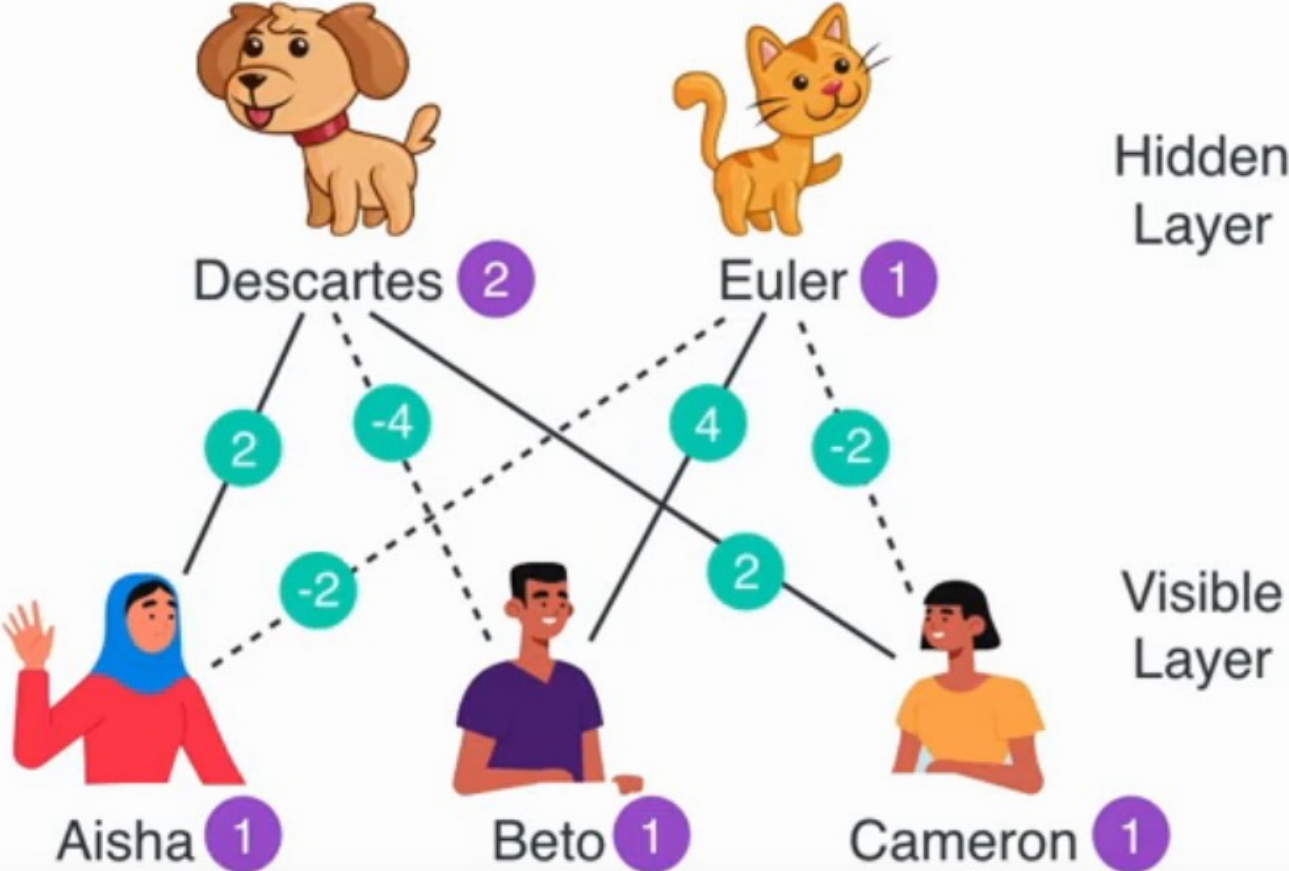


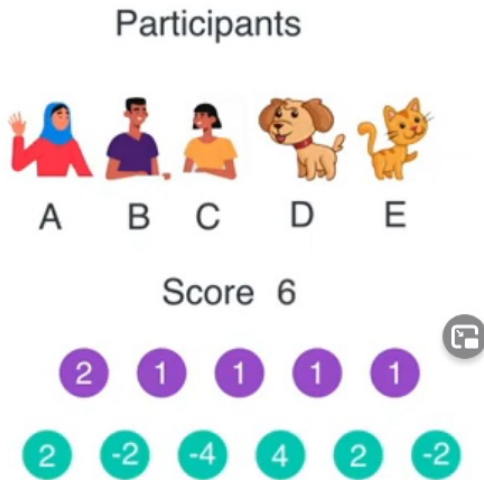
Good example:
Restricted Boltzmann
Machine

