Machine Learning for Data Science (CS4786)
Lecture 23

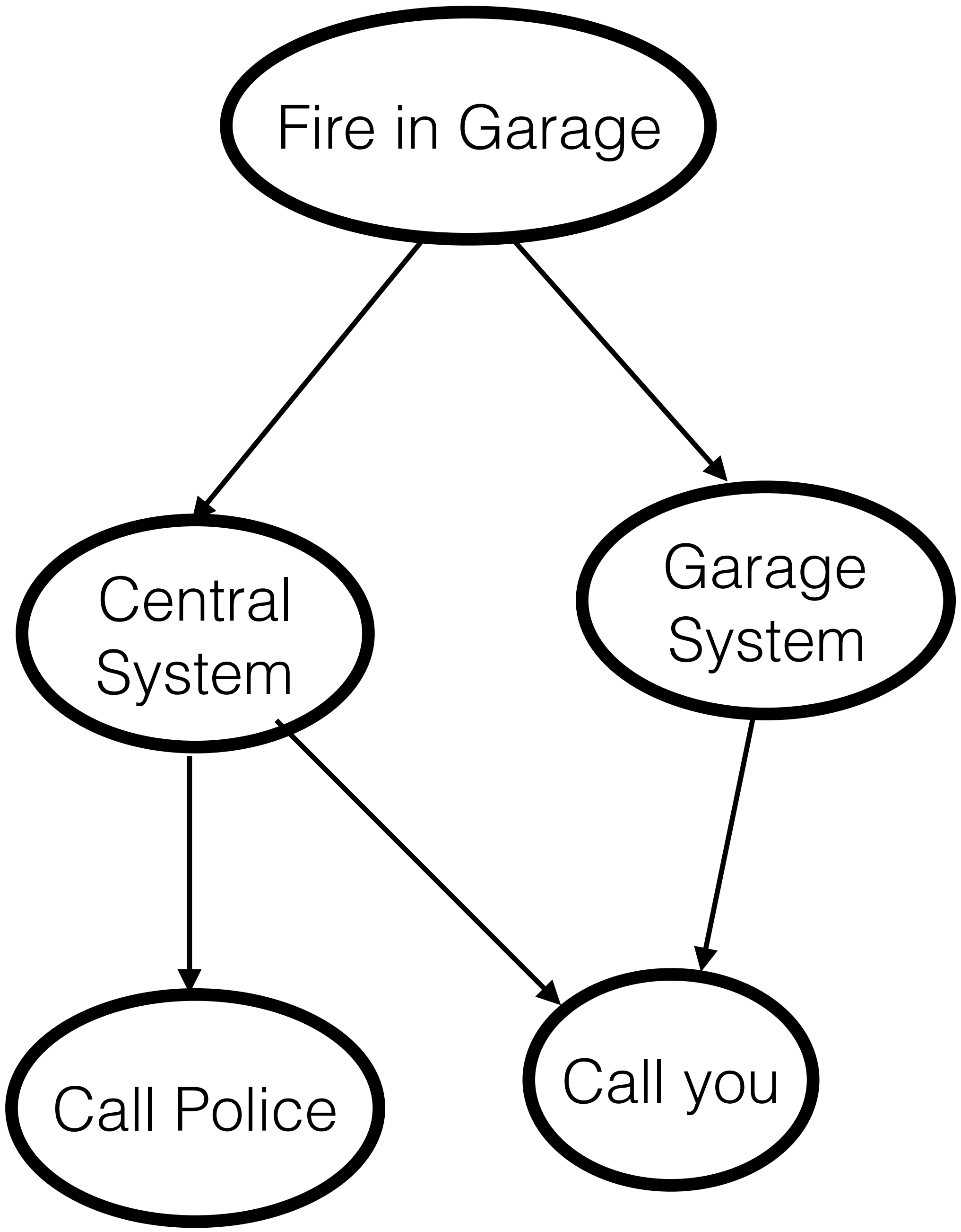
Message Passing and Learning in Graphical Models

Course Webpage :
http://www.cs.cornell.edu/Courses/cs4786/2016fa/

Announcements

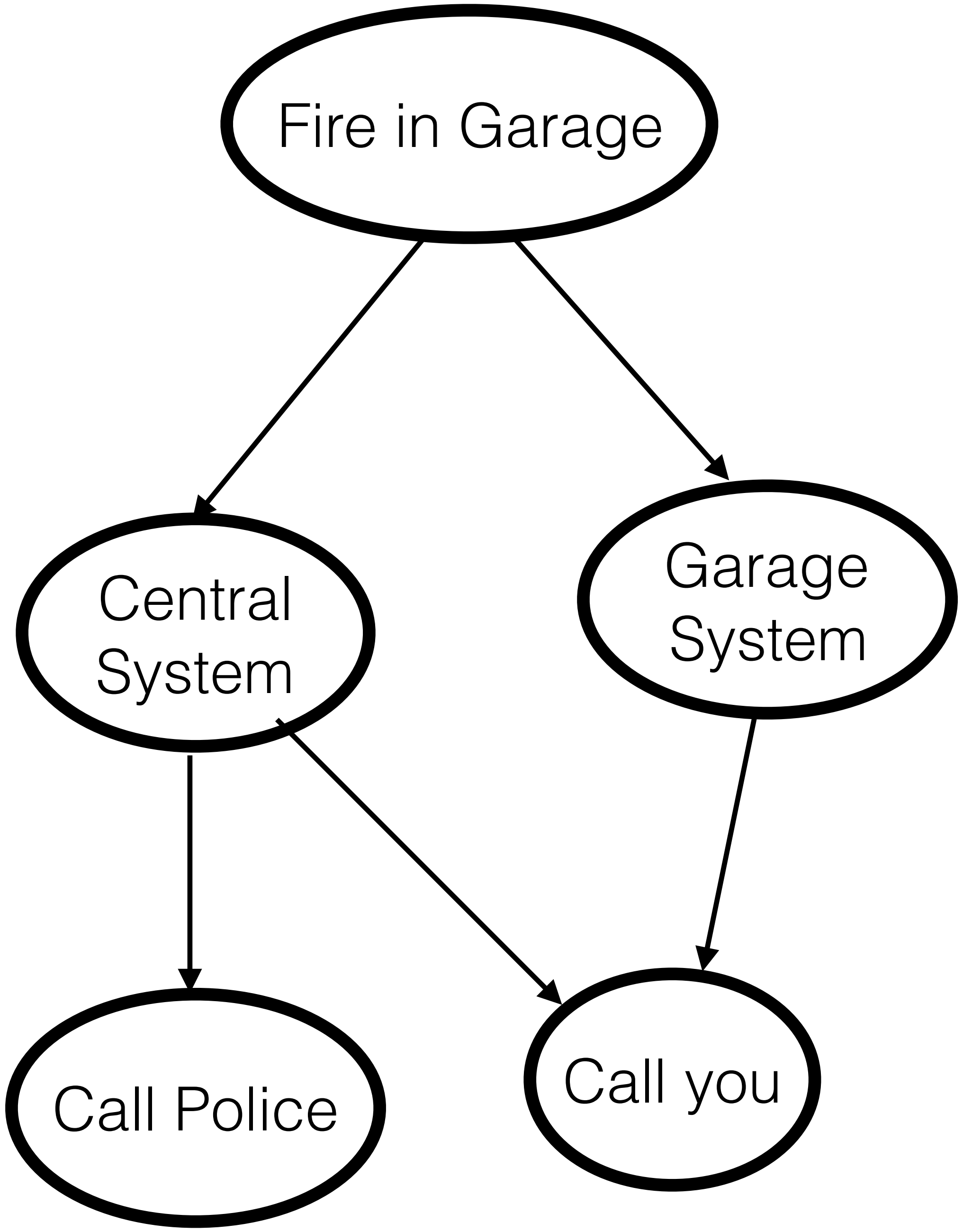
- Survey 2 is out, due by Nov 26th
- Topics lecture after thanksgiving on
  “Fair and unbiased in machine learning”

Rejection Sampling



Rejection Sampling

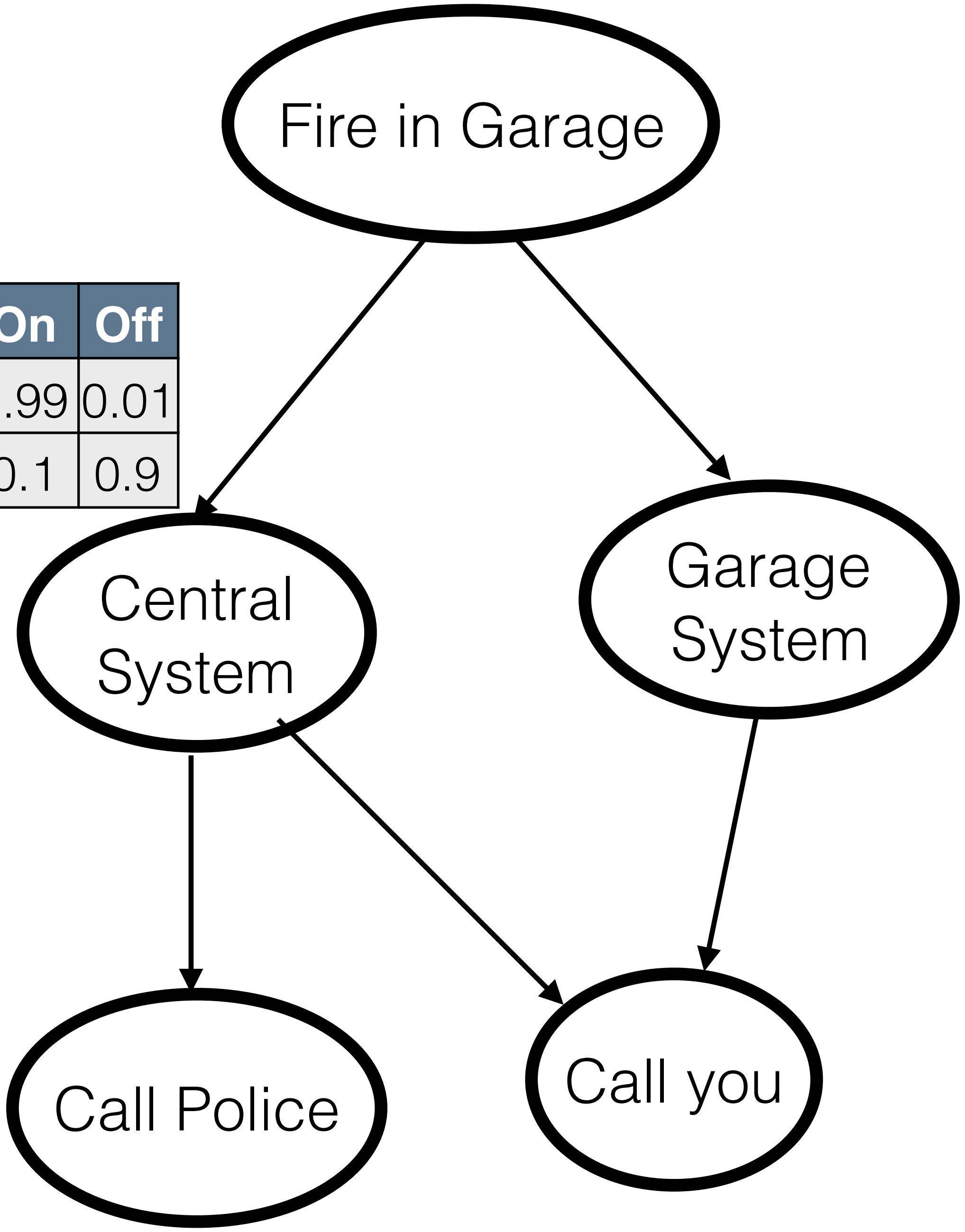
Fire	No Fire
0.05	0.95



Rejection Sampling

Fire	No Fire
0.05	0.95

CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

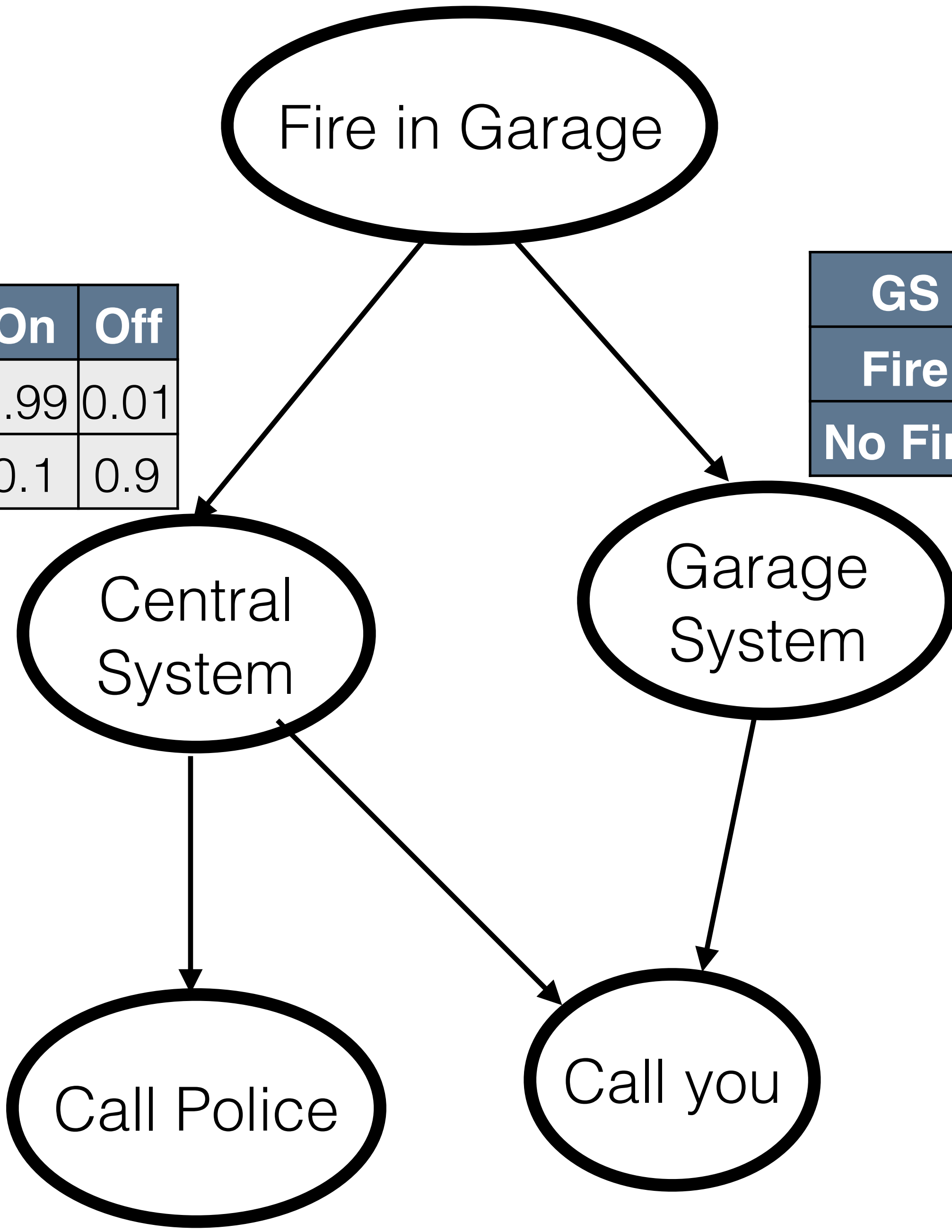


REJECTION SAMPLING

Fire	No Fire
0.05	0.95

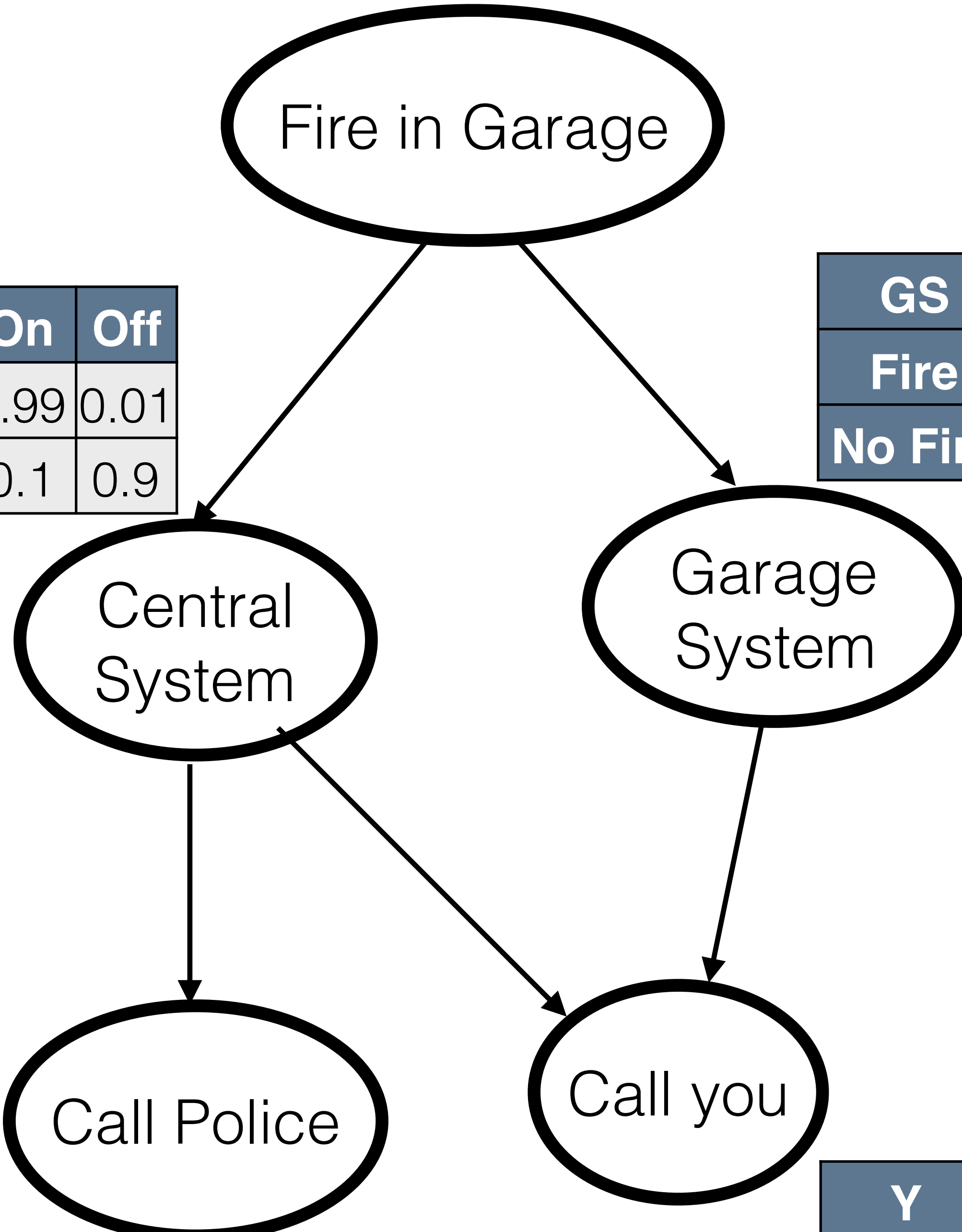
CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7



