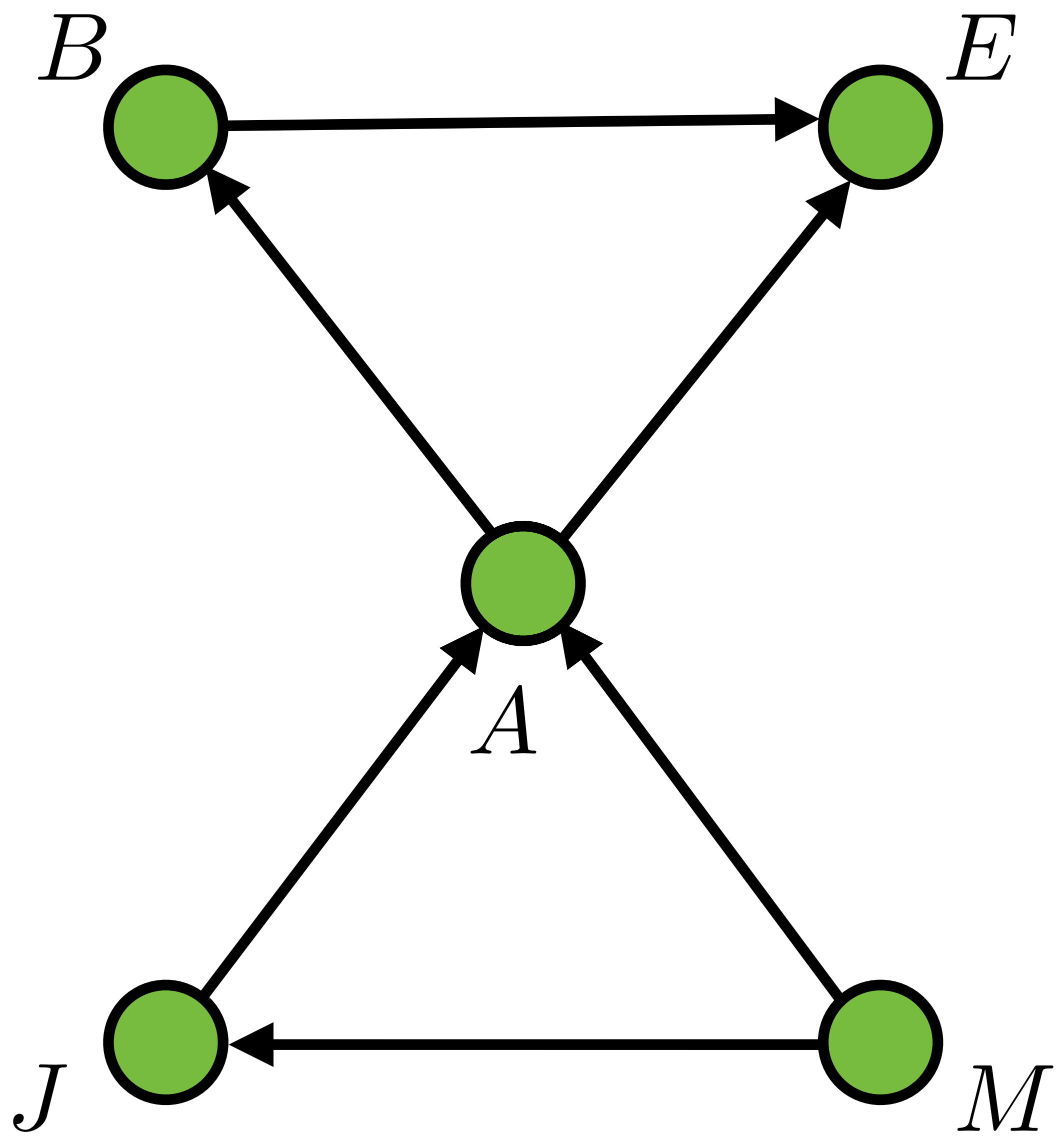
$$\{M, J, A, B, E\}$$



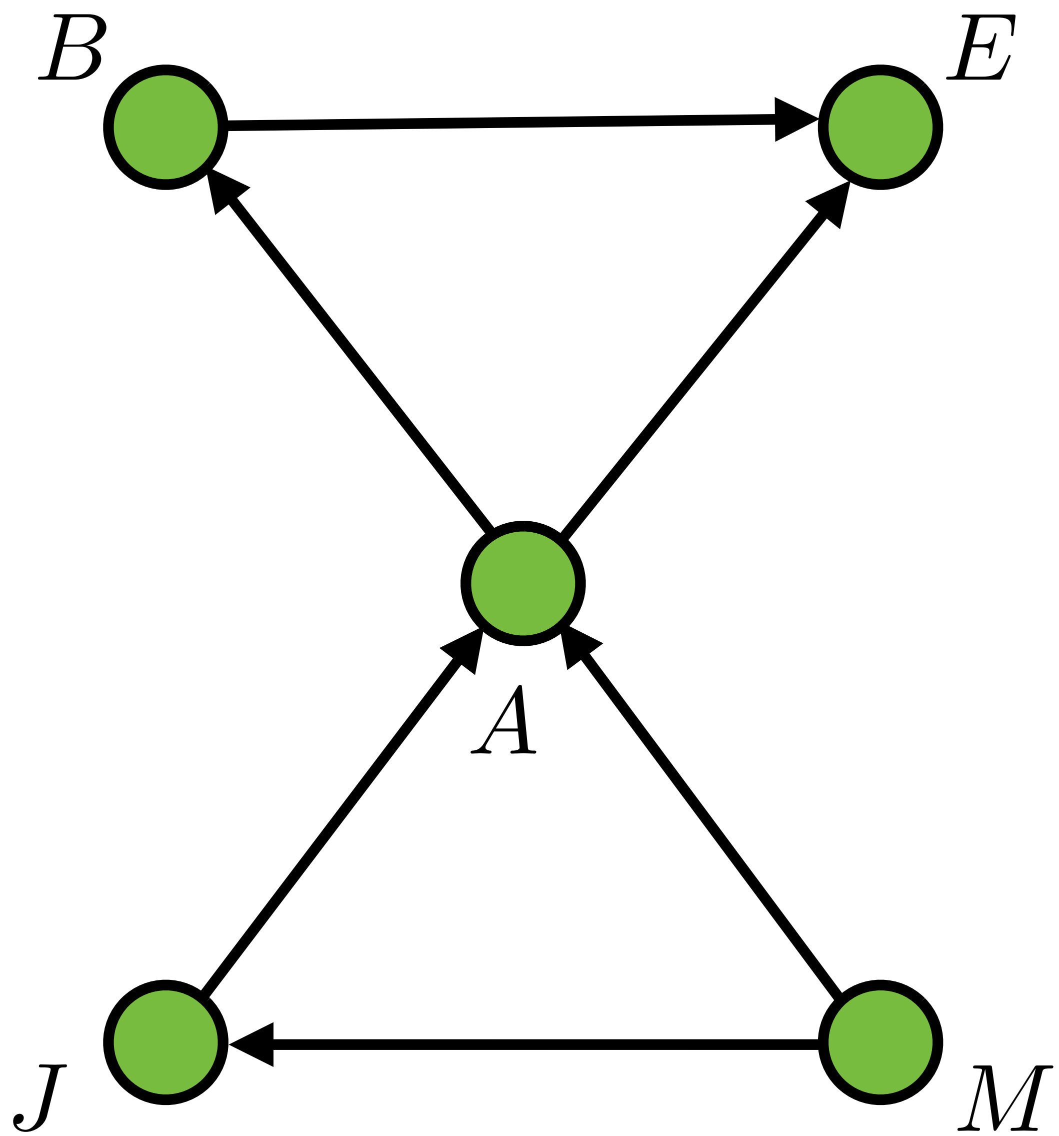
$$\{M, J, A, B, E\}$$



$$\{M, J, A, B, E\}$$



$$\{M, J, A, B, E\}$$

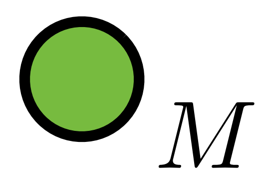


DIAGNOSTIC vs CAUSAL

$$\{M, J, E, B, A\}$$

TOTAL NUMBER OF PARAMETERS?

$$\{M, J, E, B, A\}$$



TOTAL NUMBER OF PARAMETERS?