Weights

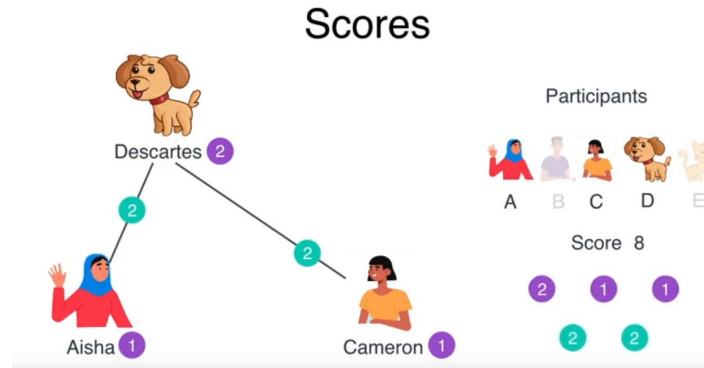


Aisha	Beto	Cameron
✓	✗	✓
✗	✓	✗
✓	✗	✓
✓	✗	✓
✗	✓	✗
✓	✗	✓
✗	✓	✗
✓	✗	✓
✓	✗	✓
✓	✗	✓

Scenario

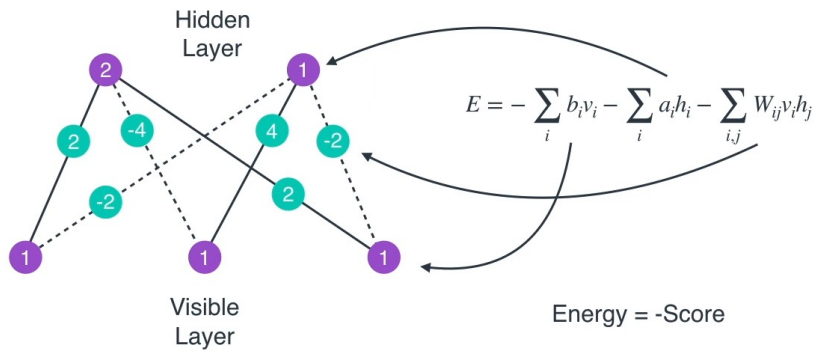


Scenario



Scenario	Score
None	0
A	1
B	1
C	1
D	2
E	1
AB	2
AC	2
AD	5
AE	0
BC	2
BD	-2
BE	7
CD	5
CE	0
DE	3
ABC	3
ABD	1
ABE	6
ACD	8
ACE	-1
ADE	4
BCD	1
BCE	6
BDE	4
CDE	4
ABCD	4
ABCE	5
ABDE	5
ACDE	5
BCDE	5
ABCDE	6

Restricted Boltzmann Machine (RBM)