REJECTION SAMPLING

Fire	No Fire
0.05	0.95



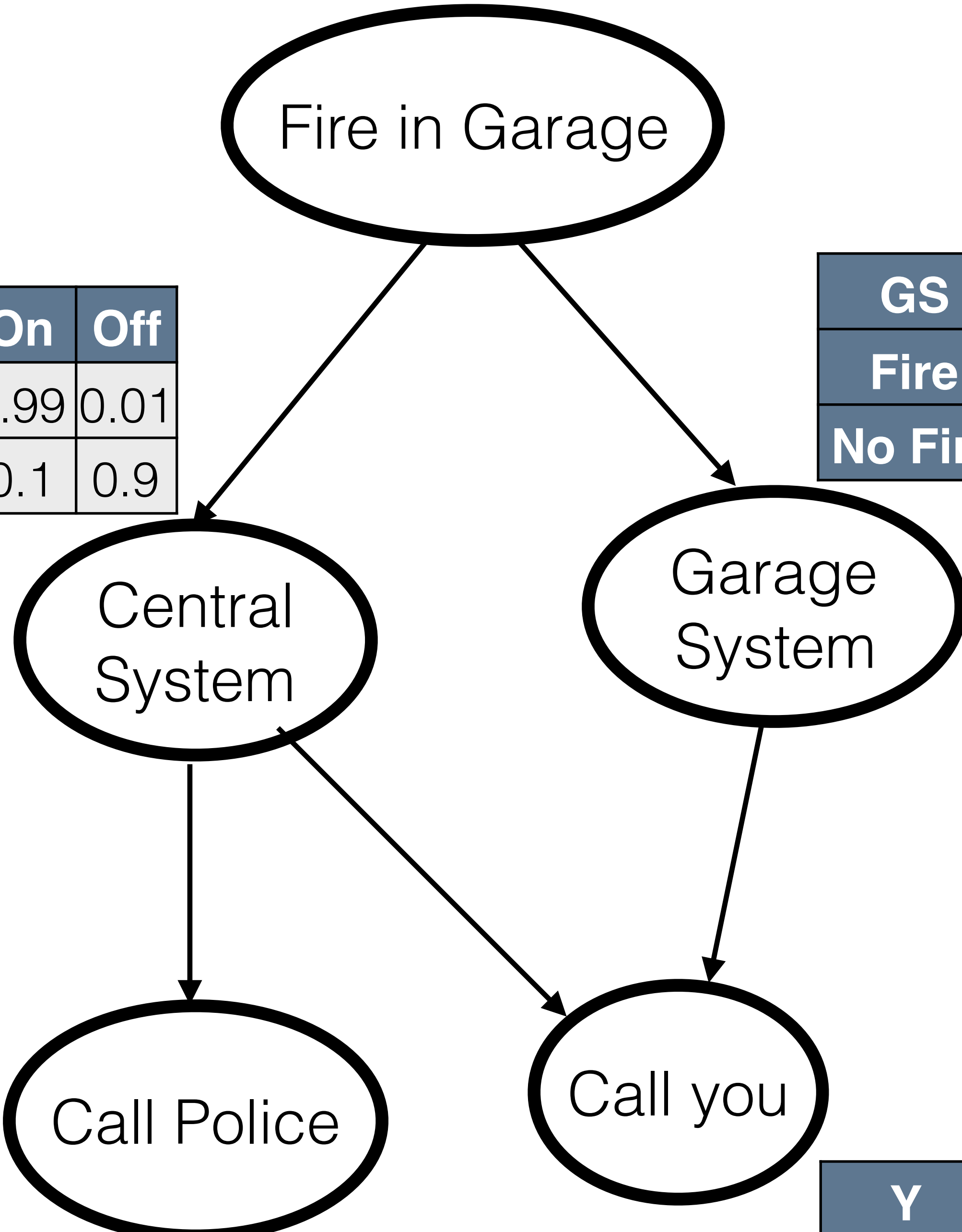
CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

Rejection Sampling

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

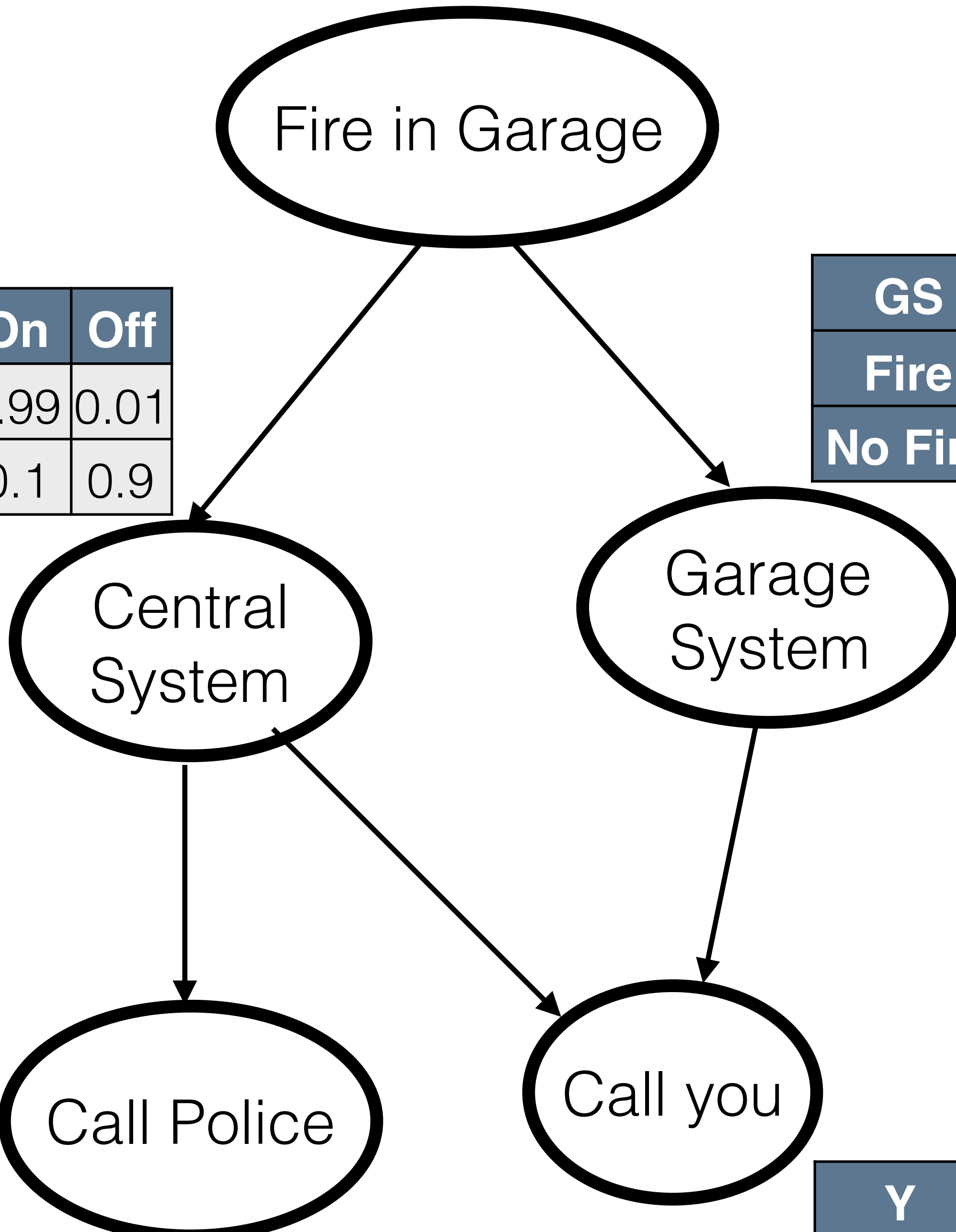
GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

Rejection Sampling

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

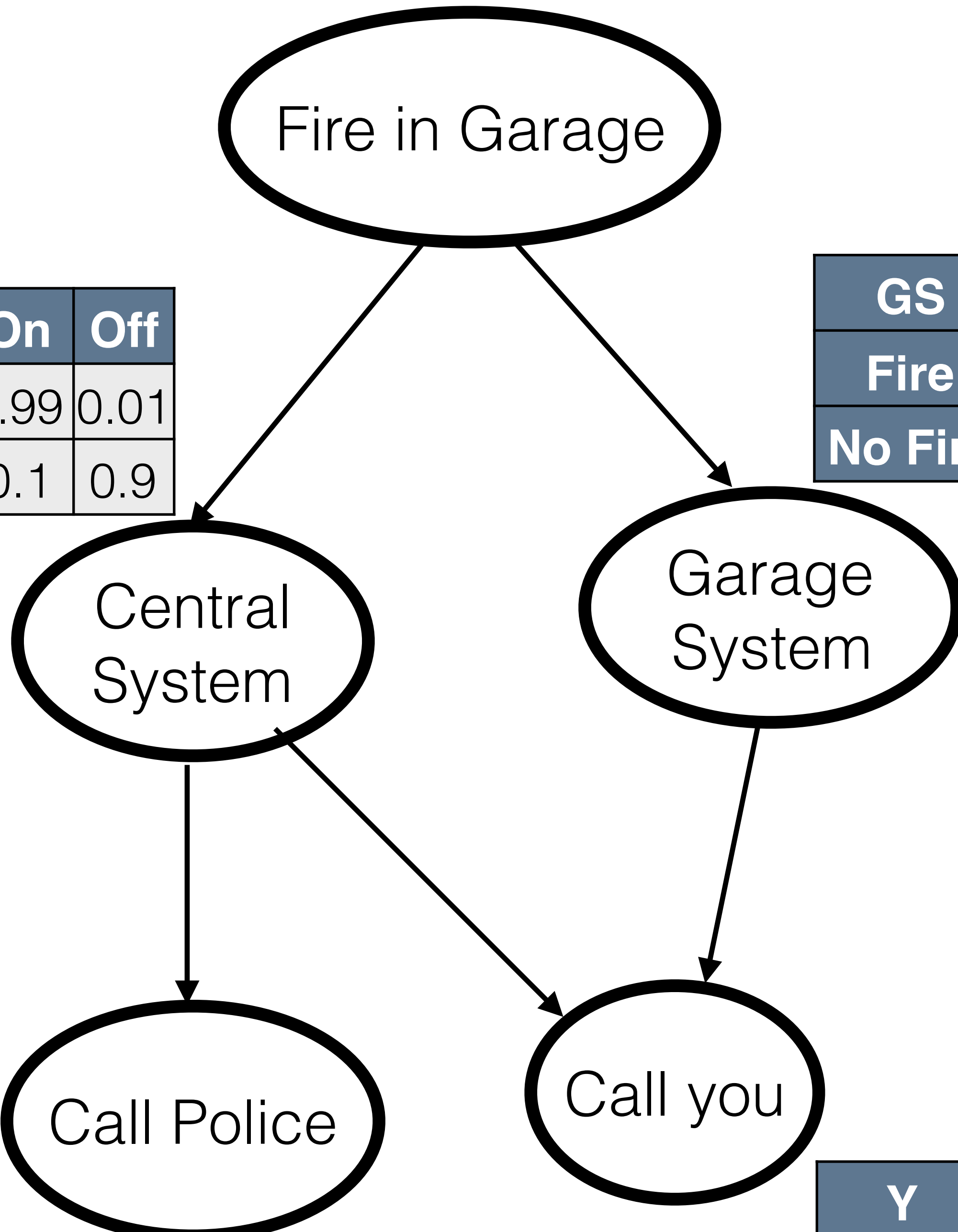
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

	F	CS	GS	P	Y

Rejection Sampling

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

P	Call	no call
CS on	1	0
CS off	0.01	0.99

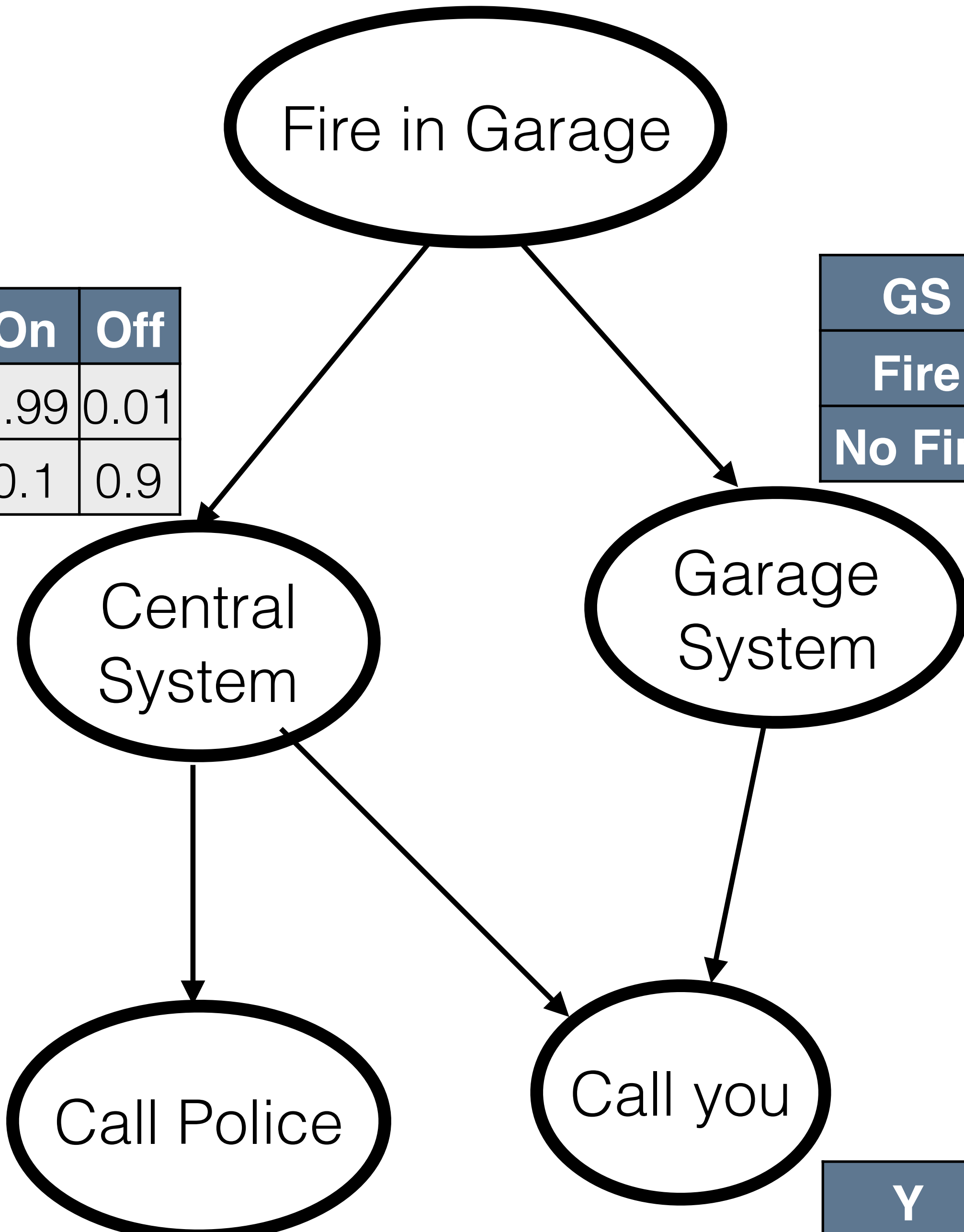
Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

	F	CS	GS	P	Y
1	0				

Rejection Sampling

Fire	No Fire
0.05	0.95

	F	CS	GS	P	Y
1	0	0			



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

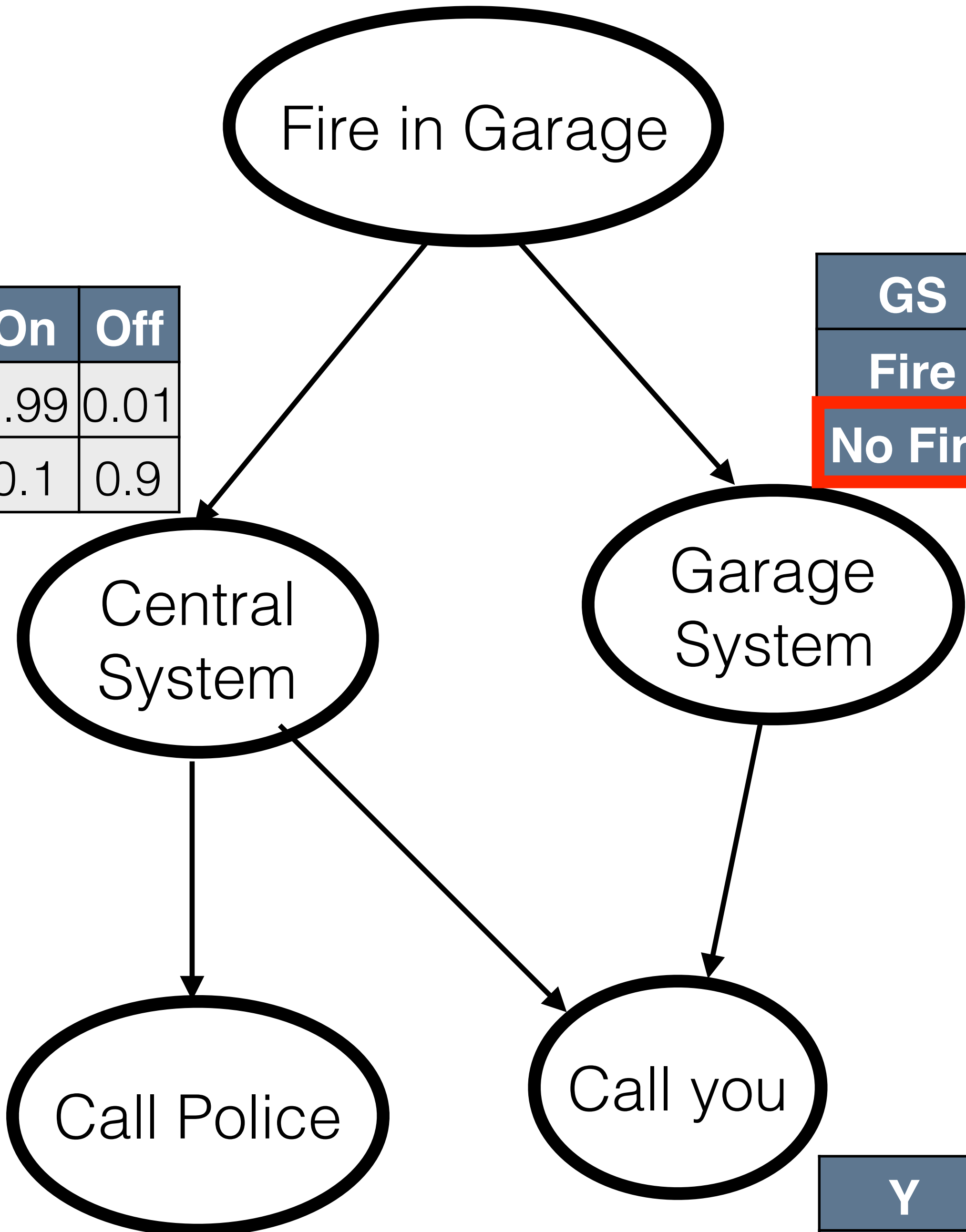
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