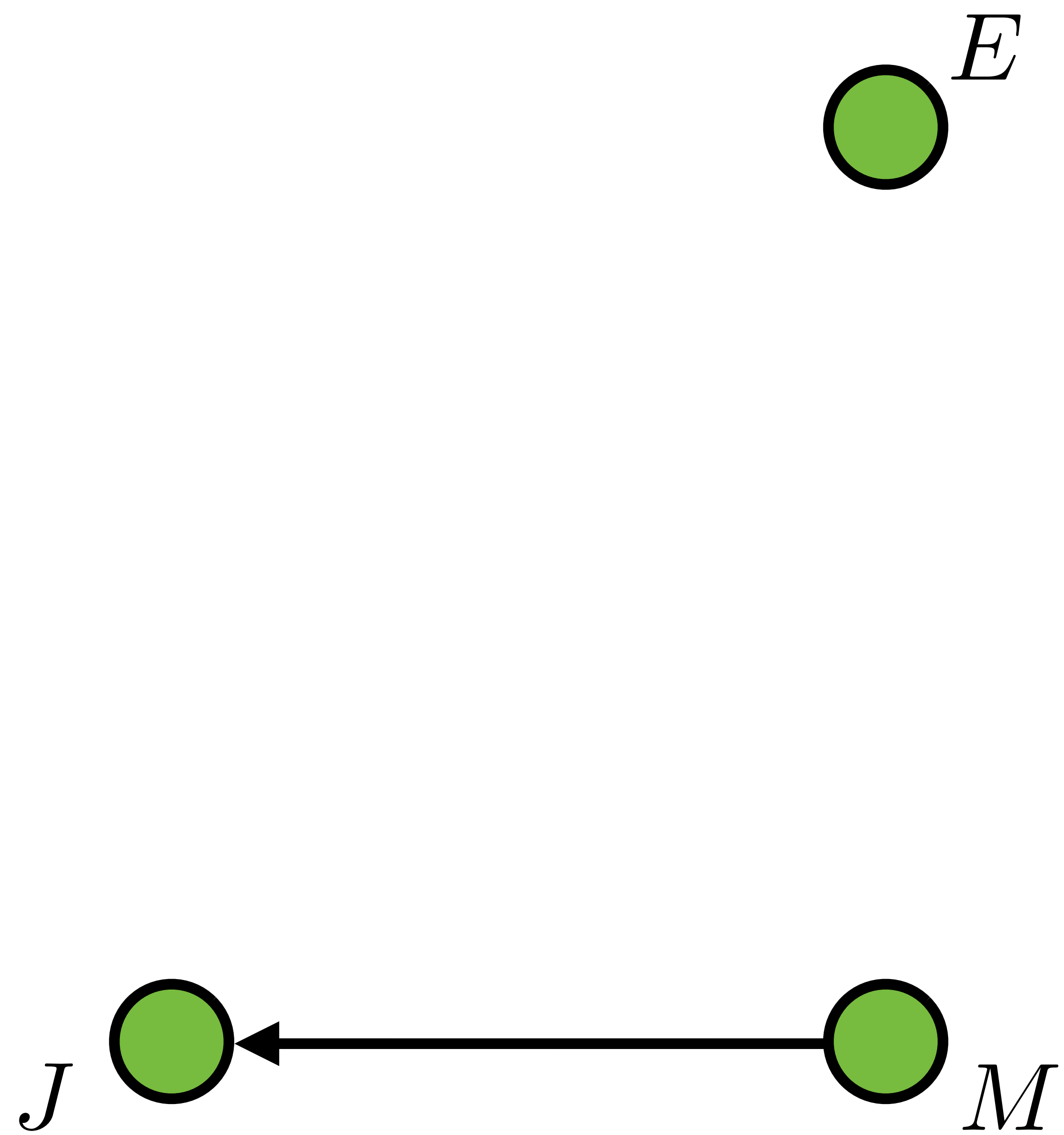
$$\{M, J, E, B, A\}$$

TOTAL NUMBER OF PARAMETERS?

$$\{M, J, E, B, A\}$$

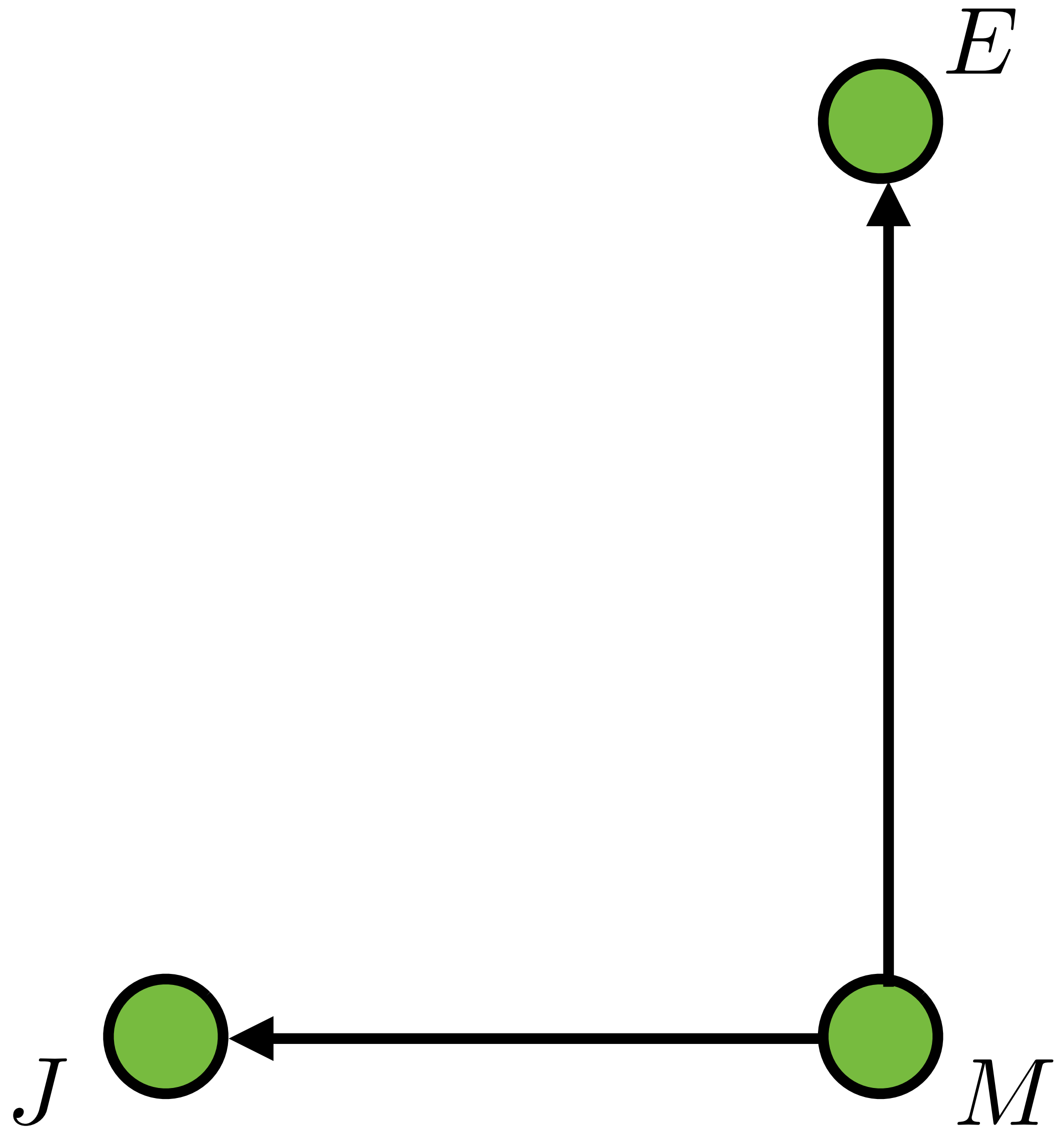
TOTAL NUMBER OF PARAMETERS?