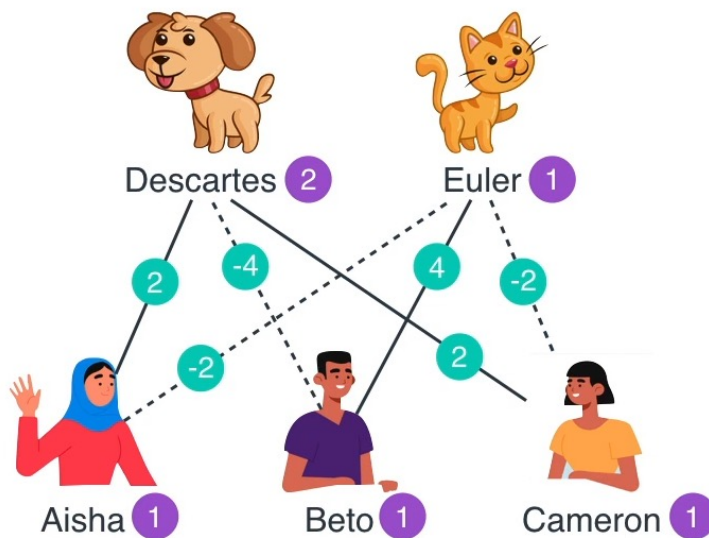


Ref:
<https://www.youtube.com/watch?v=Fkw0>

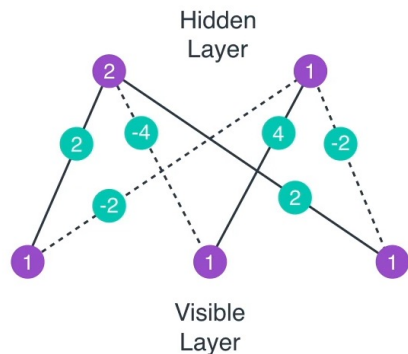
Scores to probabilities

Score	e^{score}	Normalize
1	$e^1 = 2.718$	0.665
0	$e^0 = 1$	0.245
-1	$e^{-1} = 0.368$	0.09
Sum = 4.086		Sum = 1

softmax



Energy to probability



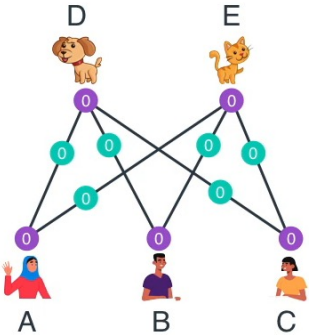
$$E = - \sum_i b_i v_i - \sum_i a_i h_i - \sum_{i,j} W_{ij} v_i h_j$$

$$p(v, h) = \frac{1}{Z} e^{-E(v, h)} \quad Z = \sum_{v, h} e^{-E(v, h)}$$

Z-Sum of all elements : softmax equation

Scenario	Score	e^{score}	Probability
None	0	1	0
A	1	2.72	0
B	1	2.72	0
C	1	2.72	0
D	2	7.38	0
E	1	2.72	0
AB	2	7.38	0
AC	2	7.38	0
AD	5	148.41	0.02
AE	0	2.72	0
BC	2	7.38	0
BD	-2	0.14	0
BE	7	1096.63	0.17
CD	5	148.41	0.02
CE	0	1	0
DE	3	20.08	0
ABC	3	20.08	0
ABD	1	2.72	0
ABE	6	403.43	0.06
ACD	8	2980.96	0.45
ACE	-1	0.37	0
ADE	4	54.6	0
BCD	1	2.72	0
BCE	6	403.43	0.06
BDE	4	54.6	0
CDE	4	54.6	0
ABCD	4	54.6	0.02
ABCE	5	148.41	0.02
ABDE	5	148.41	0.02
ACDE	5	148.41	0.02
BCDE	5	148.41	0.02
ABCDE	6	403.43	0.06

Train RBM



Scenario	Score	eScore	Probability
None	0	1	1/32
A	0	1	1/32
B	0	1	1/32
C	0	1	1/32
D	0	1	1/32
E	0	1	1/32
AB	0	1	1/32
AC	0	1	1/32
AD	0	1	1/32
AE	0	1	1/32
BC	0	1	1/32
BD	0	1	1/32
BE	0	1	1/32
CD	0	1	1/32
CE	0	1	1/32
DE	0	1	1/32

Scenario	Score	eScore	Probability
ABC	0	1	1/32
ABD	0	1	1/32
ABE	0	1	1/32
ACD	0	1	1/32
ACE	0	1	1/32
ADE	0	1	1/32
BCD	0	1	1/32
BCE	0	1	1/32
BDE	0	1	1/32
CDE	0	1	1/32
ABCD	0	1	1/32
ABCE	0	1	1/32
ABDE	0	1	1/32
ACDE	0	1	1/32
BCDE	0	1	1/32
ABCDE	0	1	1/32

Real data (scenarios)

Aisha	Beto	Cameron
✓	✗	✓
✗	✓	✗
✓	✗	✓
✓	✗	✓
✗	✓	✗
✓	✗	✓
✗	✓	✗
✓	✗	✓
✓	✗	✓
✗	✓	✗
✓	✗	✓