REJECTION SAMPLING

	F	CS	GS	P	Y
1	0	0	1		

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

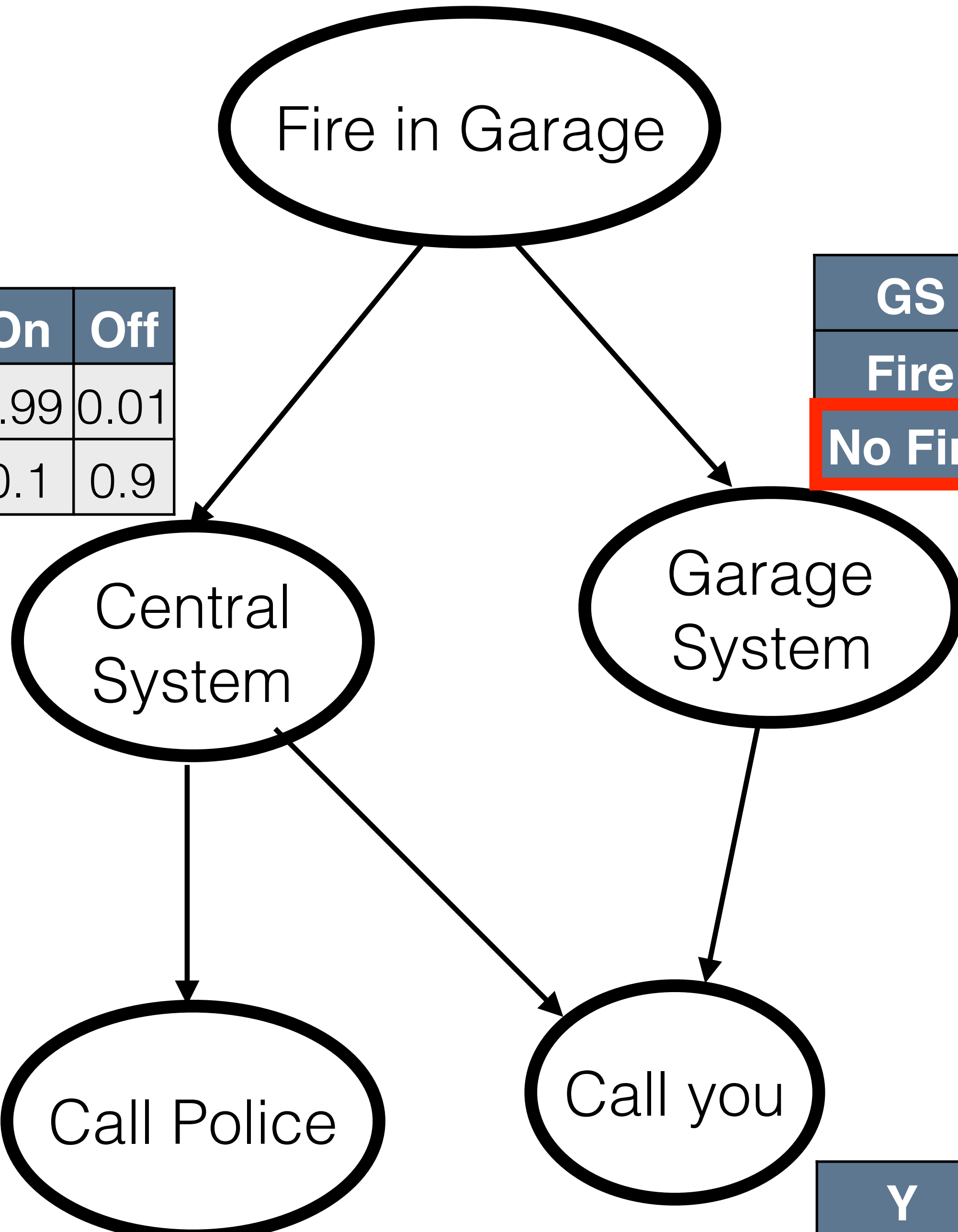
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

Rejection Sampling

	F	CS	GS	P	Y
1	0	0	1	0	

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

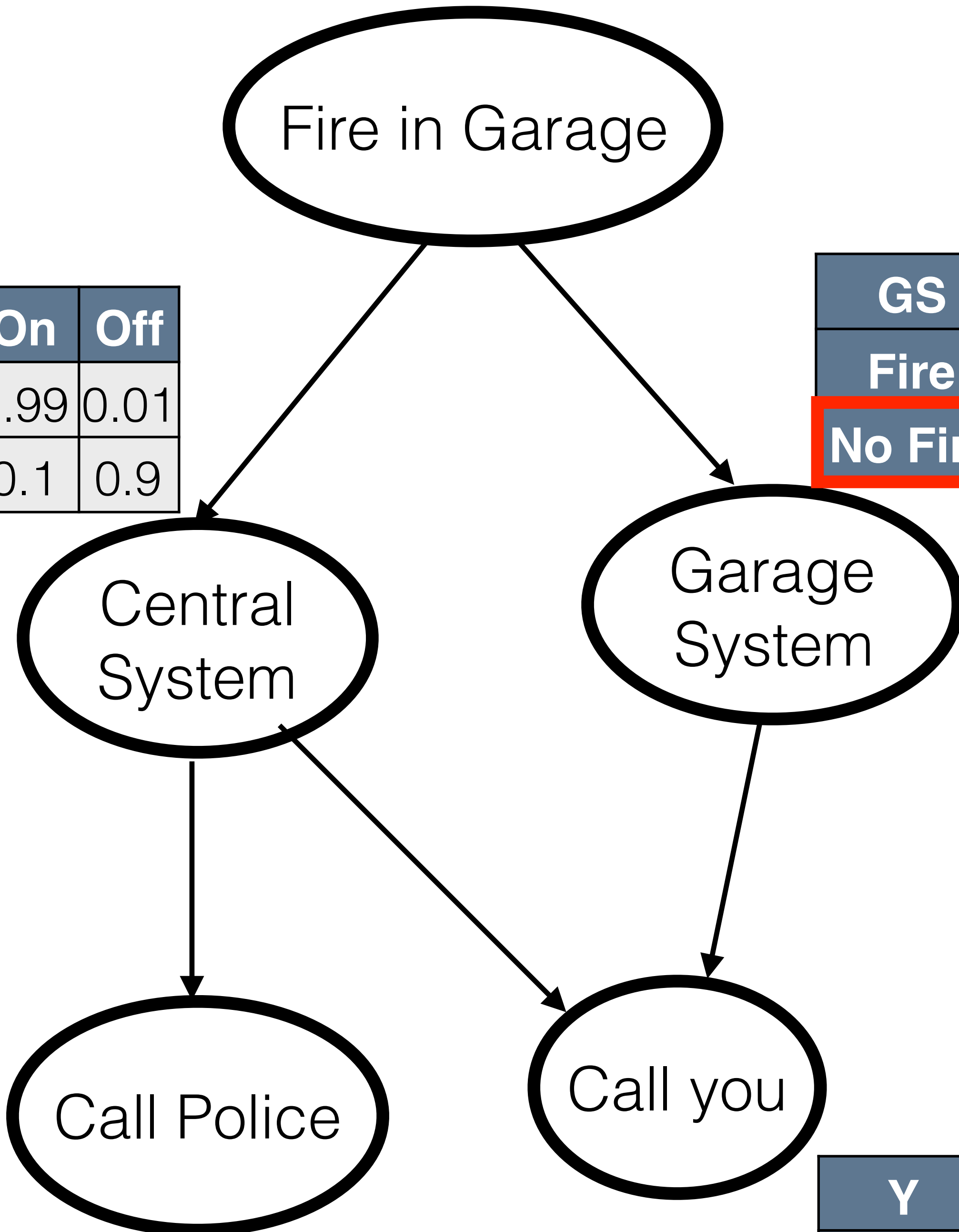
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

Rejection Sampling

	F	CS	GS	P	Y
1	0	0	1	0	1

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

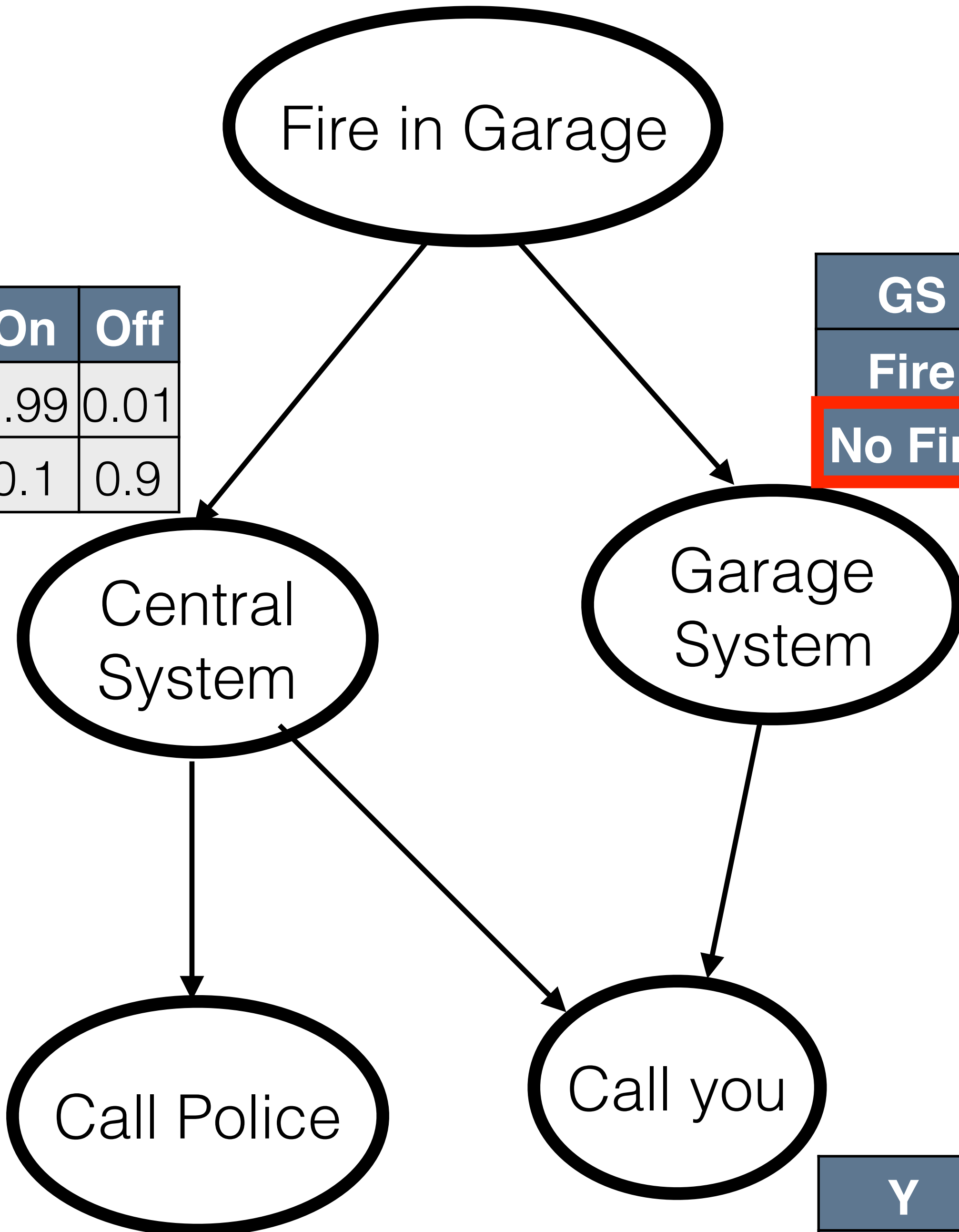
GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

Rejection Sampling

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

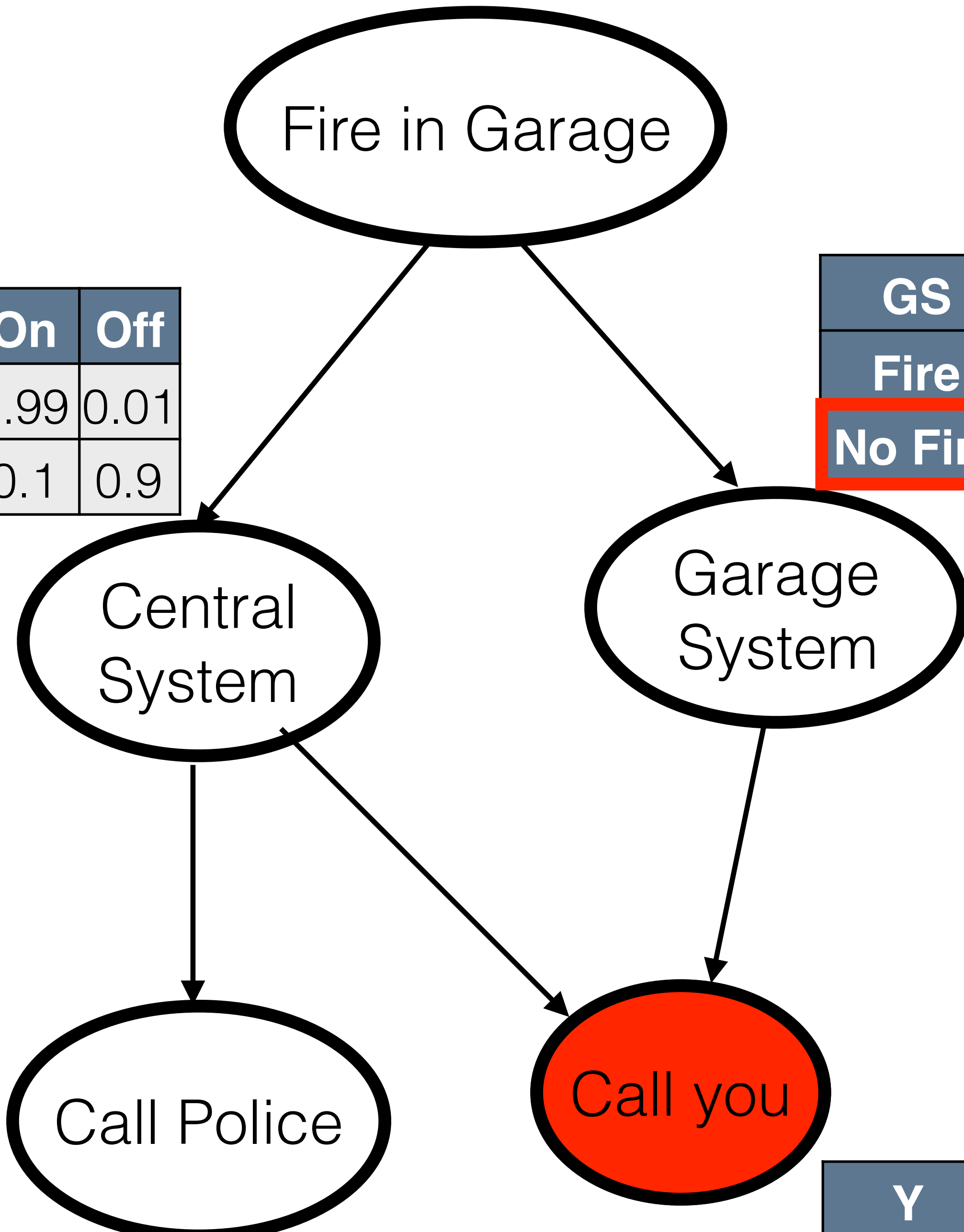
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

	F	CS	GS	P	Y
1	0	0	1	0	1
2	0	1	0	1	0
3	1	1	1	1	1
4	0	0	0	0	0
5	0	0	1	0	0
6	0	0	1	0	1
7	0	0	0	0	0
8	0	0	1	0	1
9	0	0	0	0	0
10	0	0	1	0	1
11	0	0	1	0	0
12	0	0	1	0	1
13	0	0	1	0	1
14	0	0	0	0	0
15	0	0	0	0	0
16	0	0	1	0	0
17	0	0	0	0	0
18	0	0	0	0	0
19	0	0	0	0	0
20	0	0	0	0	0
21	0	0	0	0	0
22	0	0	0	0	0
23	0	0	0	0	0
24	0	0	0	0	0
25	0	0	1	0	1