$$\{M, J, E, B, A\}$$



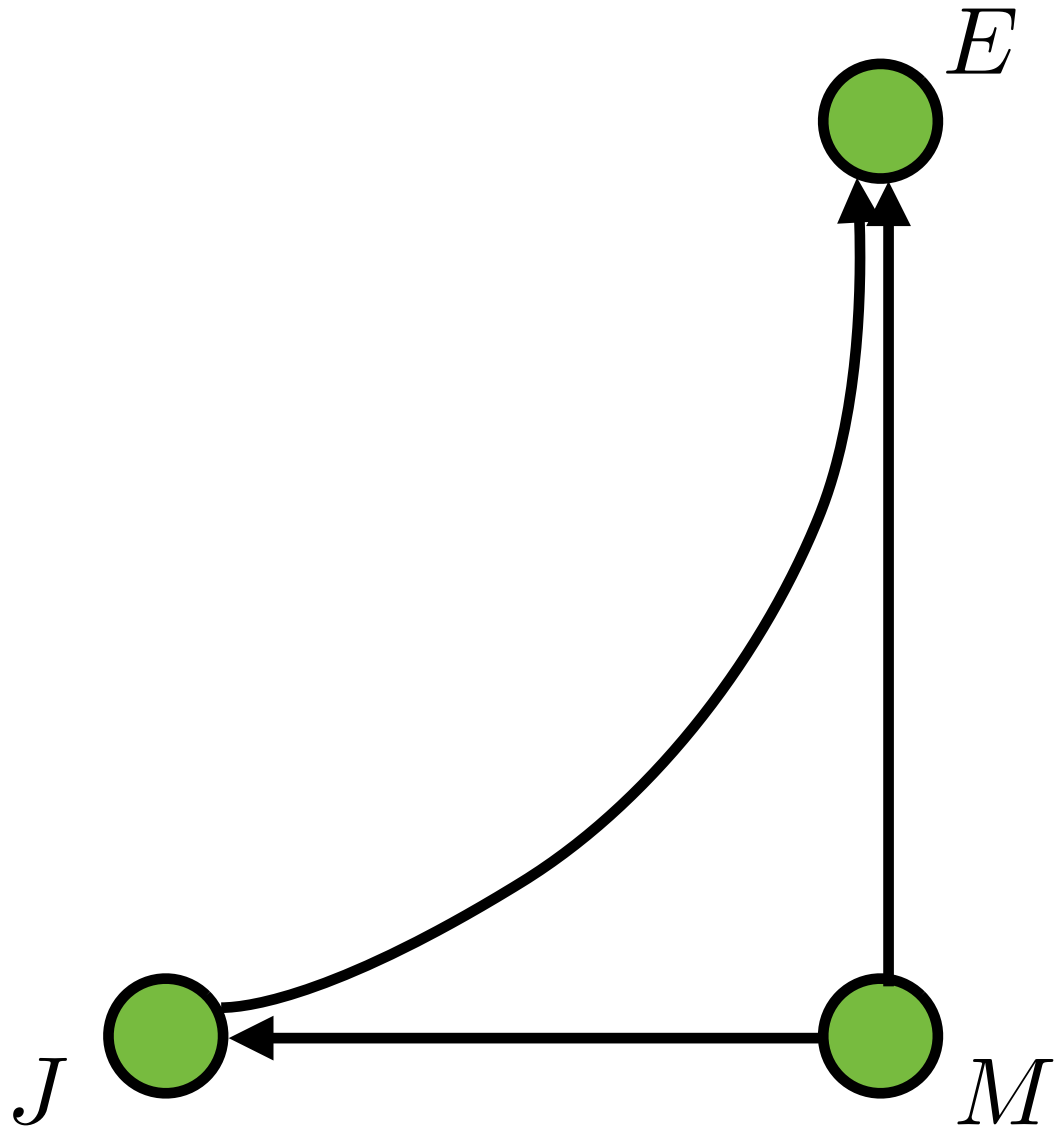
TOTAL NUMBER OF PARAMETERS?

$$\{M, J, E, B, A\}$$



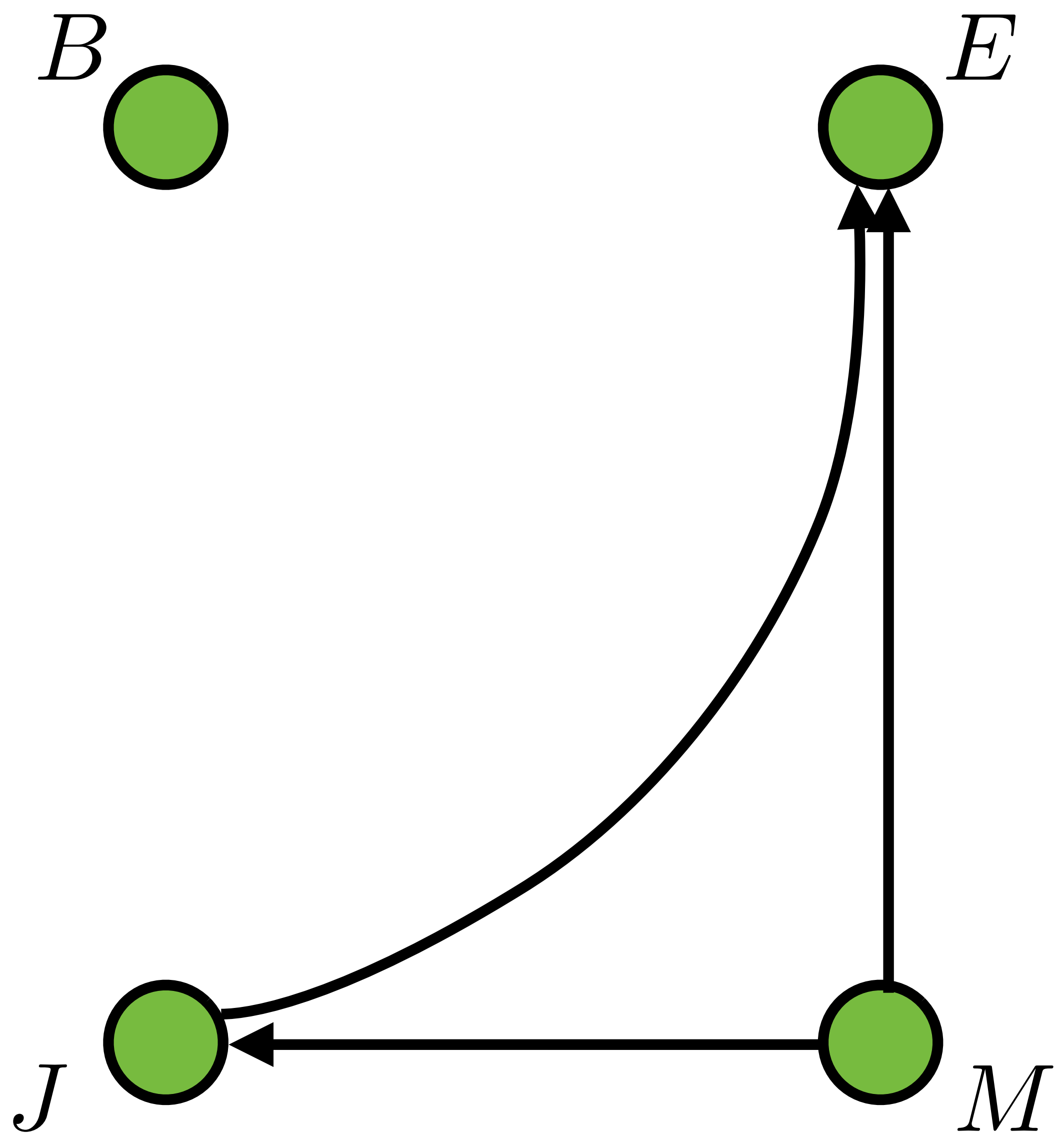
TOTAL NUMBER OF PARAMETERS?