None		ABC	
A		ABD	
B		ABE	
C		ACD	
D		ACE	
E		ADE	
AB		BCD	
AC		BCE	
AD		BDE	
AE		CDE	
BC		ABCD	
BD		ABCE	
BE		ABDE	
CD		ACDE	
CE		BCDE	
DE		ABCDE	

Initial condition

Asha	Ben	Cameron
✓	✗	✓
✗	✓	✗

B and no A,C

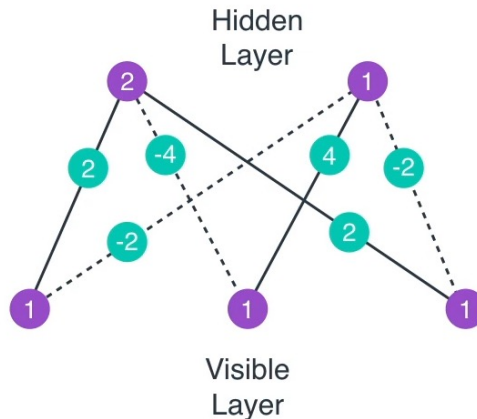


All scenarios

None	
A	
B	
C	
D	
E	
AB	
AC	
AD	
AE	
BC	
BD	
BE	
CD	
CE	
DE	

ABC	
ABD	
ABE	
ACD	
ACE	
ADE	
BCD	
BCE	
BDE	
CDE	
ABCD	
ABCE	
ABDE	
ACDE	
BCDE	
ABCDE	

Maximizing the probability of the data



$$\text{Find } \arg \max_w \prod_{v \in V} P(v)$$

$$\text{Maximize } \arg \max_w \mathbb{E}[\log P(v)]$$

$$\text{Derivative: } \frac{\partial}{\partial W} \log P(v_n)$$

$$= \mathbb{E} \left[\frac{\partial}{\partial W} - E(v, h) \mid v = v_n \right] - \mathbb{E} \left[\frac{\partial}{\partial W} - E(v, h) \right]$$

Criteria: Assign
high prob to
likelihood and
decrease to all
other

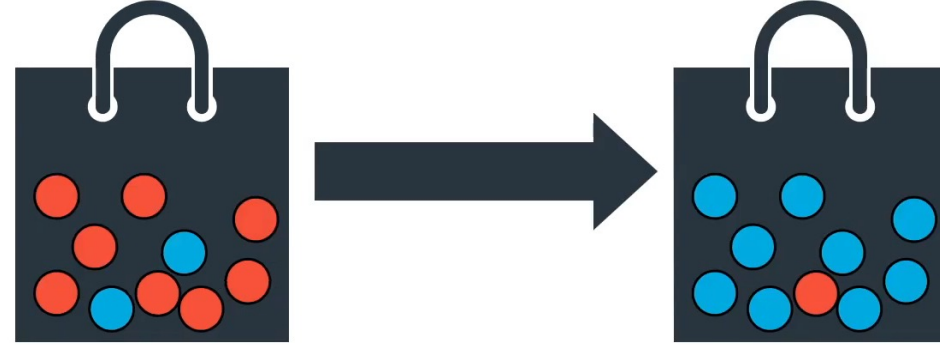
Problem

There are way too many possibilities!

$$32 = 2^5$$

Picking one from many possibility

None		ABC	
A		ABD	
B		ABE	
C		ACD	
D		ACE	
E		ADE	
AB		BCD	
AC		BCE	
AD		BDE	
AE		CDE	
BC		ABCD	
BD		ABCE	
BE		ABDE	
CD		ACDE	
CE		BCDE	
DE		ABCDE	



Solution: Gibbs sampler:
take out random red
and take in blue. Even
random take out blue. It
will same or little
improvement.