Rejection Sampling

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

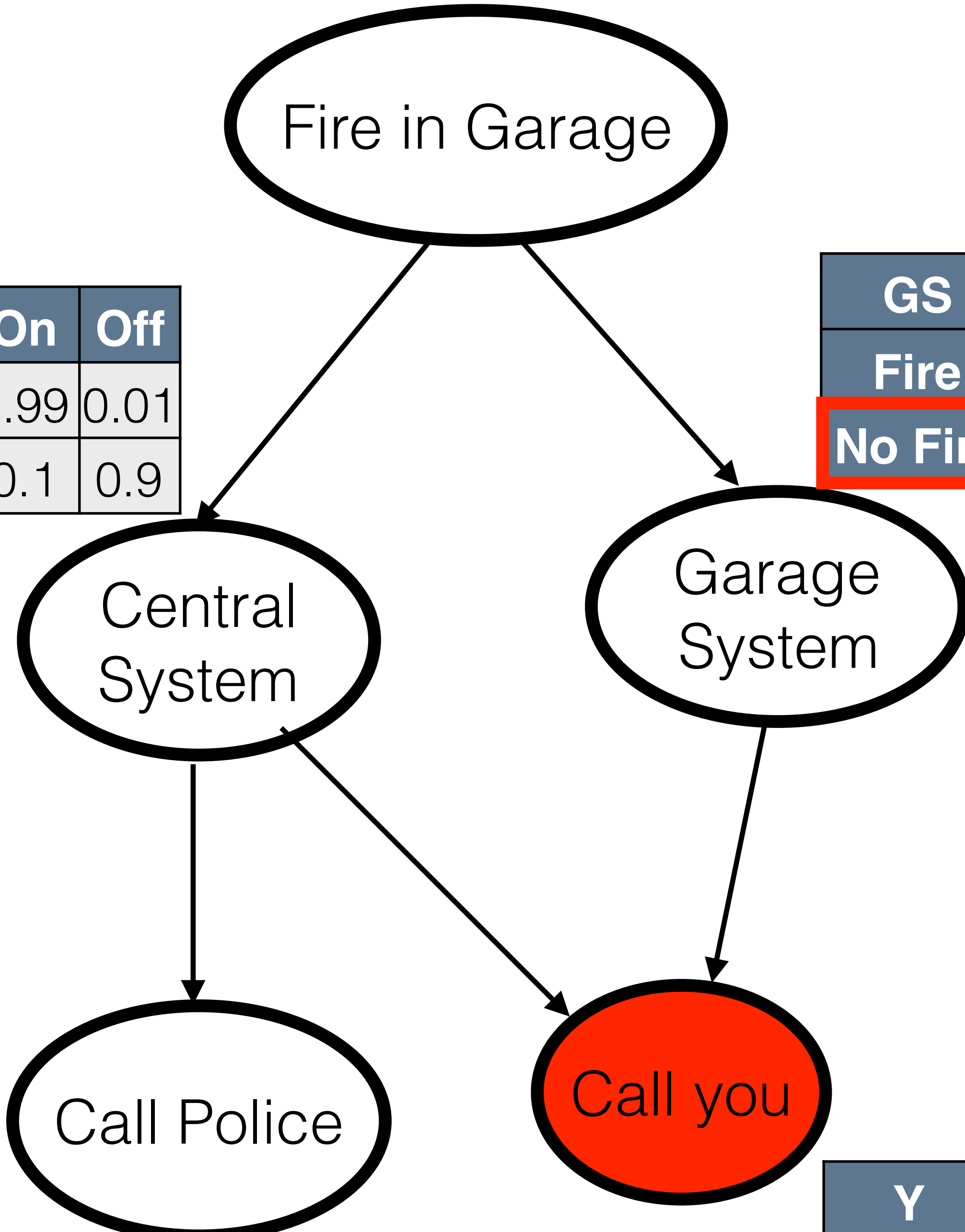
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

	F	CS	GS	P	Y
1	0	0	1	0	1
2	0	1	0	1	0
3	1	1	1	1	1
4	0	0	0	0	0
5	0	0	1	0	0
6	0	0	1	0	1
7	0	0	0	0	0
8	0	0	1	0	1
9	0	0	0	0	0
10	0	0	1	0	1
11	0	0	1	0	0
12	0	0	1	0	1
13	0	0	1	0	1
14	0	0	0	0	0
15	0	0	0	0	0
16	0	0	1	0	0
17	0	0	0	0	0
18	0	0	0	0	0
19	0	0	0	0	0
20	0	0	0	0	0
21	0	0	0	0	0
22	0	0	0	0	0
23	0	0	0	0	0
24	0	0	0	0	0
25	0	0	1	0	1

Rejection Sampling

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

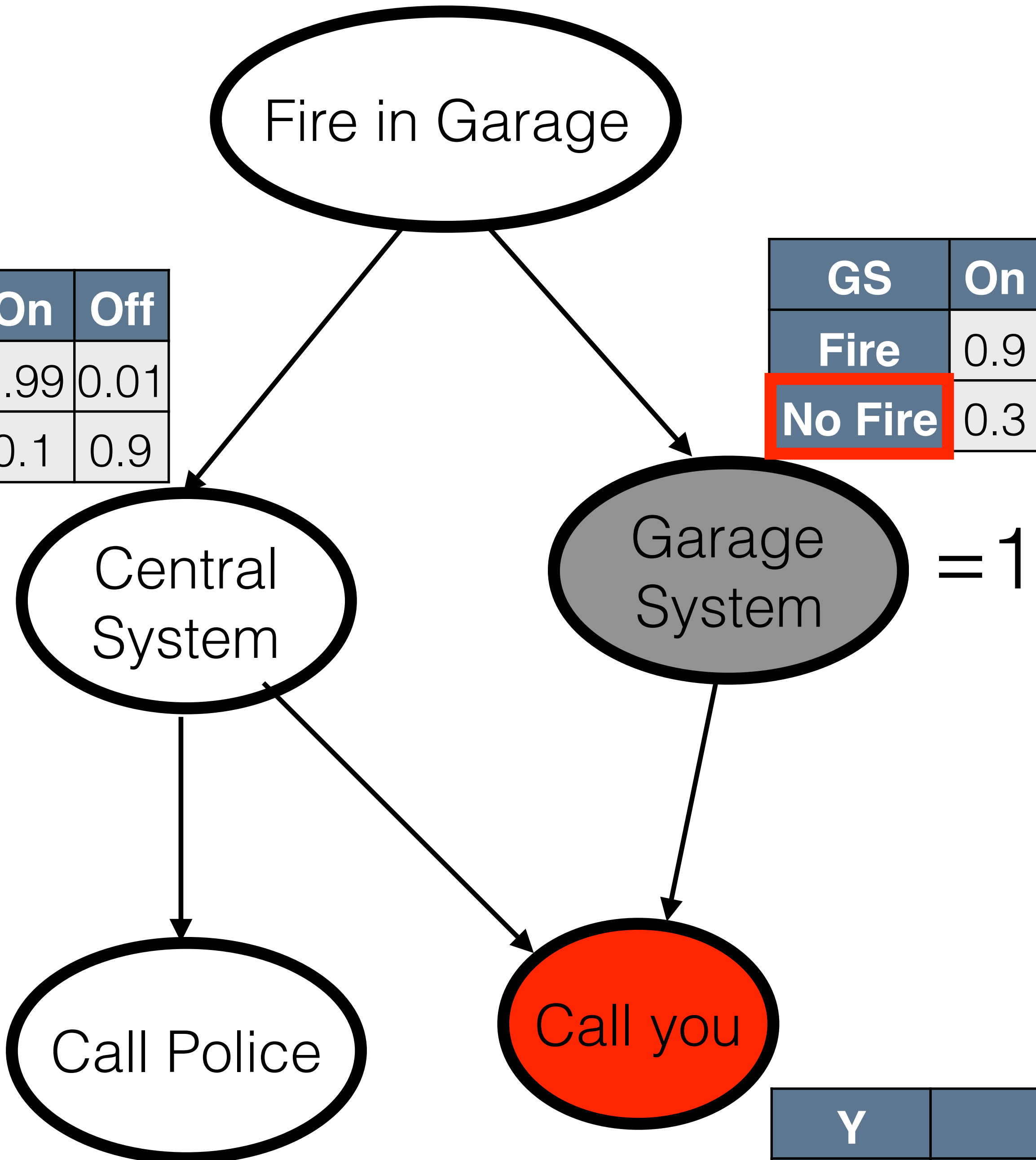
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

	F	CS	GS	P	Y
1	0	0	1	0	1
2	0	1	0	1	0
3	1	1	1	1	1
4	0	0	0	0	0
5	0	0	1	0	0
6	0	0	1	0	1
7	0	0	0	0	0
8	0	0	1	0	1
9	0	0	0	0	0
10	0	0	1	0	1
11	0	0	1	0	0
12	0	0	1	0	1
13	0	0	1	0	1
14	0	0	0	0	0
15	0	0	0	0	0
16	0	0	1	0	0
17	0	0	0	0	0
18	0	0	0	0	0
19	0	0	0	0	0
20	0	0	0	0	0
21	0	0	0	0	0
22	0	0	0	0	0
23	0	0	0	0	0
24	0	0	0	0	0
25	0	0	1	0	1

Rejection Sampling

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

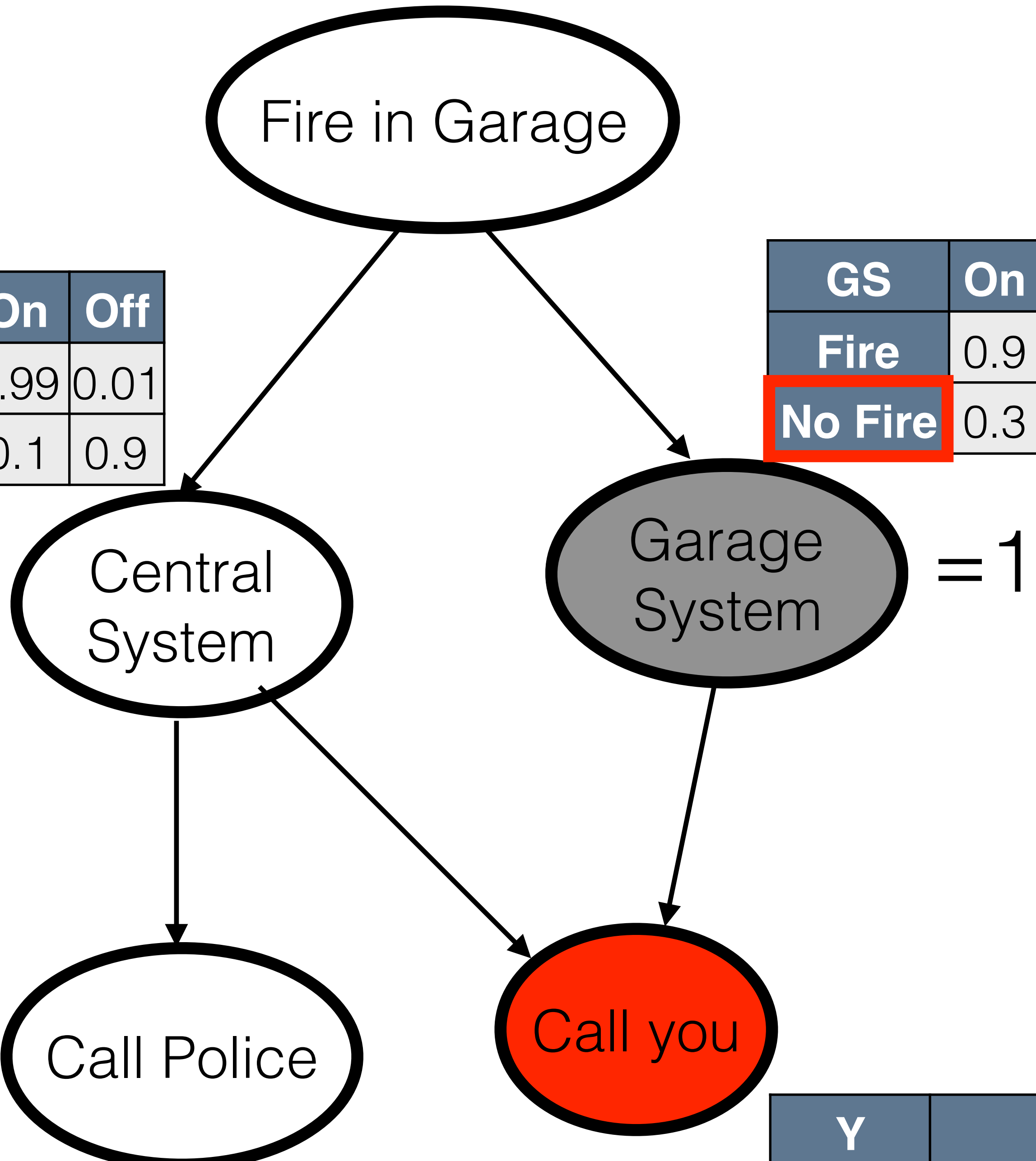
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

	F	CS	GS	P	Y
1	0	0	1	0	1
2	0	1	0	1	0
3	1	1	1	1	1
4	0	0	0	0	0
5	0	0	1	0	0
6	0	0	1	0	1
7	0	0	0	0	0
8	0	0	1	0	1
9	0	0	0	0	0
10	0	0	1	0	1
11	0	0	1	0	0
12	0	0	1	0	1
13	0	0	1	0	1
14	0	0	0	0	0
15	0	0	0	0	0
16	0	0	1	0	0
17	0	0	0	0	0
18	0	0	0	0	0
19	0	0	0	0	0
20	0	0	0	0	0
21	0	0	0	0	0
22	0	0	0	0	0
23	0	0	0	0	0
24	0	0	0	0	0
25	0	0	1	0	1

Rejection Sampling

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

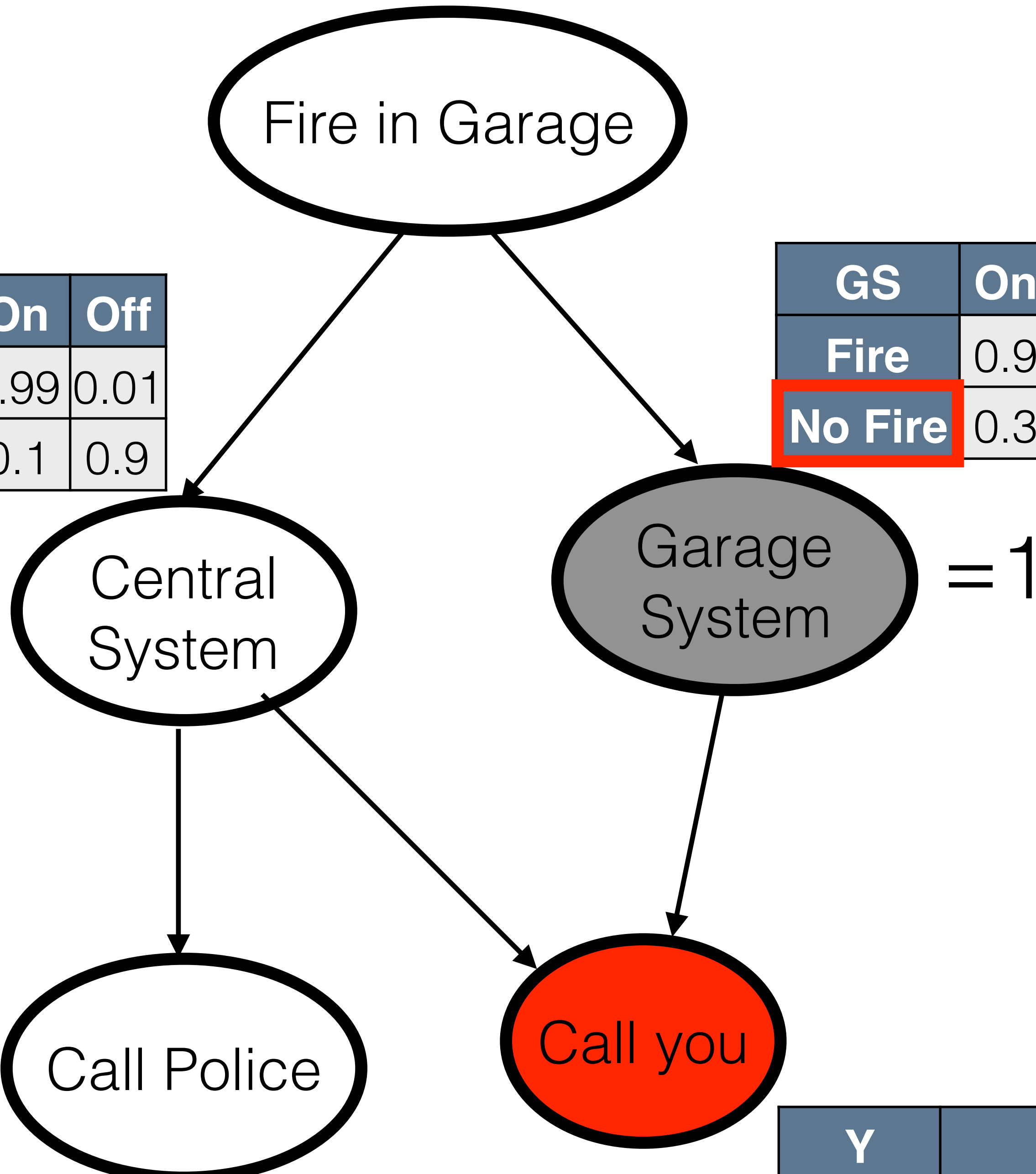
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

	F	CS	GS	P	Y
1	0	0	1	0	1
2	0	1	0	1	0
3	1	1	1	1	1
4	0	0	0	0	0
5	0	0	1	0	0
6	0	0	1	0	1
7	0	0	0	0	0
8	0	0	1	0	1
9	0	0	0	0	0
10	0	0	1	0	1
11	0	0	1	0	0
12	0	0	1	0	1
13	0	0	1	0	1
14	0	0	0	0	0
15	0	0	0	0	0
16	0	0	1	0	0
17	0	0	0	0	0
18	0	0	0	0	0
19	0	0	0	0	0
20	0	0	0	0	0
21	0	0	0	0	0
22	0	0	0	0	0
23	0	0	0	0	0
24	0	0	0	0	0
25	0	0	1	0	1