$$\{M, J, E, B, A\}$$



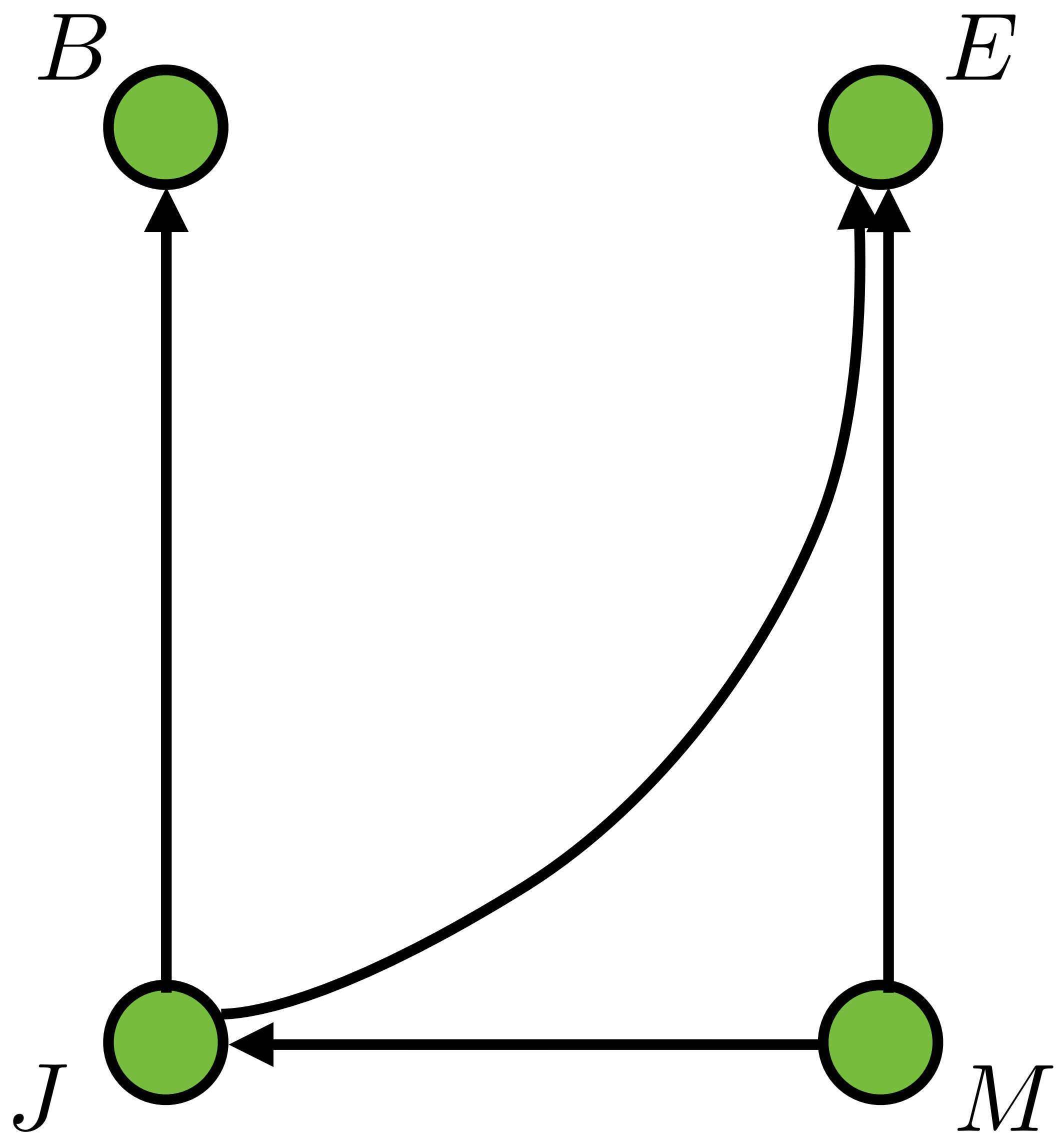
TOTAL NUMBER OF PARAMETERS?

$$\{M, J, E, B, A\}$$



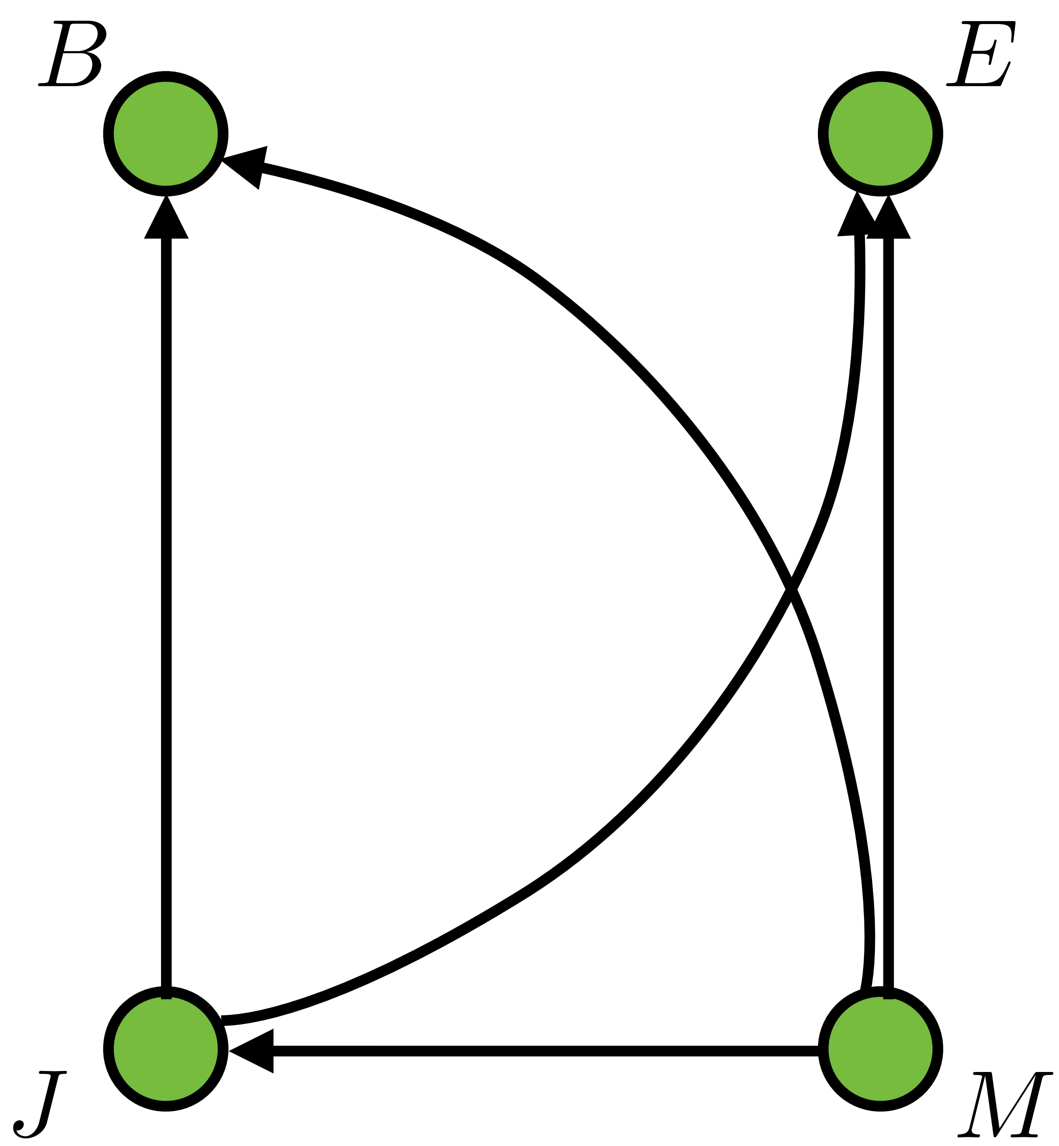
TOTAL NUMBER OF PARAMETERS?