Aisha	Beto	Cameron
✓	✗	✓
✗	✓	✗

None ██████
 A ██████
 B ██████
 C ██████
 D ██████
 E ██████
 AB ██████
 AC ██████
 AD ██████
 AE ██████
 BC ██████
 BD ██████
 BE ██████
 CD ██████
 CE ██████
 DE ██████

ABC ██████
 ABD ██████
 ABE ██████
 ACD ██████
 ACE ██████
 ADE ██████
 BCD ██████
 BCE ██████
 BDE ██████
 CDE ██████
 ABCD ██████
 ABCE ██████
 ABDE ██████
 ↓ ACDE ██████
 BCDE ██████
 ABCDE ██████



Aisha	Beto	Cameron
✓	✗	✓
✗	✓	✗
✓	✗	✓
✓	✗	✓
✗	✓	✗
✓	✗	✓
✗	✓	✗
✓	✗	✓
✓	✗	✓

None |
 A |
 B |
 C |
 D |
 E |
 AB |
 AC |
 AD |
 AE |
 BC |
 BD |
 BE ██████
 CD |
 CE |
 DE |

ABC |
 ABD |
 ABE |
 ACD ██████
 ACE |
 ADE |
 BCD |
 BCE |
 BDE |
 CDE |
 ABCD |
 ABCE |
 ABDE |
 ACDE |
 BCDE |
 ABCDE |

Choose random one to decrease instead of all

Aisha	Beto	Cameron
✓	✗	✓
✗	✓	✗
✓	✗	✓
✓	✗	✓
✗	✓	✗
✓	✗	✓
✗	✓	✗
✓	✗	✓
✓	✗	✓
✓	✗	✓

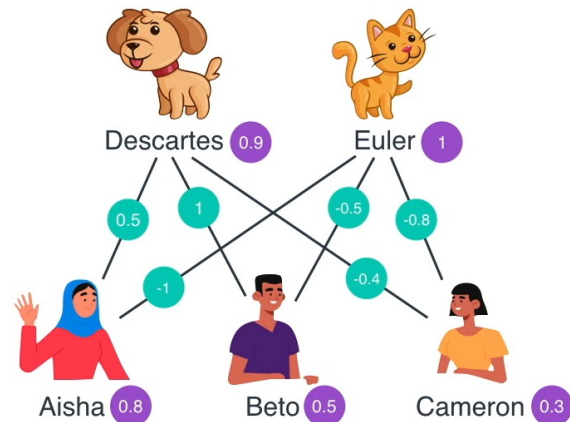
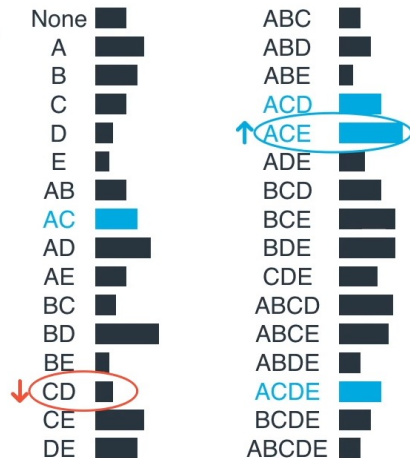
None |
 A |
 B | Beto
 C |
 D |
 E |
 AB |
 AC | Aisha, Cameron
 AD |
 AE |
 BC |
 BD | Beto
 BE ██████ Beto
 CD |
 CE |
 DE |

ABC |
 ABD |
 ABE |
 ACD ██████ Aisha, Cameron
 ACE | Aisha, Cameron
 ADE | Aisha, Cameron
 BCD |
 BCE |
 BDE | Beto
 ABCD |
 ABCE |
 ABDE |
 ACDE ██████ Aisha, Cameron
 BCDE |
 ABCDE |

Picking random items from here is really hard!

How to pick a random **one** with conditions

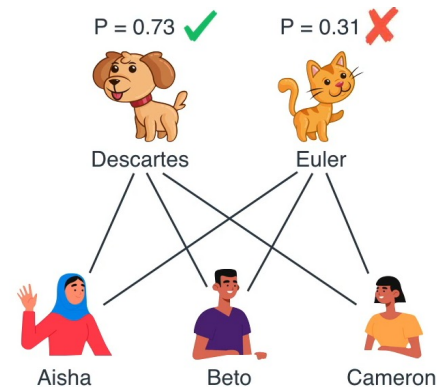
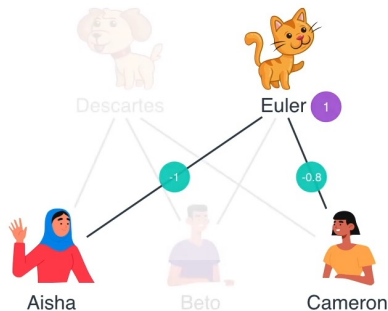
How to pick a completely random **one**