Rejection Sampling

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

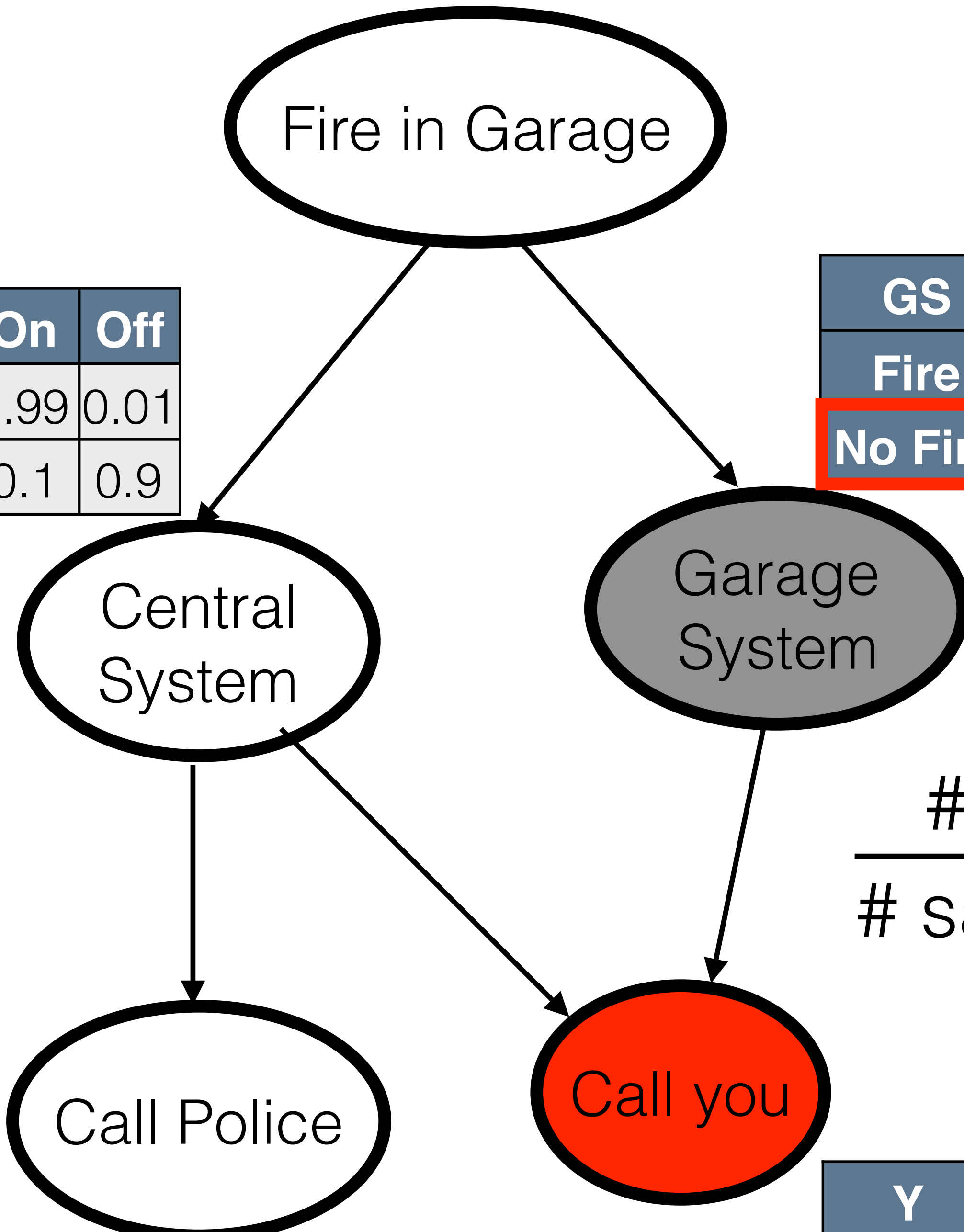
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

	F	CS	GS	P	Y
1	0	0	1	0	1
2	0	1	0	1	0
3	1	1	1	1	1
4	0	0	0	0	0
5	0	0	1	0	0
6	0	0	1	0	1
7	0	0	0	0	0
8	0	0	1	0	1
9	0	0	0	0	0
10	0	0	1	0	1
11	0	0	1	0	0
12	0	0	1	0	1
13	0	0	1	0	1
14	0	0	0	0	0
15	0	0	0	0	0
16	0	0	1	0	0
17	0	0	0	0	0
18	0	0	0	0	0
19	0	0	0	0	0
20	0	0	0	0	0
21	0	0	0	0	0
22	0	0	0	0	0
23	0	0	0	0	0
24	0	0	0	0	0
25	0	0	1	0	1

Rejection Sampling

Fire	No Fire
0.05	0.95



$=1$

CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

$$\frac{\#\text{ calls}}{\#\text{ samples}} = \frac{8}{11}$$

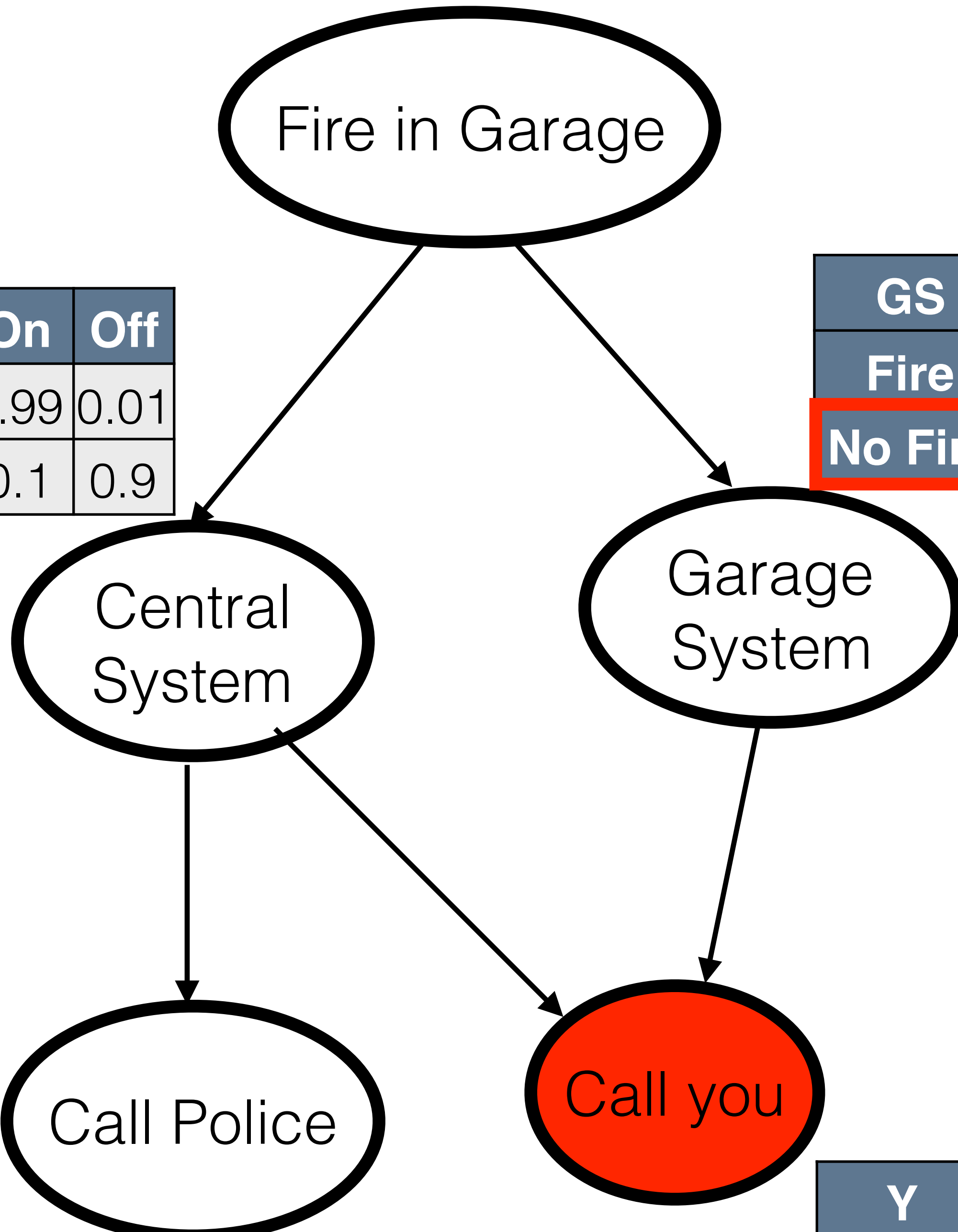
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

	F	CS	GS	P	Y
1	0	0	1	0	1
2	0	1	0	1	0
3	1	1	1	1	1
4	0	0	0	0	0
5	0	0	1	0	0
6	0	0	1	0	1
7	0	0	0	0	0
8	0	0	1	0	1
9	0	0	0	0	0
10	0	0	1	0	1
11	0	0	1	0	0
12	0	0	1	0	1
13	0	0	1	0	1
14	0	0	0	0	0
15	0	0	0	0	0
16	0	0	1	0	0
17	0	0	0	0	0
18	0	0	0	0	0
19	0	0	0	0	0
20	0	0	0	0	0
21	0	0	0	0	0
22	0	0	0	0	0
23	0	0	0	0	0
24	0	0	0	0	0
25	0	0	1	0	1

Rejection Sampling

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

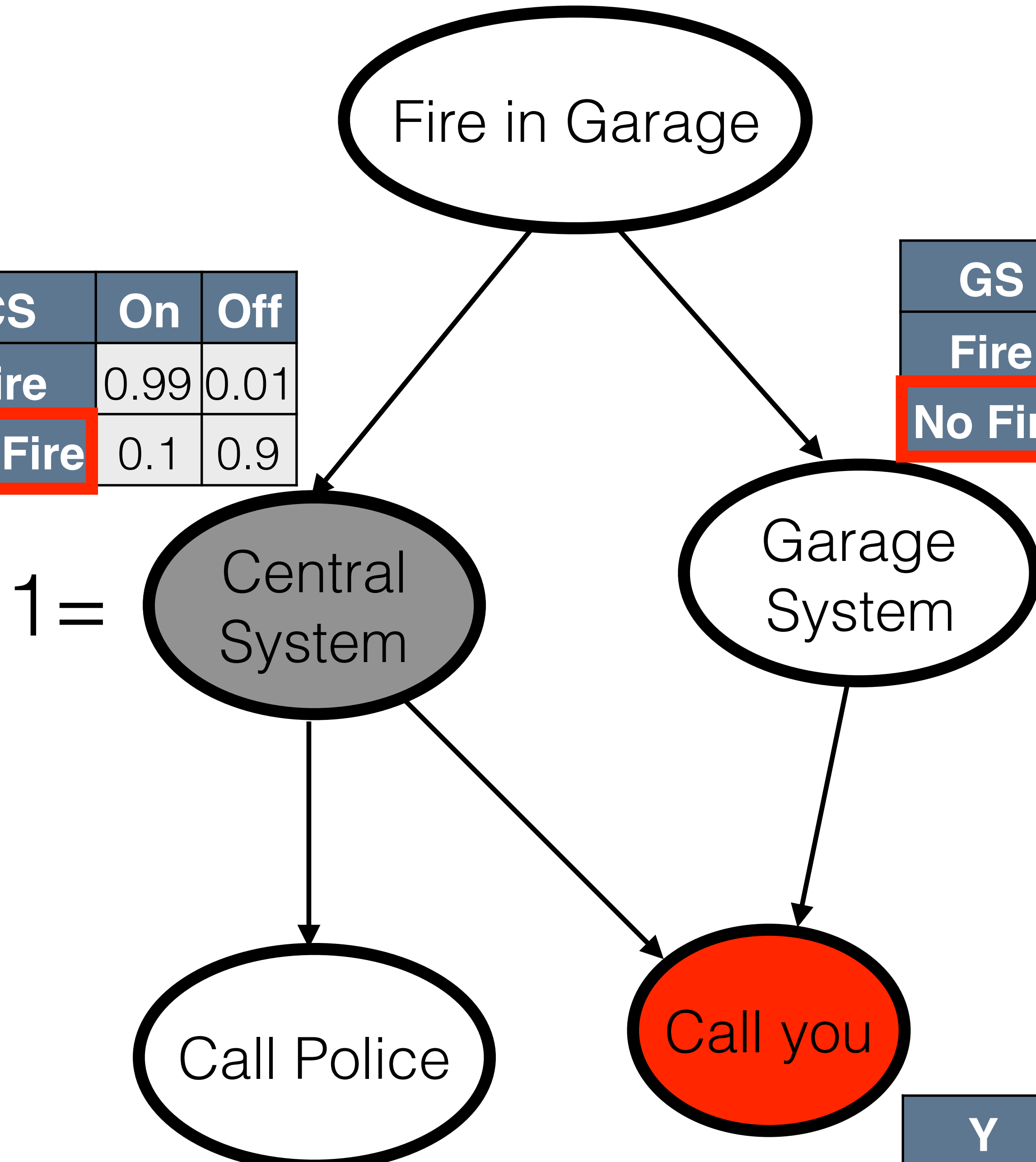
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

	F	CS	GS	P	Y
1	0	0	1	0	1
2	0	1	0	1	0
3	1	1	1	1	1
4	0	0	0	0	0
5	0	0	1	0	0
6	0	0	1	0	1
7	0	0	0	0	0
8	0	0	1	0	1
9	0	0	0	0	0
10	0	0	1	0	1
11	0	0	1	0	0
12	0	0	1	0	1
13	0	0	1	0	1
14	0	0	0	0	0
15	0	0	0	0	0
16	0	0	1	0	0
17	0	0	0	0	0
18	0	0	0	0	0
19	0	0	0	0	0
20	0	0	0	0	0
21	0	0	0	0	0
22	0	0	0	0	0
23	0	0	0	0	0
24	0	0	0	0	0
25	0	0	1	0	1

Rejection Sampling

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

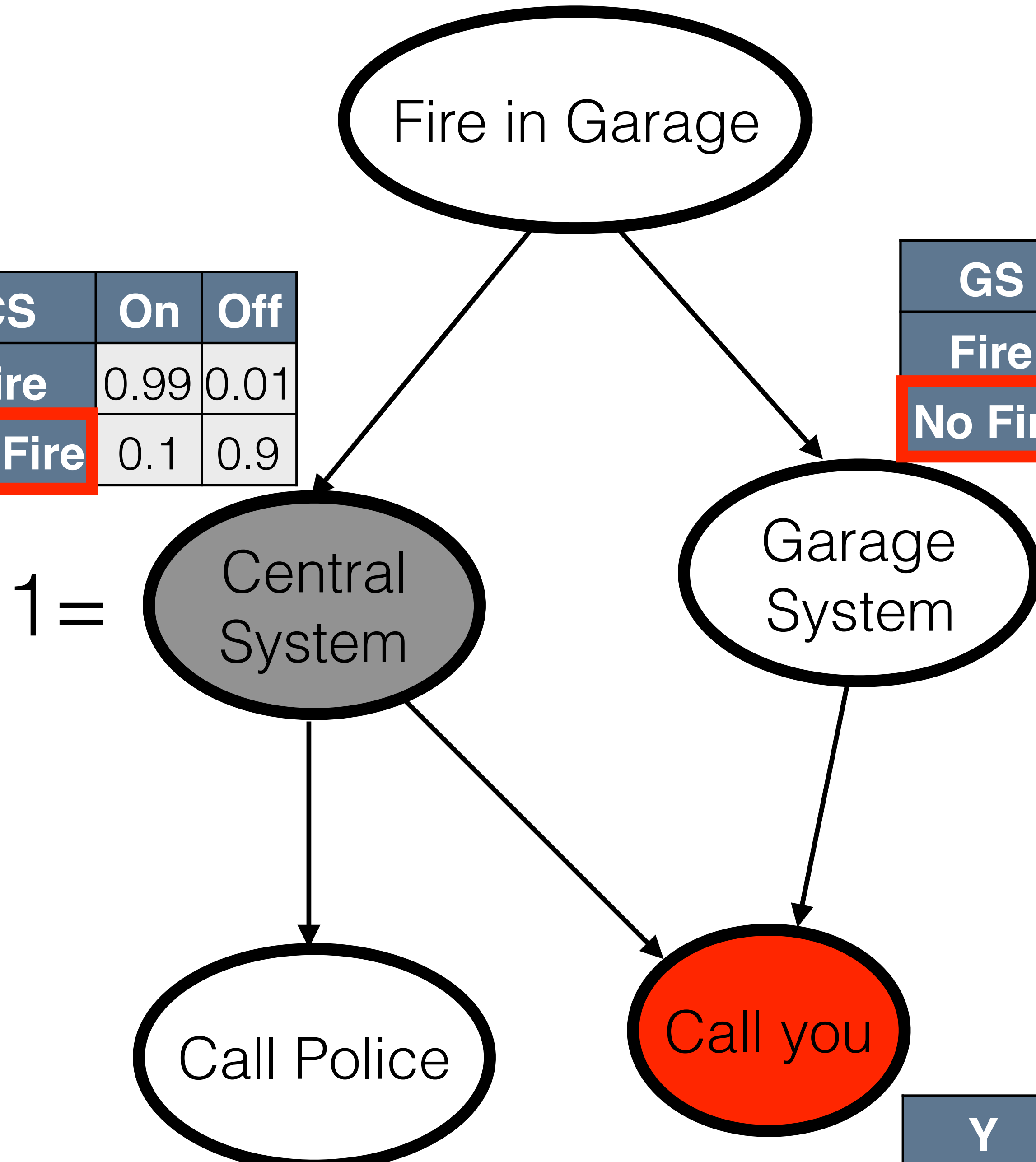
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

	F	CS	GS	P	Y
1	0	0	1	0	1
2	0	1	0	1	0
3	1	1	1	1	1
4	0	0	0	0	0
5	0	0	1	0	0
6	0	0	1	0	1
7	0	0	0	0	0
8	0	0	1	0	1
9	0	0	0	0	0
10	0	0	1	0	1
11	0	0	1	0	0
12	0	0	1	0	1
13	0	0	1	0	1
14	0	0	0	0	0
15	0	0	0	0	0
16	0	0	1	0	0
17	0	0	0	0	0
18	0	0	0	0	0
19	0	0	0	0	0
20	0	0	0	0	0
21	0	0	0	0	0
22	0	0	0	0	0
23	0	0	0	0	0
24	0	0	0	0	0
25	0	0	1	0	1