$$\{M, J, E, B, A\}$$



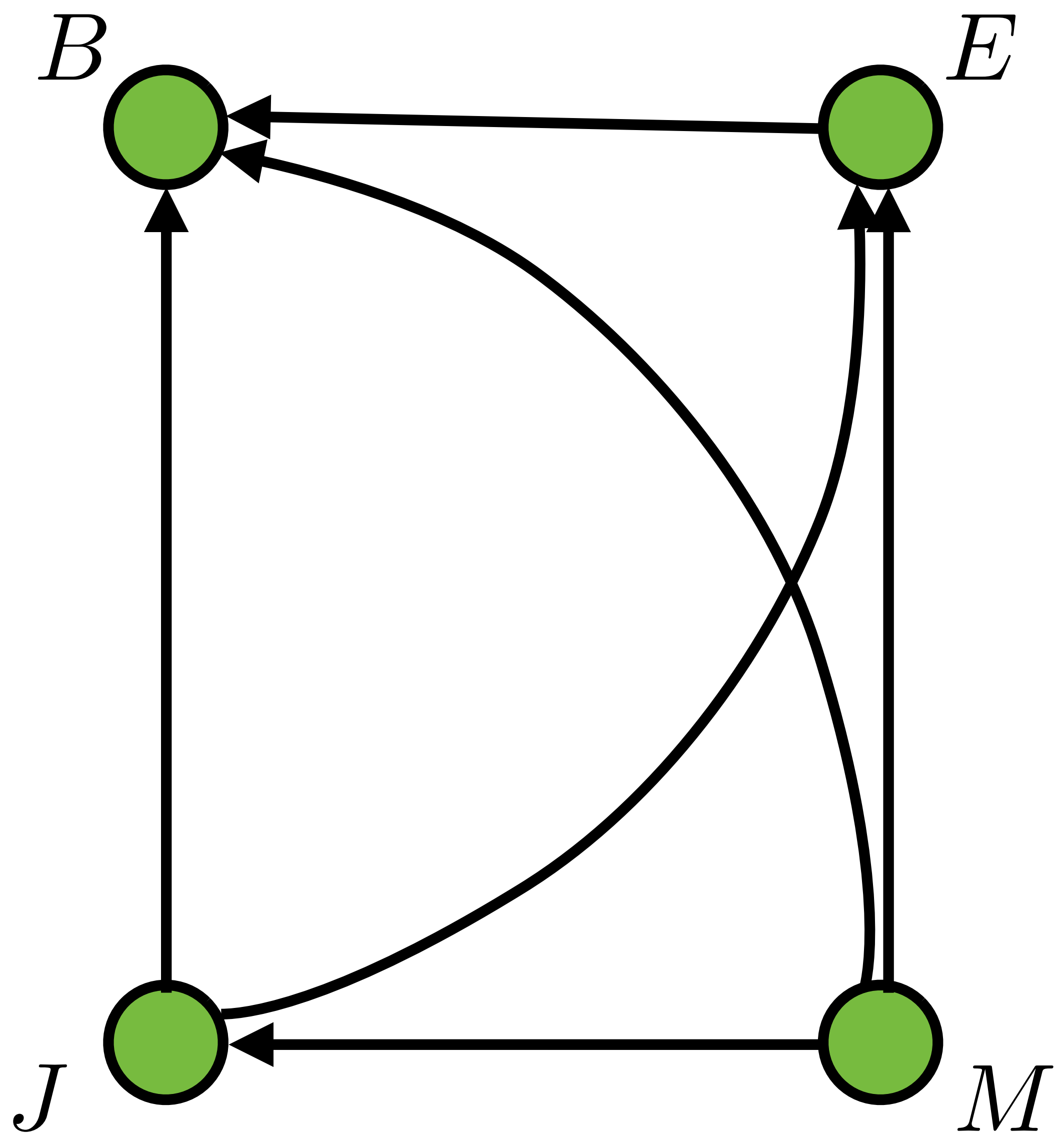
TOTAL NUMBER OF PARAMETERS?

$$\{M, J, E, B, A\}$$



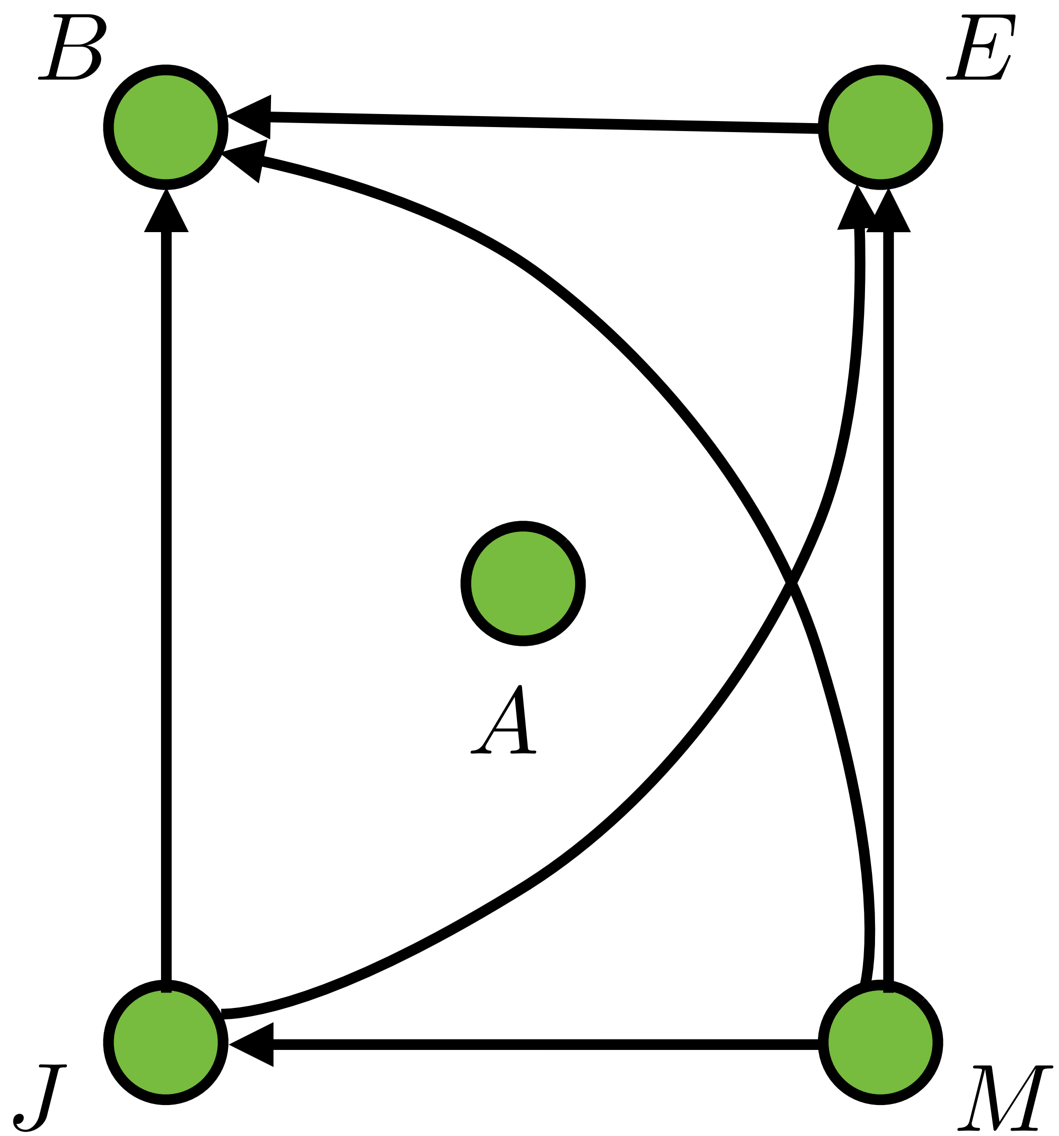
TOTAL NUMBER OF PARAMETERS?