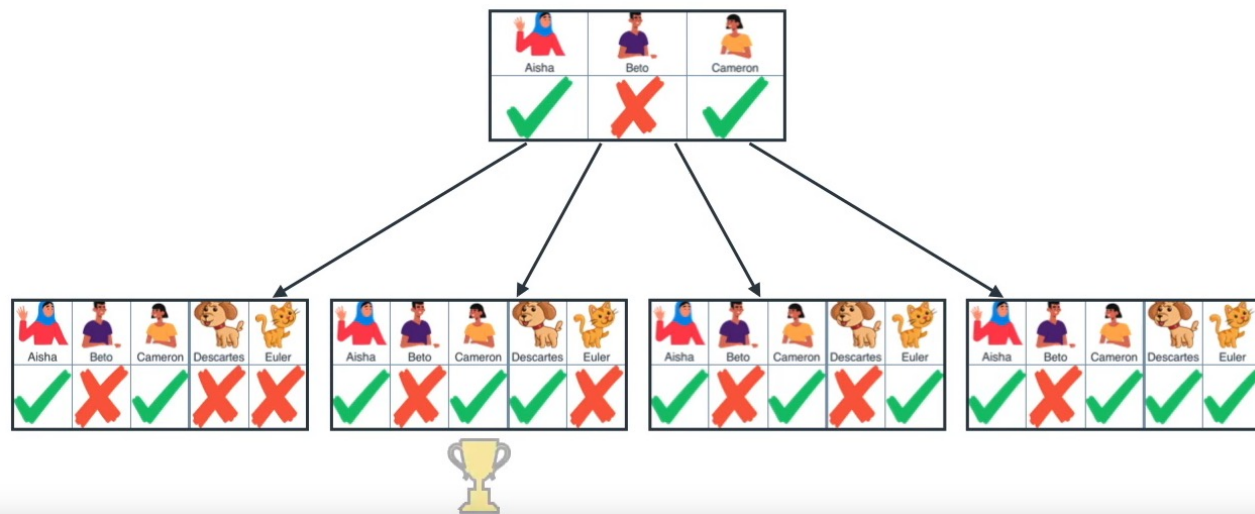
Independent sampling:

$$P(\text{Dog}) = \sigma(1) = 0.73$$

$$P(\text{Cat}) = \sigma(-0.8) = 0.31$$



Gibbs Sampling



Select 1 out of 4
using independent
sampling

Choose completely random



How to pick a random **one** with conditions

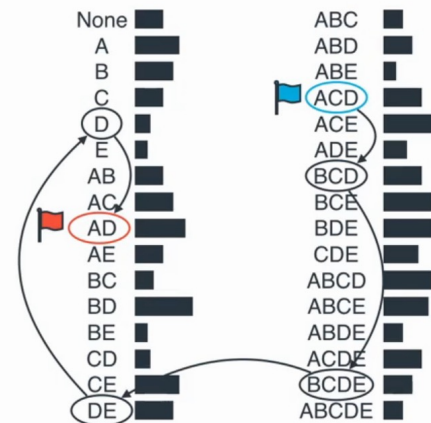
How to pick a completely random **one**



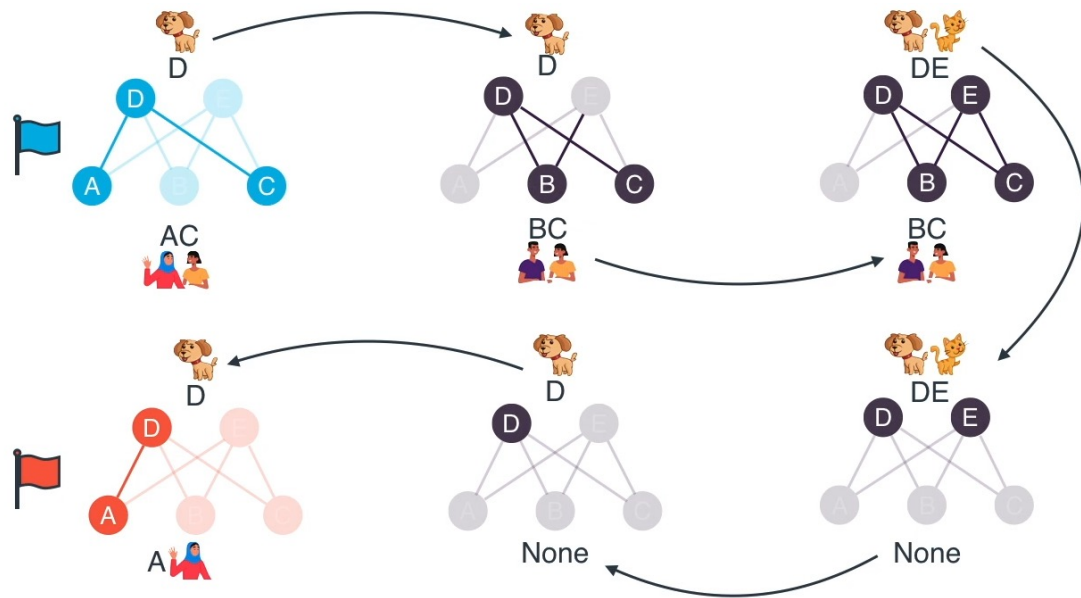
Gibbs Sampling



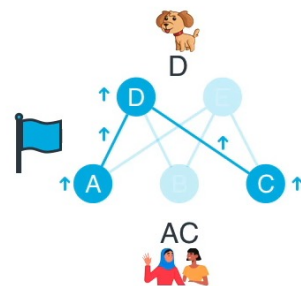
How to pick a totally random sample from this distribution



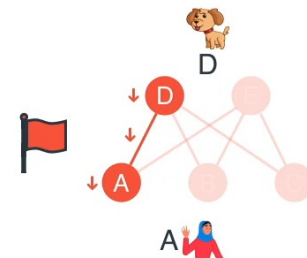
no. steps is random hyperparameter



Increase scores



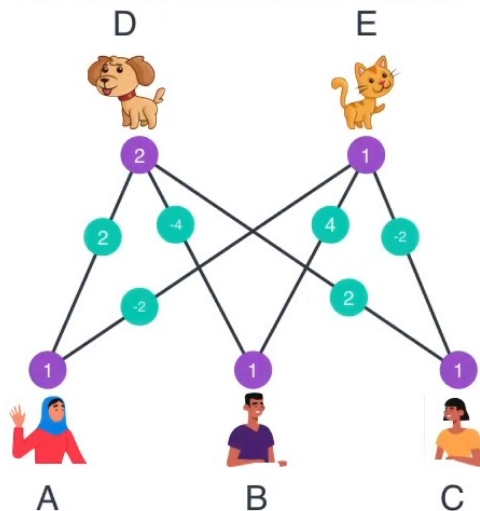
Decrease scores



Data

Aisha	Beto	Cameron
✓	✗	✓
✗	✓	✗
✓	✗	✓
✓	✗	✓
✗	✓	✗
✓	✗	✓
✗	✓	✗
✓	✗	✓
✓	✗	✓

Restricted Boltzmann Machine



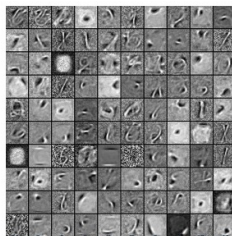
Probabilities

None	ABC
A	ABD
B	ABE
C	ACD
D	ACE
E	ADE
AB	BCD
AC	BCE
AD	BDE
AE	CDE
BC	ABCD
BD	ABCE
BE	ABDE
CD	ACDE
CE	BCDE
DE	ABCDE

Generated Data

Aisha	Beto	Cameron
✓	✗	✓
✗	✓	✗
✗	✗	✗
✗	✓	✗
✓	✗	✓
✓	✗	✓
✓	✗	✓
✓	✗	✓
✓	✗	✓
✓	✗	✓

Images



Hidden Layer



Visible Layer

Application:

Thank you