Rejection Sampling

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

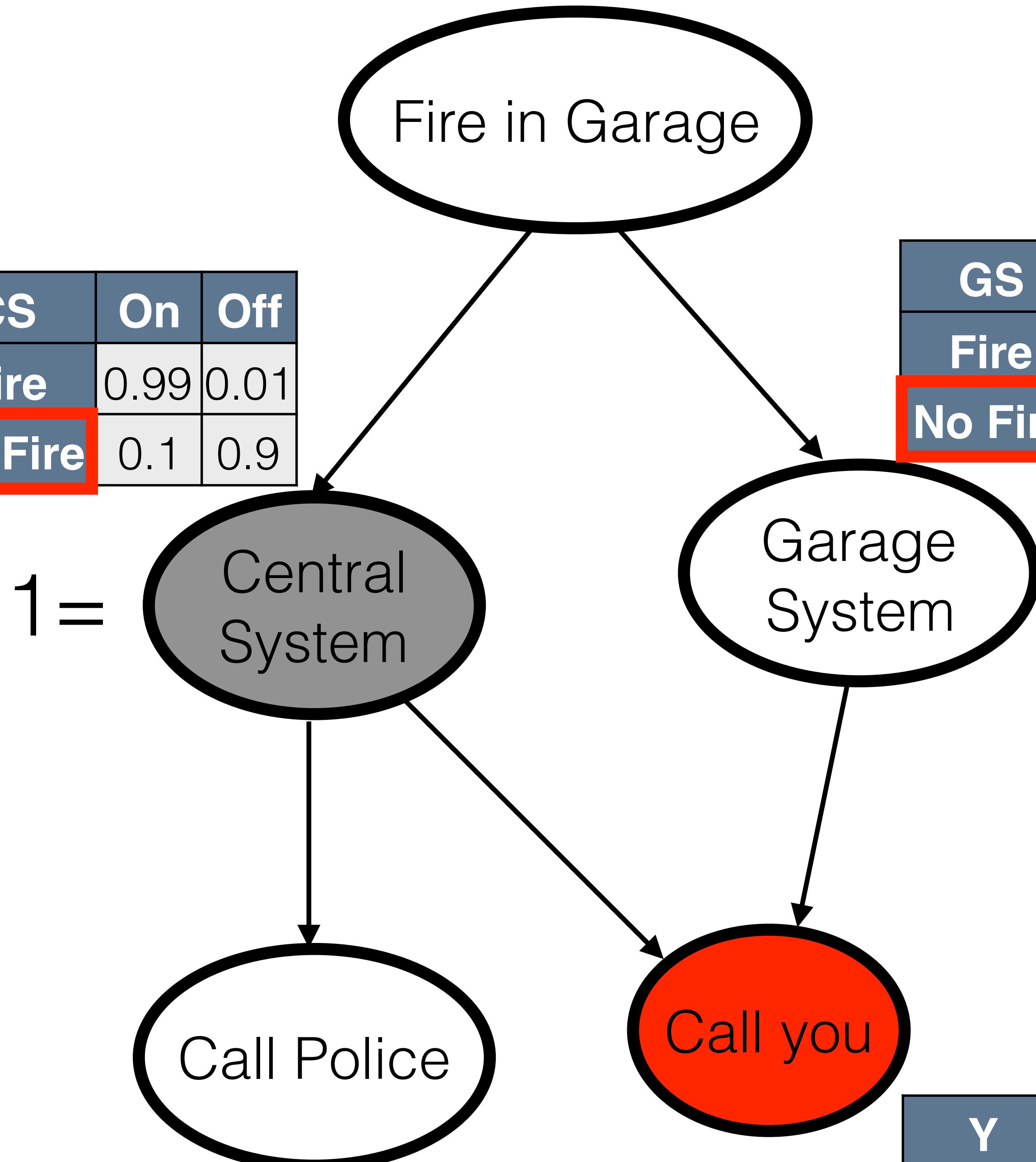
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

	F	CS	GS	P	Y
1	0	0	1	0	1
2	0	1	0	1	0
3	1	1	1	1	1
4	0	0	0	0	0
5	0	0	1	0	0
6	0	0	1	0	1
7	0	0	0	0	0
8	0	0	1	0	1
9	0	0	0	0	0
10	0	0	1	0	1
11	0	0	1	0	0
12	0	0	1	0	1
13	0	0	1	0	1
14	0	0	0	0	0
15	0	0	0	0	0
16	0	0	1	0	0
17	0	0	0	0	0
18	0	0	0	0	0
19	0	0	0	0	0
20	0	0	0	0	0
21	0	0	0	0	0
22	0	0	0	0	0
23	0	0	0	0	0
24	0	0	0	0	0
25	0	0	1	0	1

Rejection Sampling

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

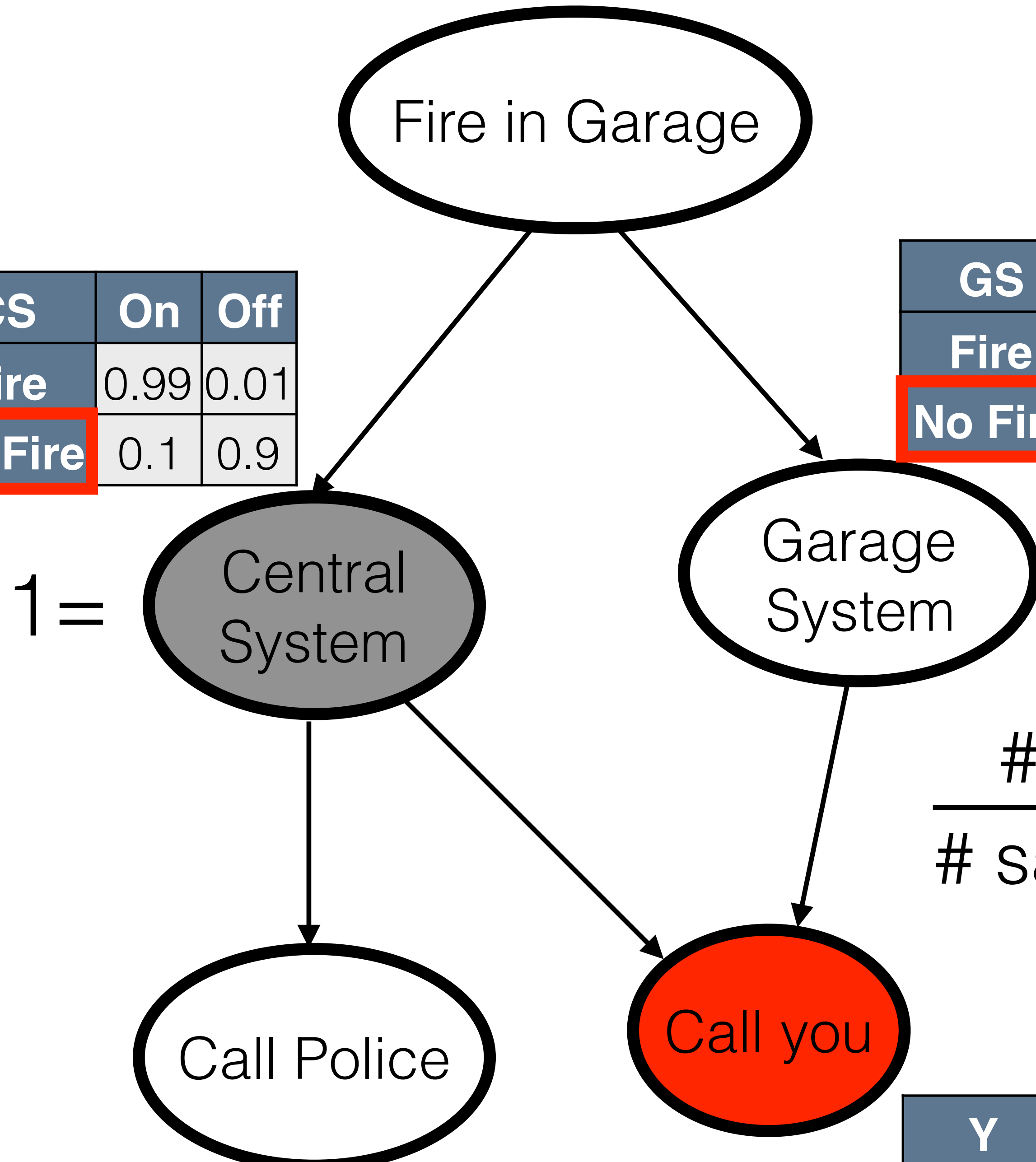
	F	CS	GS	P	Y
1	0	0	1	0	1
2	0	1	0	1	0
3	1	1	1	1	1
4	0	0	0	0	0
5	0	0	1	0	0
6	0	0	1	0	1
7	0	0	0	0	0
8	0	0	1	0	1
9	0	0	0	0	0
10	0	0	1	0	1
11	0	0	1	0	0
12	0	0	1	0	1
13	0	0	1	0	1
14	0	0	0	0	0
15	0	0	0	0	0
16	0	0	1	0	0
17	0	0	0	0	0
18	0	0	0	0	0
19	0	0	0	0	0
20	0	0	0	0	0
21	0	0	0	0	0
22	0	0	0	0	0
23	0	0	0	0	0
24	0	0	0	0	0
25	0	0	1	0	1

Rejection Sampling

Fire	No Fire
0.05	0.95

CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7



$$\frac{\#\ \text{calls}}{\#\ \text{samples}} = \frac{1}{2}$$

P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

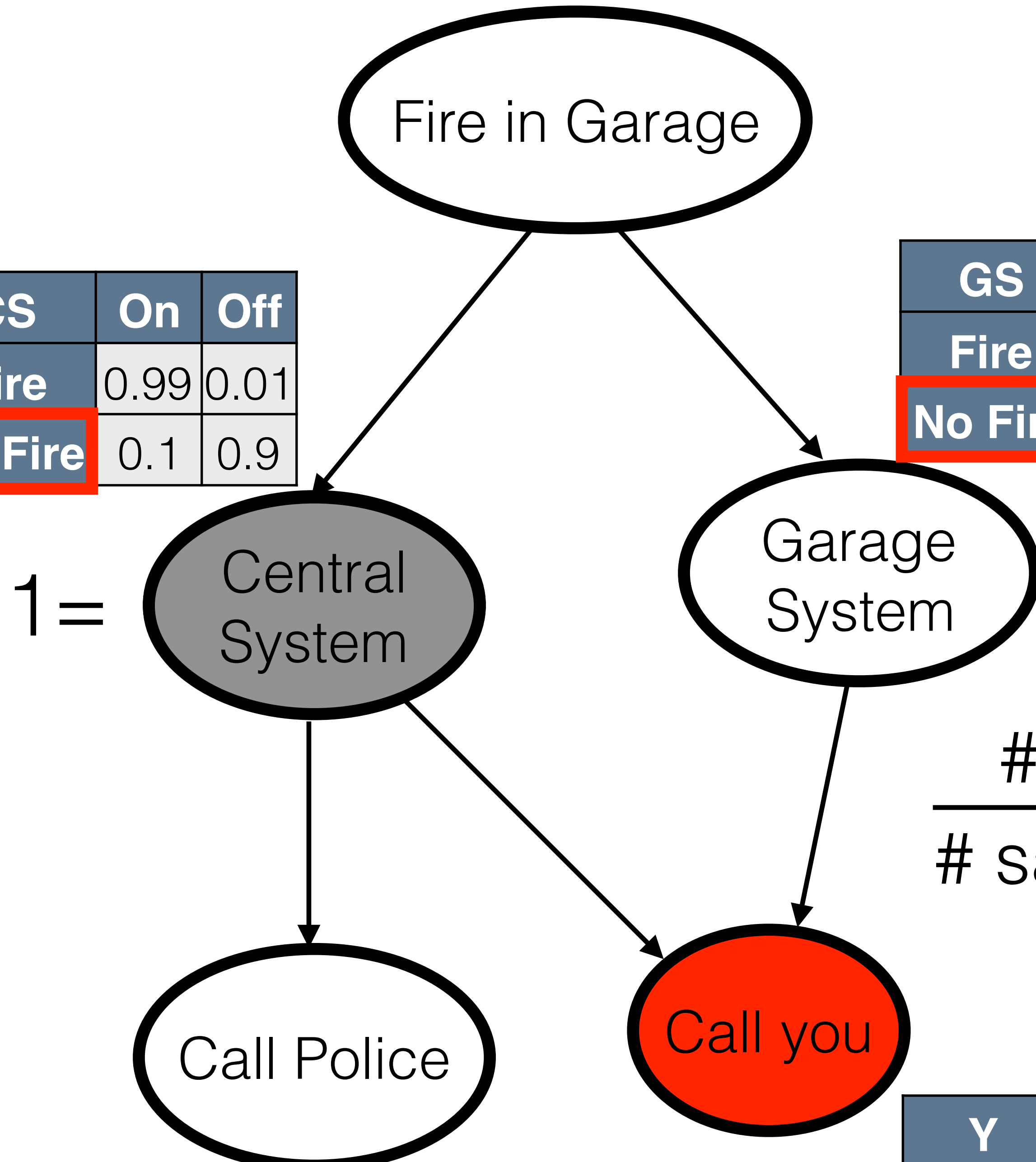
	F	CS	GS	P	Y
1	0	0	1	0	1
2	0	1	0	1	0
3	1	1	1	1	1
4	0	0	0	0	0
5	0	0	1	0	0
6	0	0	1	0	1
7	0	0	0	0	0
8	0	0	1	0	1
9	0	0	0	0	0
10	0	0	1	0	1
11	0	0	1	0	0
12	0	0	1	0	1
13	0	0	1	0	1
14	0	0	0	0	0
15	0	0	0	0	0
16	0	0	1	0	0
17	0	0	0	0	0
18	0	0	0	0	0
19	0	0	0	0	0
20	0	0	0	0	0
21	0	0	0	0	0
22	0	0	0	0	0
23	0	0	0	0	0
24	0	0	0	0	0
25	0	0	1	0	1

Rejection Sampling

Fire	No Fire
0.05	0.95

CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7



???!!!

$$\frac{\#\ \text{calls}}{\#\ \text{samples}} = \frac{1}{2}$$

P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

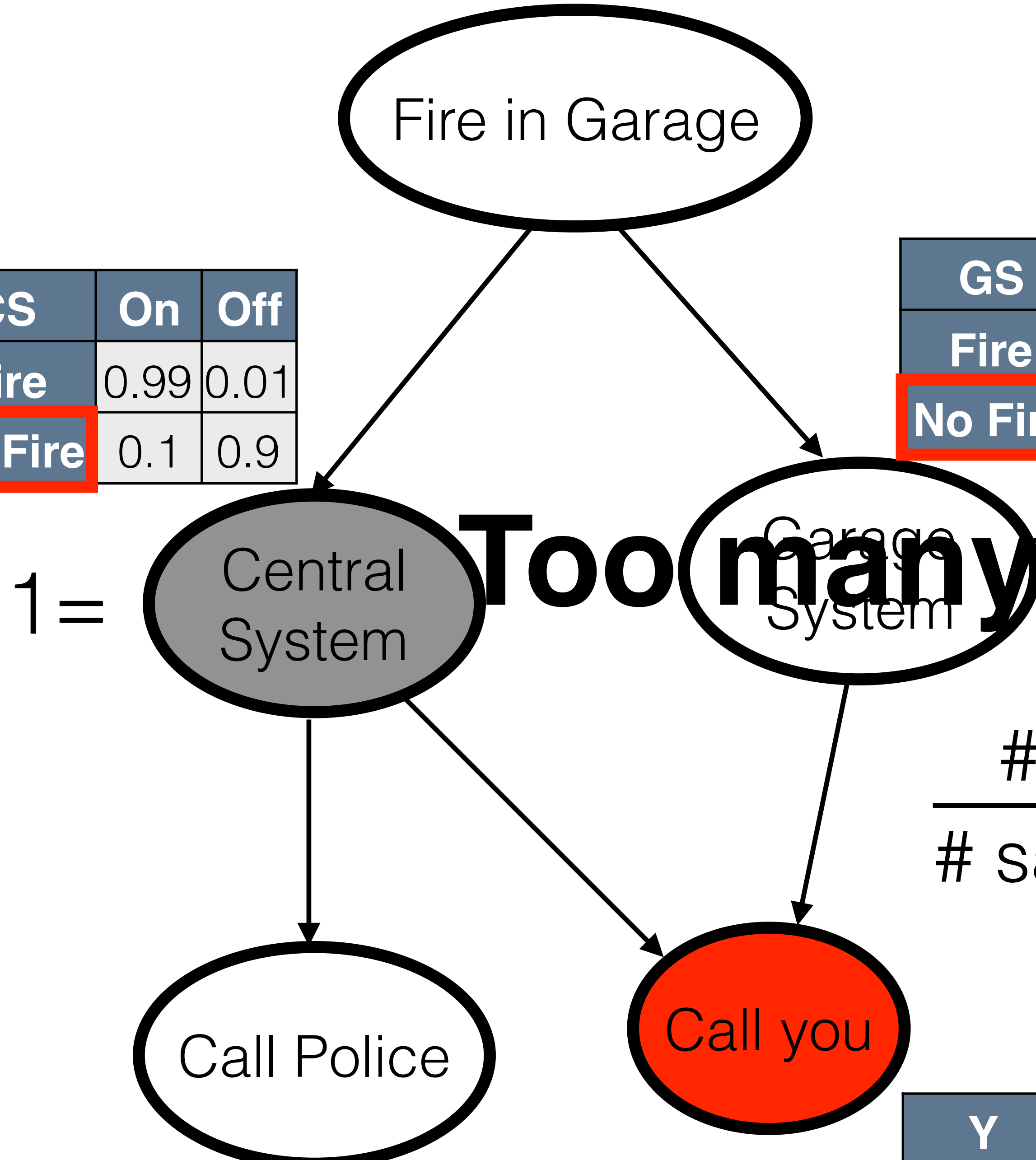
	F	CS	GS	P	Y
1	0	0	1	0	1
2	0	1	0	1	0
3	1	1	1	1	1
4	0	0	0	0	0
5	0	0	1	0	0
6	0	0	1	0	1
7	0	0	0	0	0
8	0	0	1	0	1
9	0	0	0	0	0
10	0	0	1	0	1
11	0	0	1	0	0
12	0	0	1	0	1
13	0	0	1	0	1
14	0	0	0	0	0
15	0	0	0	0	0
16	0	0	1	0	0
17	0	0	0	0	0
18	0	0	0	0	0
19	0	0	0	0	0
20	0	0	0	0	0
21	0	0	0	0	0
22	0	0	0	0	0
23	0	0	0	0	0
24	0	0	0	0	0
25	0	0	1	0	1