$$\{M, J, E, B, A\}$$



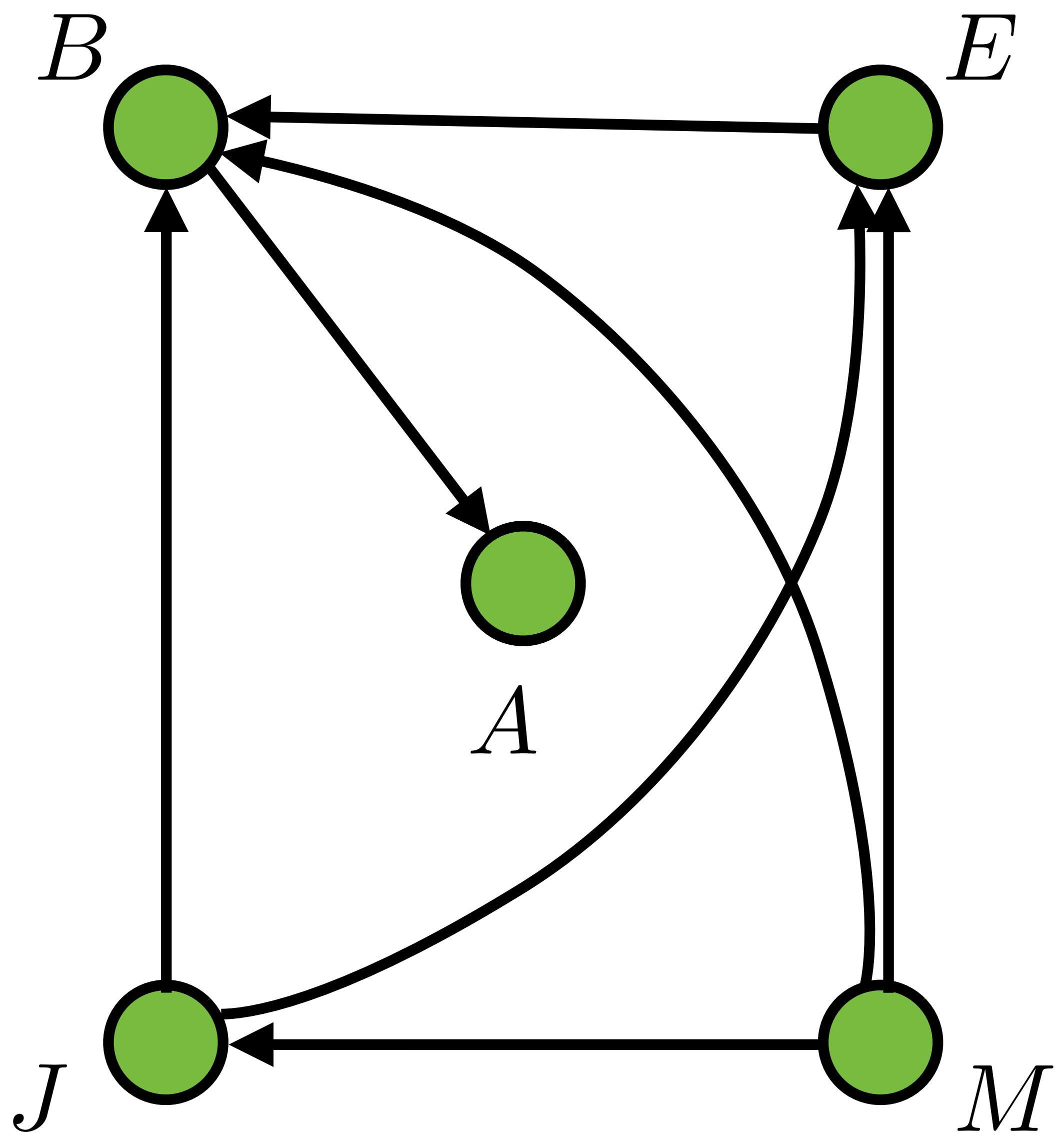
TOTAL NUMBER OF PARAMETERS?

$$\{M, J, E, B, A\}$$



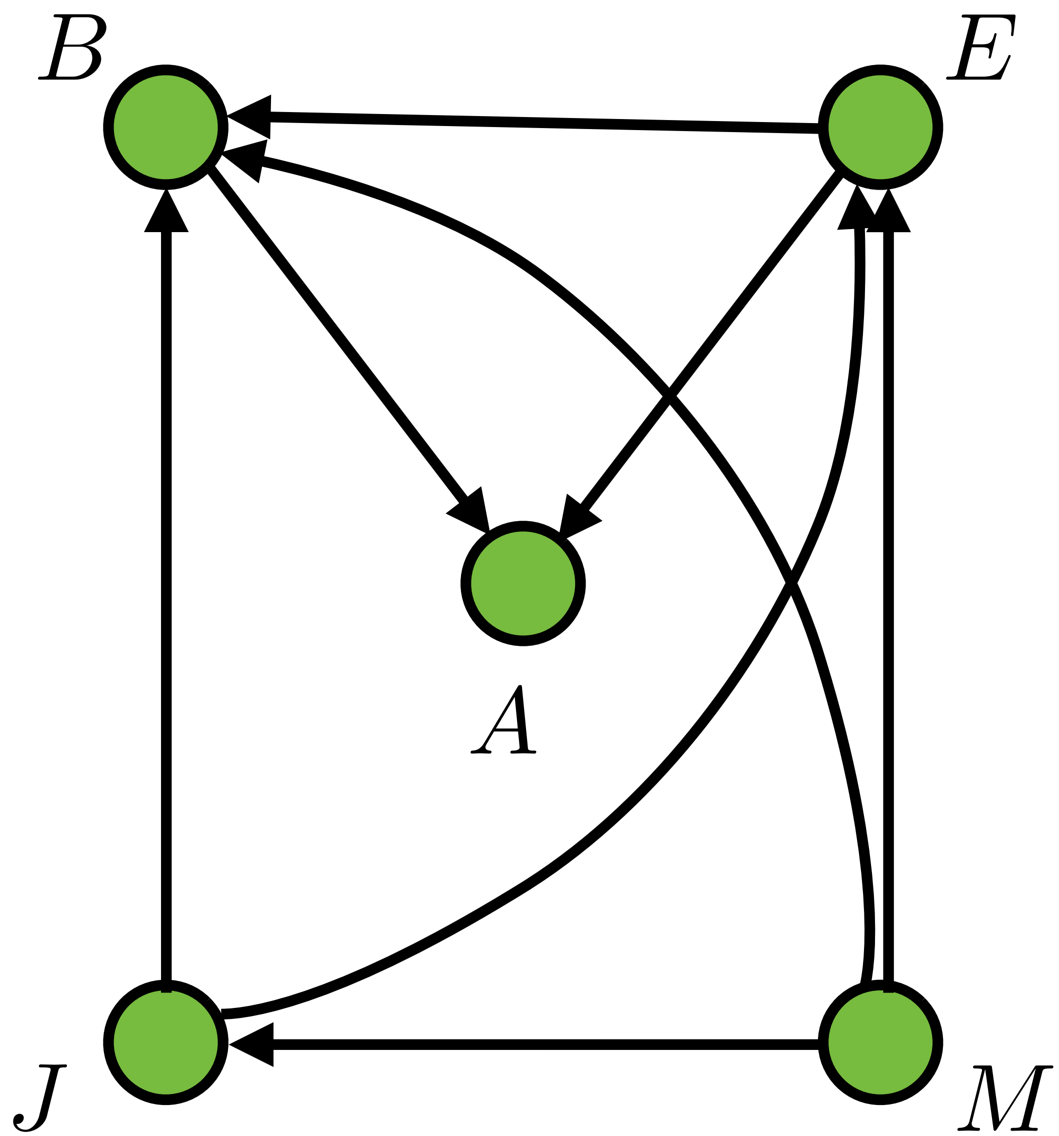
TOTAL NUMBER OF PARAMETERS?