Rejection Sampling

Fire	No Fire
0.05	0.95

CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7



Too many rejections :(

???!!!

$$\frac{\#\ \text{calls}}{\#\ \text{samples}} = \frac{1}{2}$$

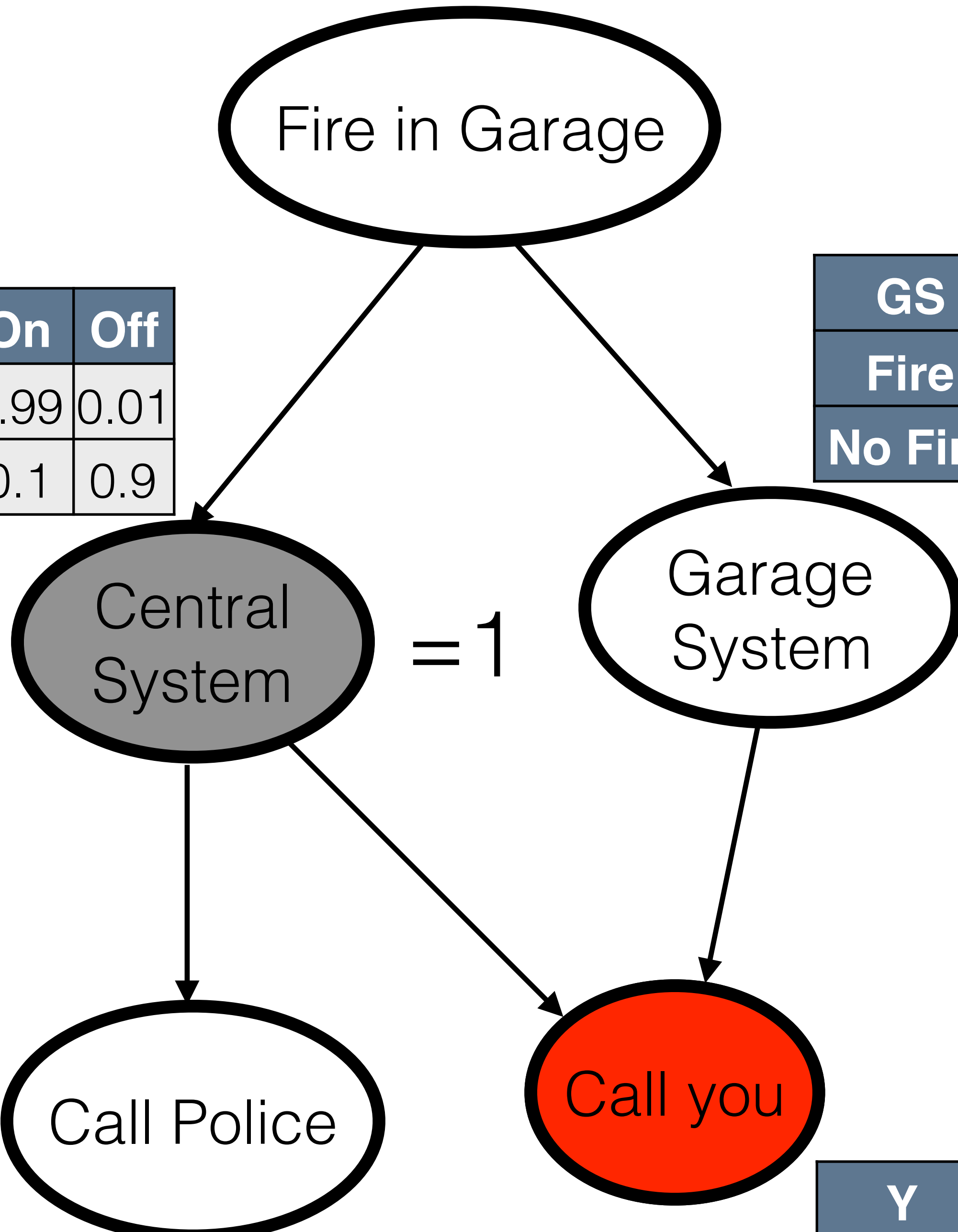
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

	F	CS	GS	P	Y
1	0	0	1	0	1
2	0	1	0	1	0
3	1	1	1	1	1
4	0	0	0	0	0
5	0	0	1	0	0
6	0	0	1	0	1
7	0	0	0	0	0
8	0	0	1	0	1
9	0	0	0	0	0
10	0	0	1	0	1
11	[illegible]	0	1	0	0
12	0	0	1	0	1
13	0	0	1	0	1
14	0	0	0	0	0
15	0	0	0	0	0
16	0	0	1	0	0
17	0	0	0	0	0
18	0	0	0	0	0
19	0	0	0	0	0
20	0	0	0	0	0
21	0	0	0	0	0
22	0	0	0	0	0
23	0	0	0	0	0
24	0	0	0	0	0
25	0	0	1	0	1

IMPORTANCE SAMPLING

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

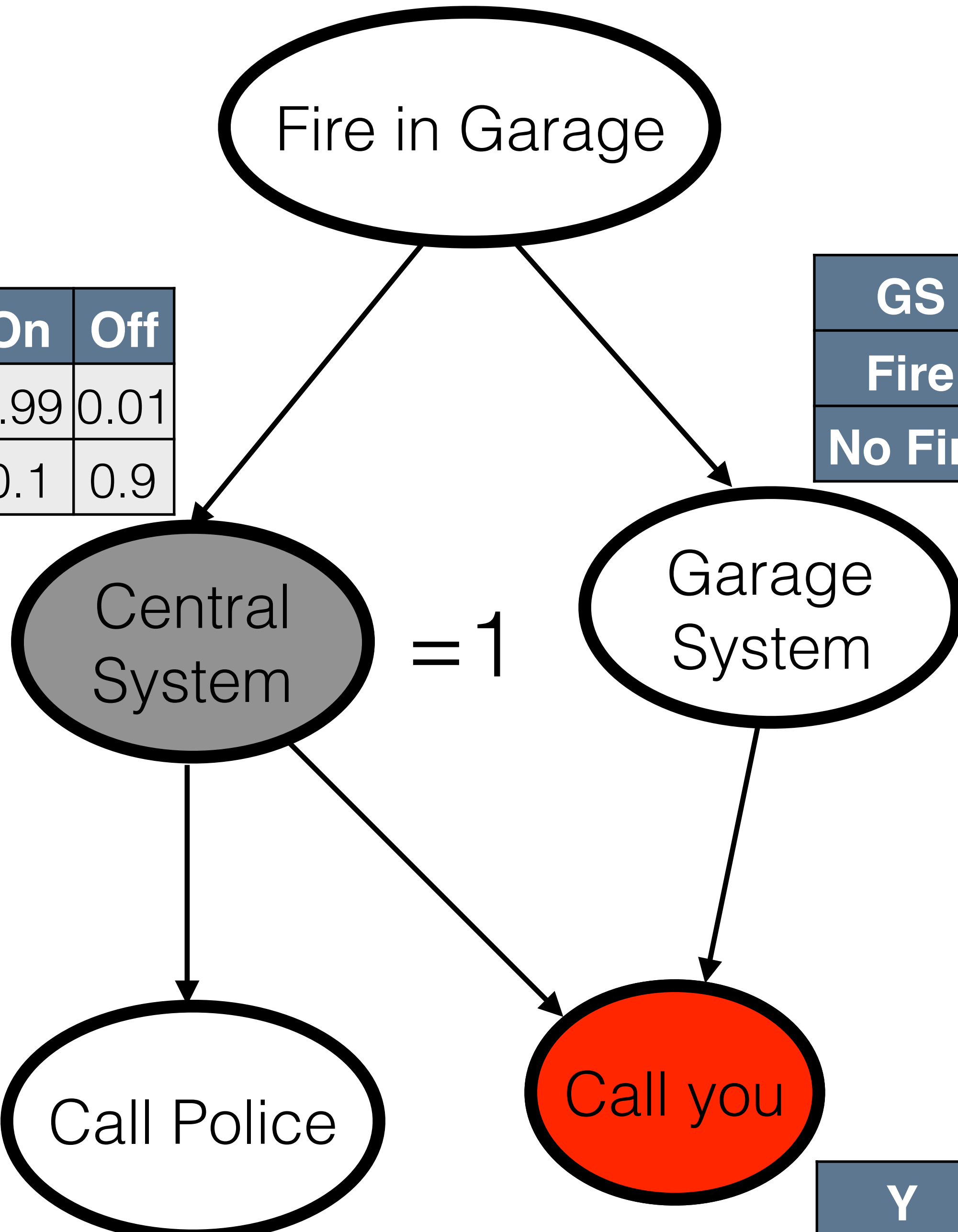
GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

Importance Sampling

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

P	Call	no call
CS on	1	0
CS off	0.01	0.99

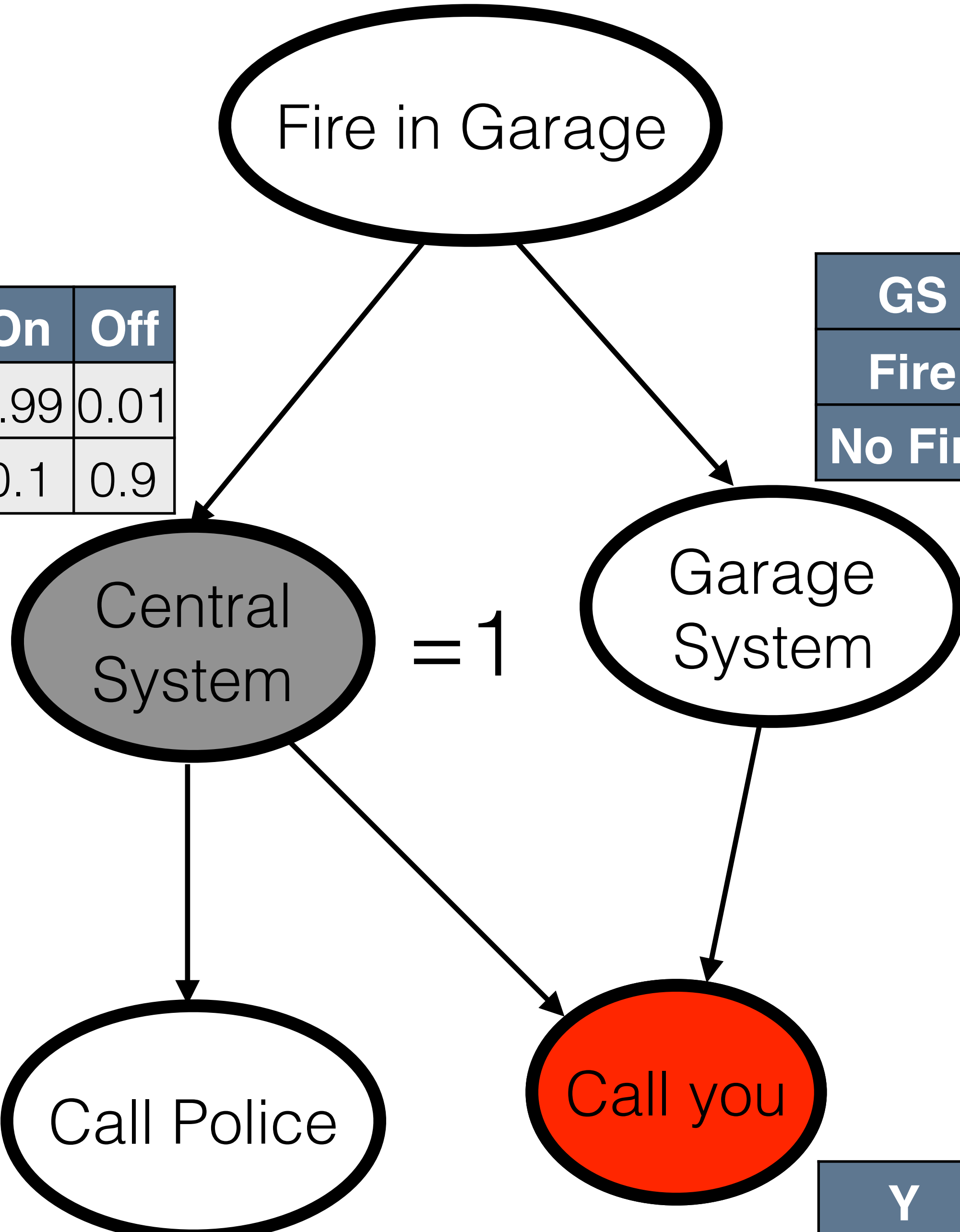
Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

	F	CS	GS	P	Y
1					

Importance Sampling

Fire	No Fire
0.05	0.95

	F	CS	GS	P	Y
1	0				



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

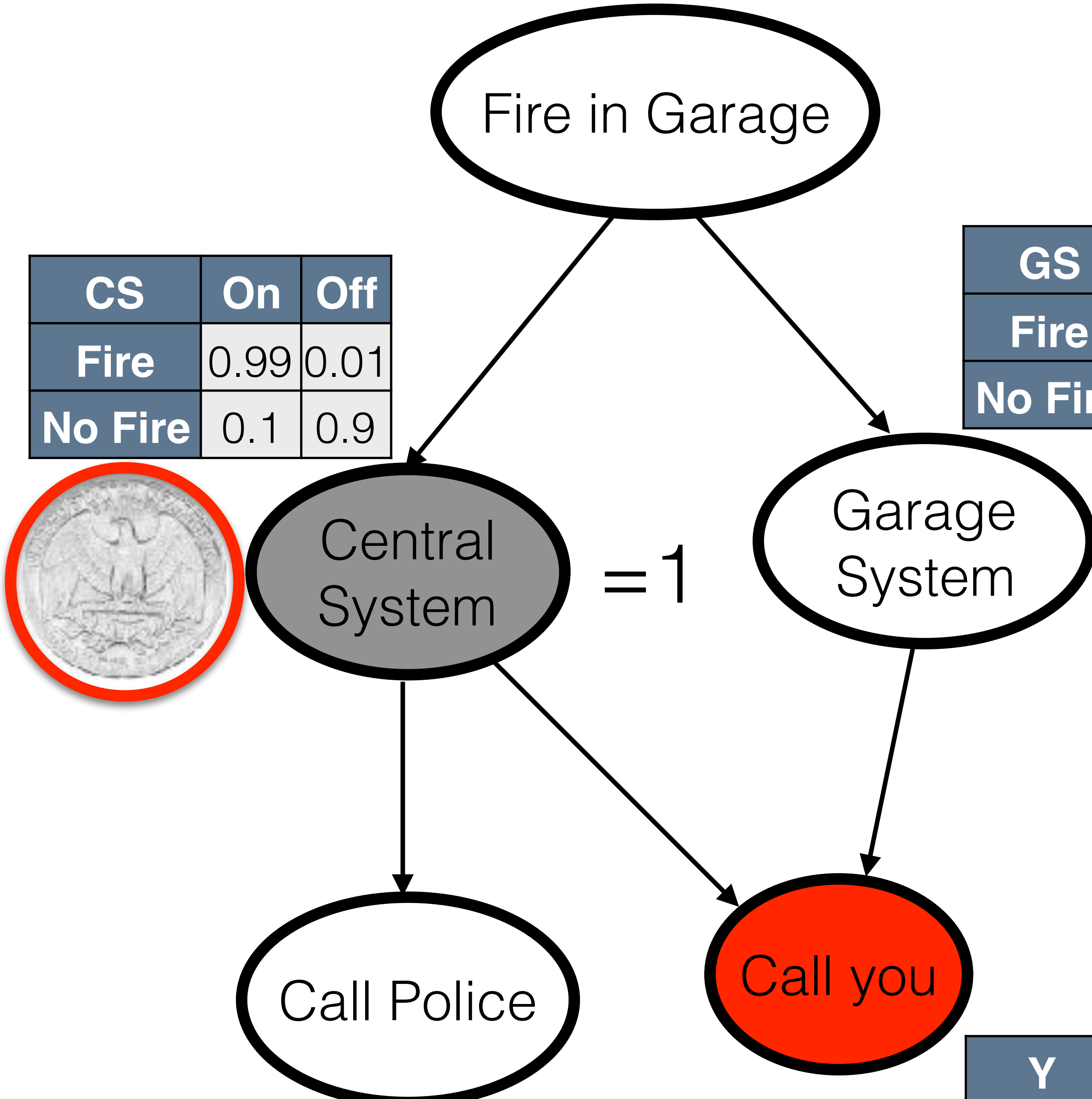
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