$$\{M, J, E, B, A\}$$



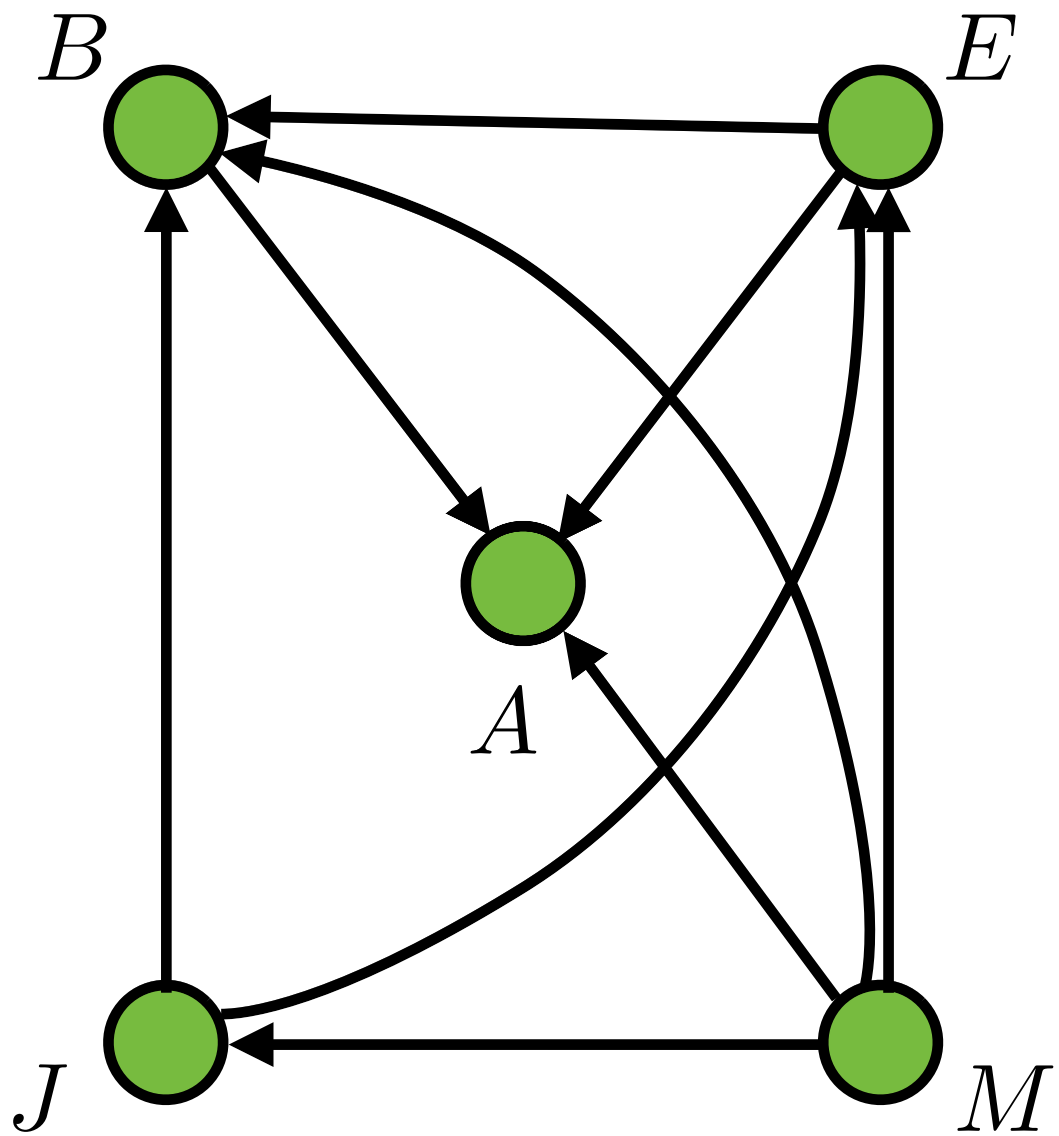
TOTAL NUMBER OF PARAMETERS?

$$\{M, J, E, B, A\}$$



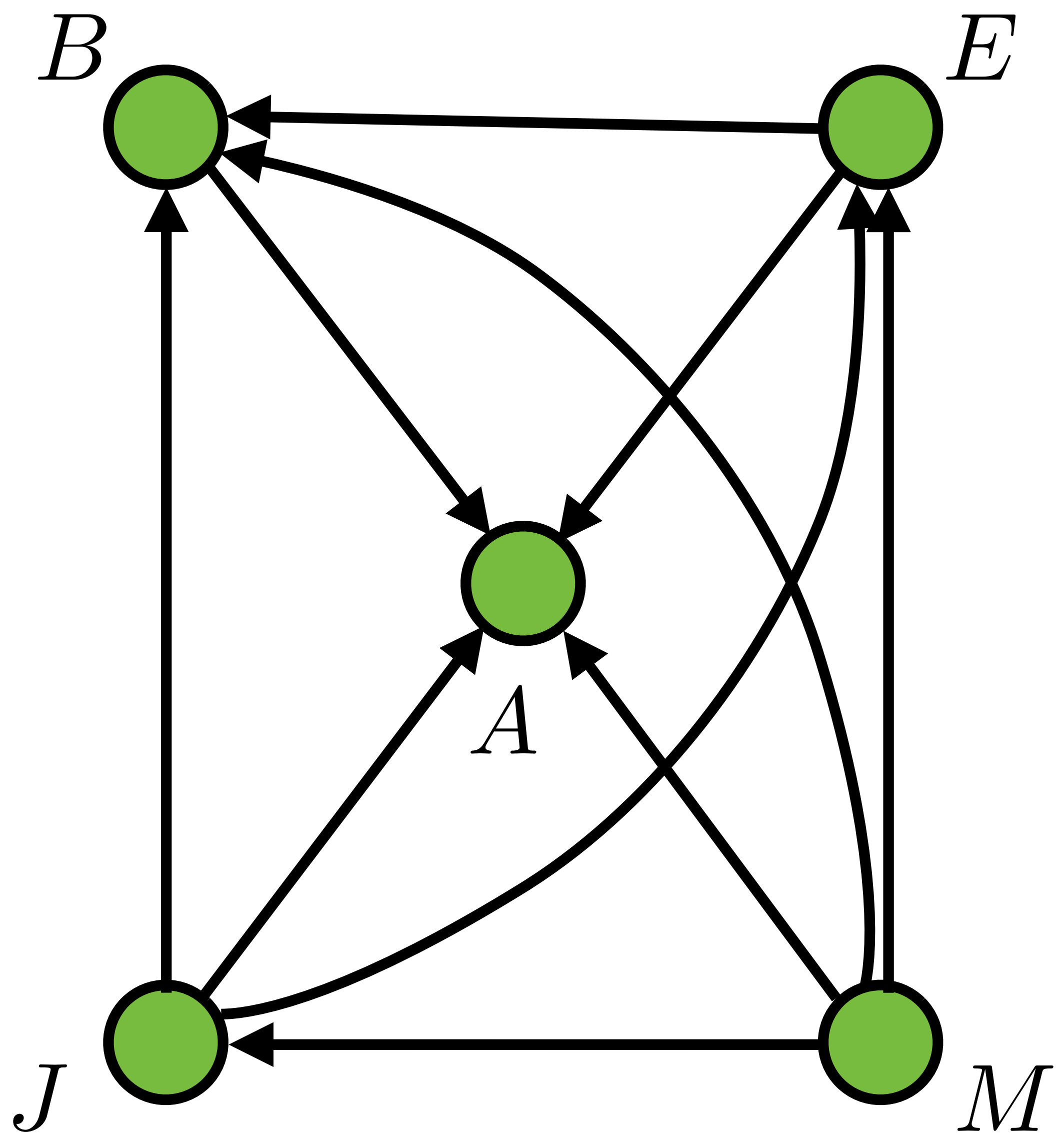
TOTAL NUMBER OF PARAMETERS?

$$\{M, J, E, B, A\}$$



TOTAL NUMBER OF PARAMETERS?

$$\{M, J, E, B, A\}$$



TOTAL NUMBER OF PARAMETERS?

COMPACTNESS

COMPACTNESS

- EACH RANDOM VARIABLE IS INFLUENCE BY *K* PARENTS

COMPACTNESS

- EACH RANDOM VARIABLE IS INFLUENCE BY *K* PARENTS
- CONDITIONAL PROBABILITIES FOR EACH BOOLEAN RV: 2^K

COMPACTNESS

- EACH RANDOM VARIABLE IS INFLUENCE BY K PARENTS
- CONDITIONAL PROBABILITIES FOR EACH BOOLEAN RV: 2^K
- CONDITIONAL PROBABILITIES FOR ALL BOOLEAN RV: $N2^K$

COMPACTNESS

- EACH RANDOM VARIABLE IS INFLUENCE BY *K* PARENTS
- CONDITIONAL PROBABILITIES FOR EACH BOOLEAN RV: 2^K
- CONDITIONAL PROBABILITIES FOR ALL BOOLEAN RV: $N2^K$
- JOINT PROBABILITY FOR ALL BOOLEAN RV: 2^N

COMPACTNESS

- EACH RANDOM VARIABLE IS INFLUENCE BY *K* PARENTS
- CONDITIONAL PROBABILITIES FOR EACH BOOLEAN RV: 2^K
- CONDITIONAL PROBABILITIES FOR ALL BOOLEAN RV: $N2^K$
- JOINT PROBABILITY FOR ALL BOOLEAN RV: 2^N

$N = 20$ BELIEF NETWORK $= 640$

$K = 5$ JOINT PROBABILITY $= 1,048,576$

COMPACTNESS

- EACH RANDOM VARIABLE IS INFLUENCE BY *K* PARENTS
- CONDITIONAL PROBABILITIES FOR EACH BOOLEAN RV: 2^K
- CONDITIONAL PROBABILITIES FOR ALL BOOLEAN RV: $N2^K$
- JOINT PROBABILITY FOR ALL BOOLEAN RV: 2^N

$N = 20$ BELIEF NETWORK $= 640$

$K = 5$ JOINT PROBABILITY $= 1,048,576$

BRAIN SCALE

$N = 10^{11}$ BELIEF NETWORK $= 2^{7036}$

$K = 7000$ JOINT PROBABILITY $= 2^{10^{11}}$

COMPACTNESS

- EACH RANDOM VARIABLE IS INFLUENCE BY *K* PARENTS
- CONDITIONAL PROBABILITIES FOR EACH BOOLEAN RV: 2^K
- CONDITIONAL PROBABILITIES FOR ALL BOOLEAN RV: $N2^K$
- JOINT PROBABILITY FOR ALL BOOLEAN RV: 2^N

$N = 20$ BELIEF NETWORK $= 640$

$K = 5$ JOINT PROBABILITY $= 1,048,576$

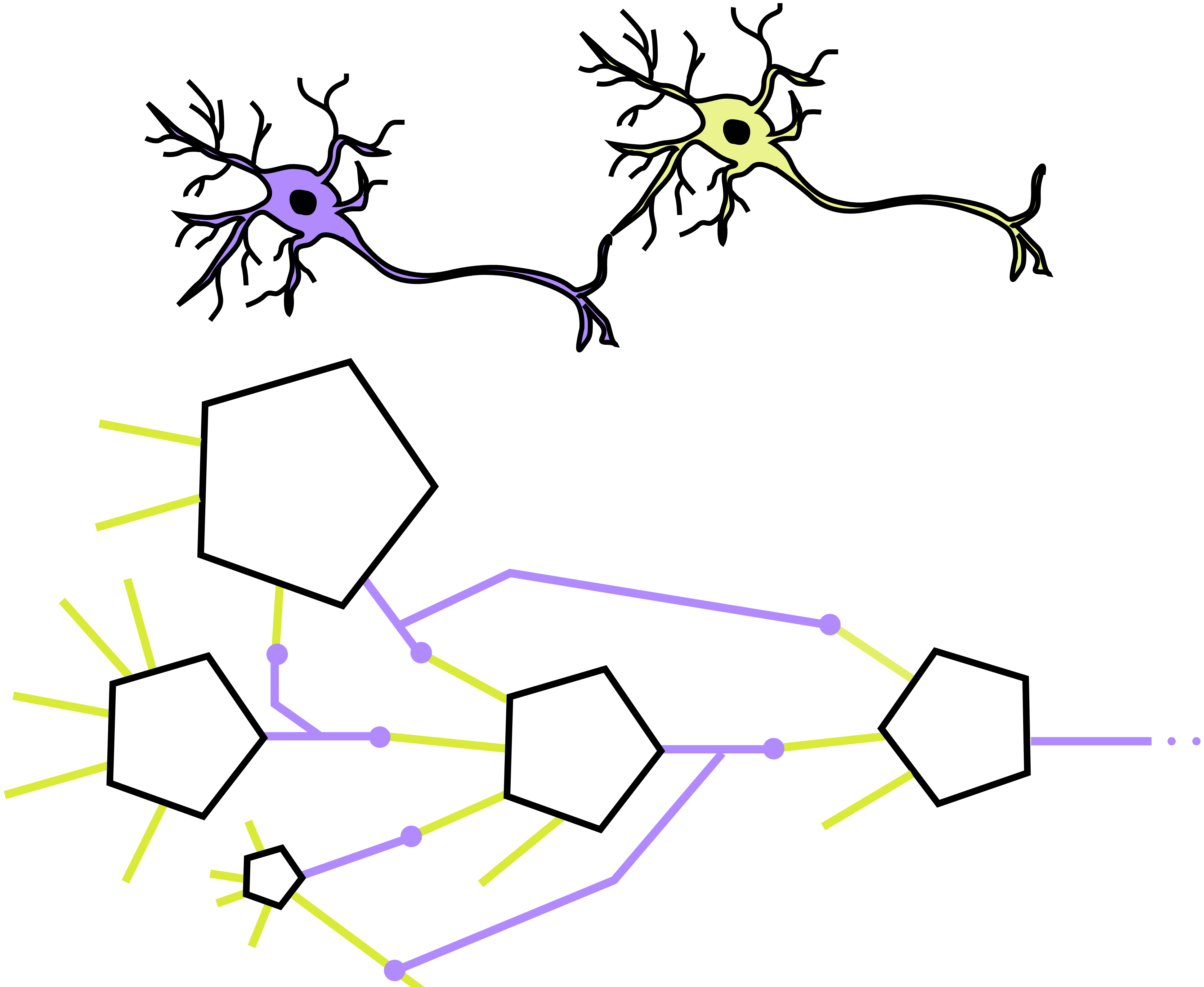
BRAIN SCALE

$N = 10^{11}$ BELIEF NETWORK $= 2^{7036}$

$K = 7000$ JOINT PROBABILITY $= 2^{10^{11}}$

ESTIMATED NUMBER OF ATOMS IN THE UNIVERSE < 2^{266}

IS THE BRAIN A BELIEF NETWORK?



CONDITIONAL INDEPENDENCE

$$p(a|b, c) = p(a|c)$$

CONDITIONAL INDEPENDENCE

$$p(a|b,c) = p(a|c)$$

$$\begin{aligned} p(a,b|c) &= p(a|b,c)p(b|c) \\ &= p(a|c)p(b|c) \end{aligned}$$

CONDITIONAL INDEPENDENCE

$$p(a|b,c) = p(a|c)$$

$$\begin{aligned} p(a,b|c) &= p(a|b,c)p(b|c) \\ &= p(a|c)p(b|c) \end{aligned}$$

$$a \perp b|c$$

CONDITIONAL INDEPENDENCE

$$p(a|b,c) = p(a|c)$$

$$\begin{aligned} p(a,b|c) &= p(a|b,c)p(b|c) \\ &= p(a|c)p(b|c) \end{aligned}$$

$$a \perp b|c$$

"*a* and *b* are independent given *c*"