Importance Sampling

	F	CS	GS	P	Y
1	0	1			

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

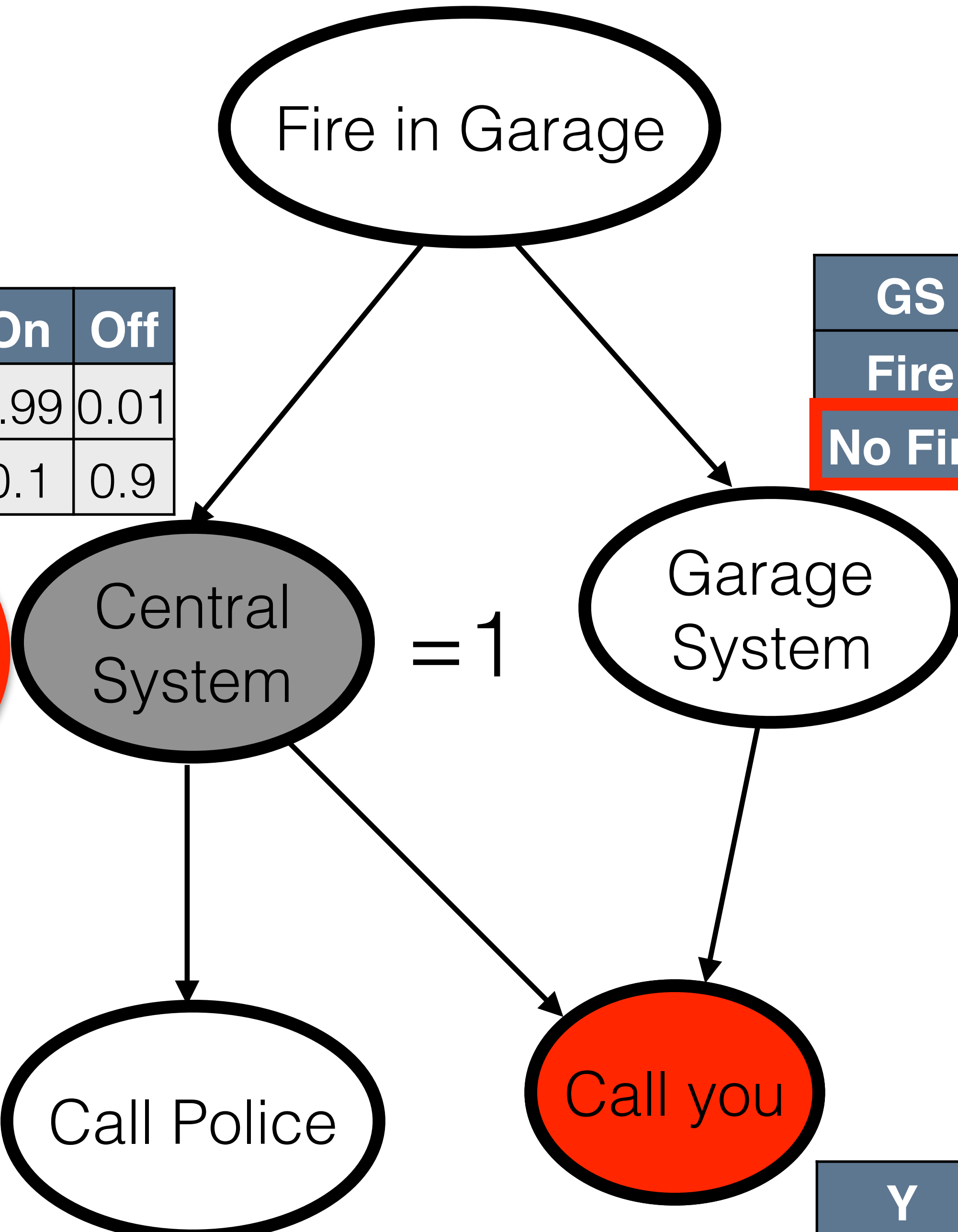
Importance Sampling

Fire	No Fire
0.05	0.95

	F	CS	GS	P	Y
1	0	1	0		

CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7



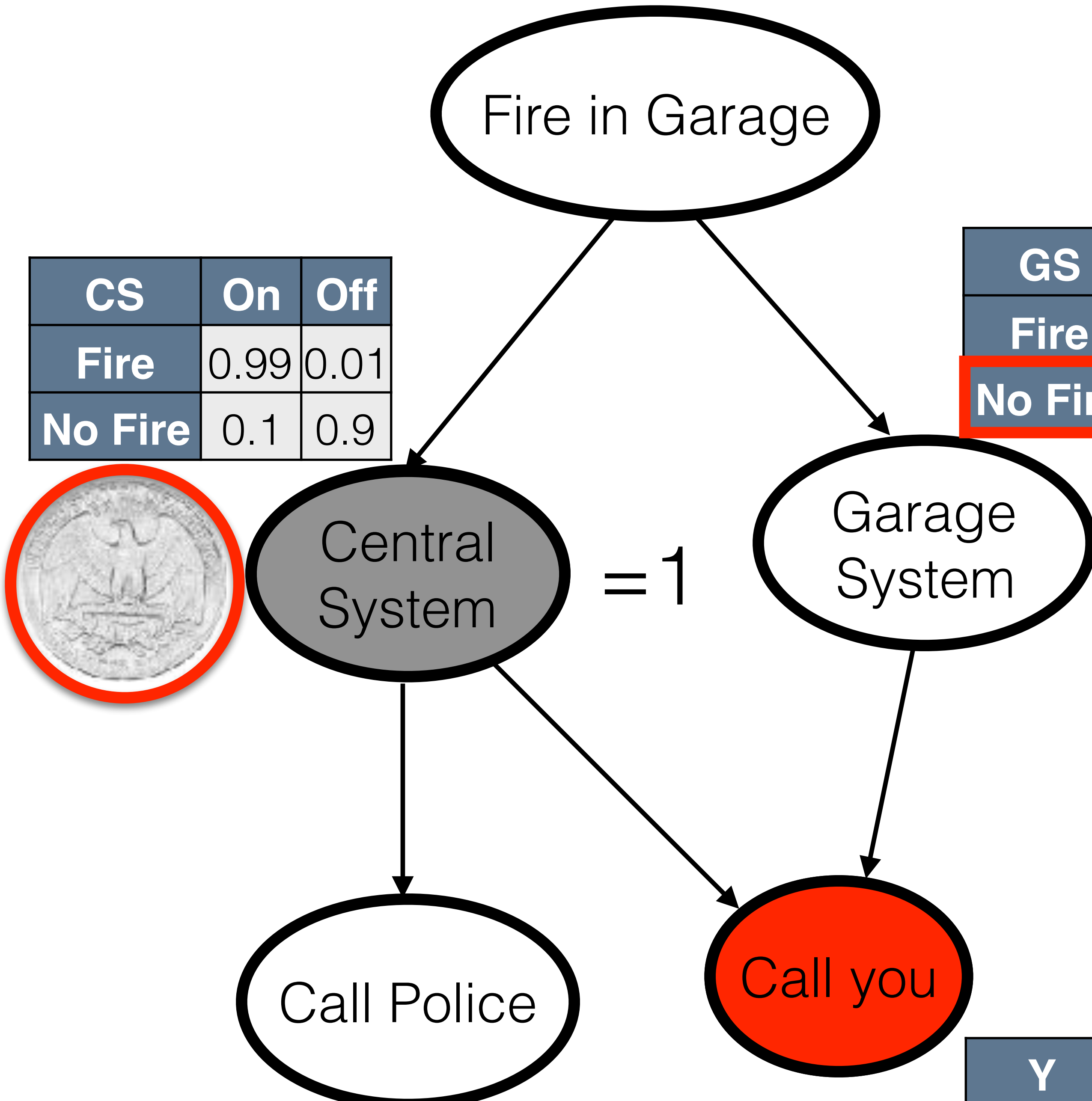
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

Importance Sampling

Fire	No Fire
0.05	0.95

	F	CS	GS	P	Y
1	0	1	0	1	



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

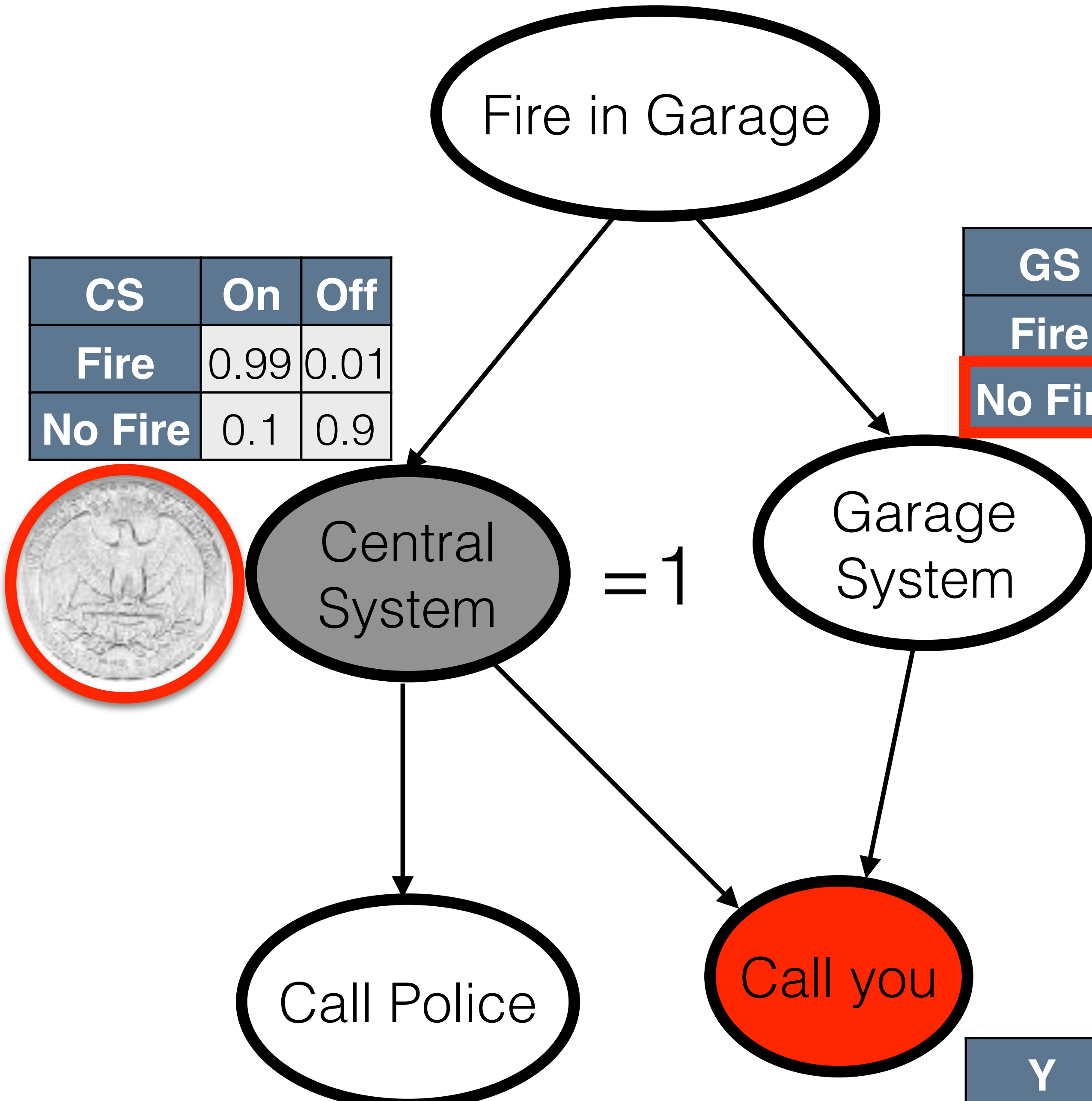
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

Importance Sampling

	F	CS	GS	P	Y
1	0	1	0	1	0

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

P	Call	no call
CS on	1	0
CS off	0.01	0.99

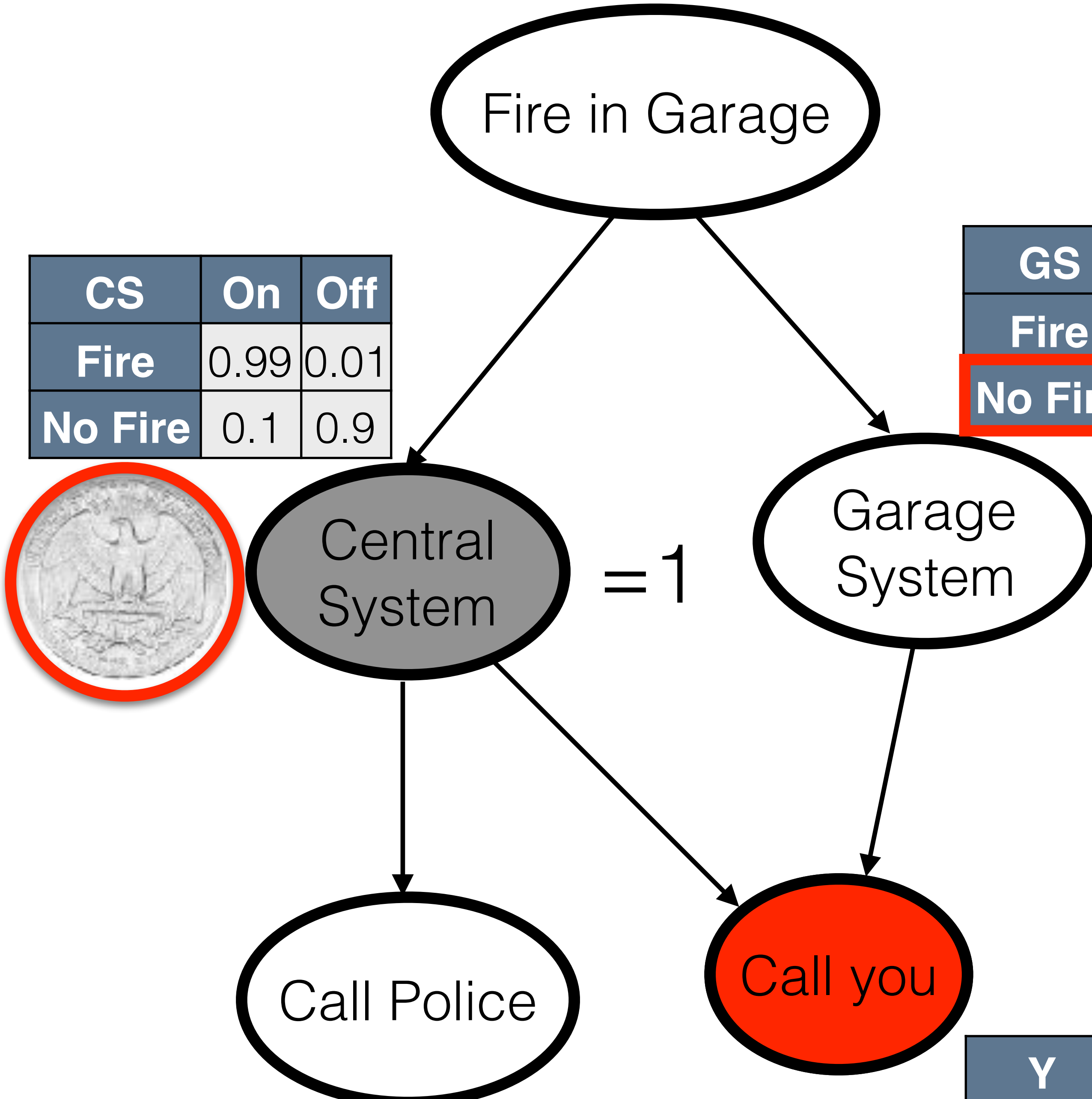
Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

Importance Sampling

Fire	No Fire
0.05	0.95

CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7



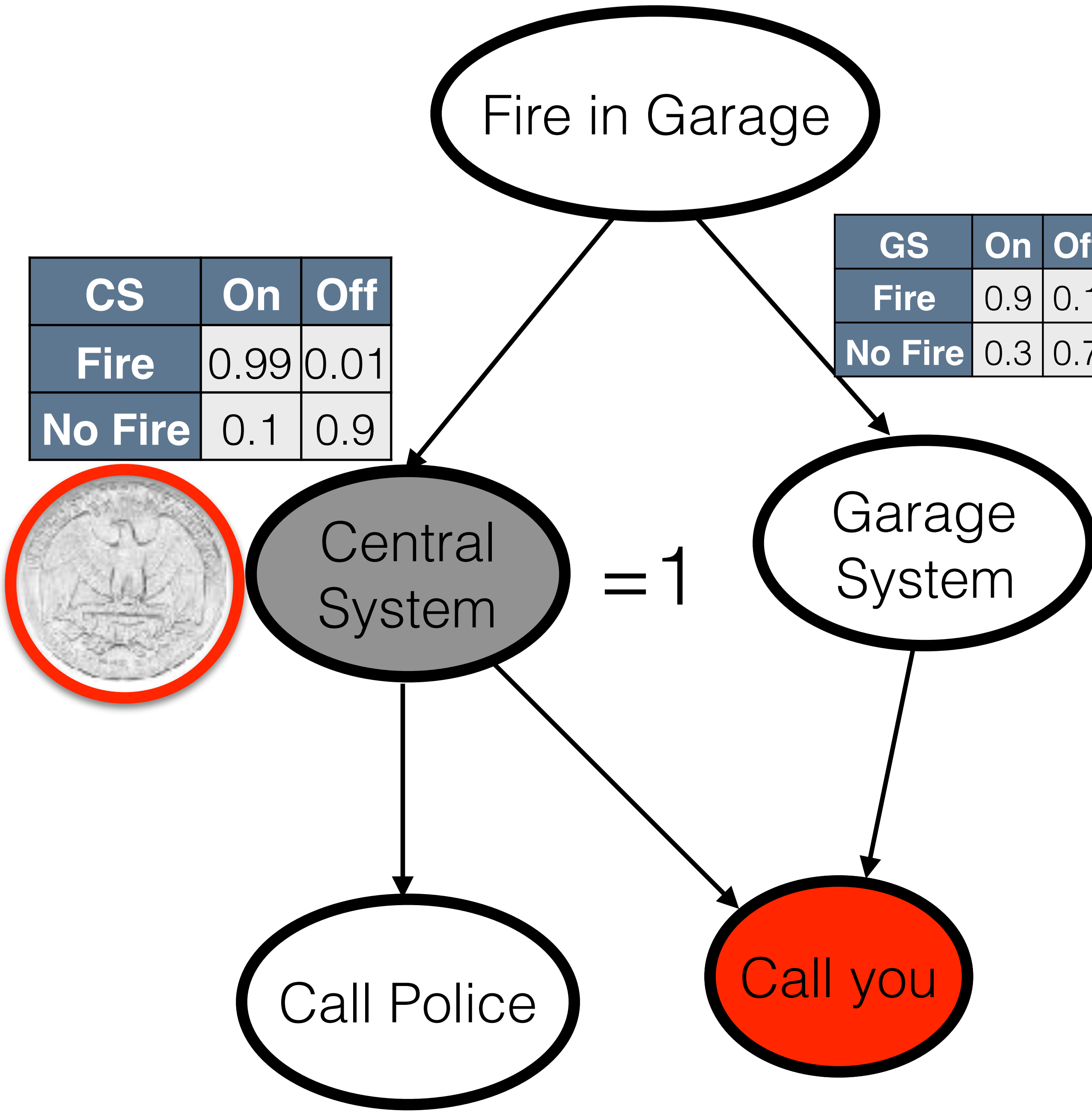
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

	F	CS	GS	P	Y
1	0	1	0	1	0
2	0	1	1	1	1
3	1	1	1	1	1
4	0	1	0	1	1
5	0	1	0	1	1
6	0	1	0	1	1
7	0	1	1	1	1
8	0	1	0	1	0
9	0	1	0	1	1
10	0	1	1	1	1
11	0	1	0	1	1
12	0	1	1	1	0
13	0	1	0	1	1
14	0	1	1	1	1
15	0	1	0	1	1
16	0	1	0	1	1
17	0	1	0	1	1
18	0	1	0	1	1
19	0	1	1	1	1
20	0	1	0	1	1
21	0	1	0	1	1
22	0	1	0	1	1
23	0	1	1	1	1
24	0	1	0	1	1
25	0	1	1	1	1

Importance Sampling

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

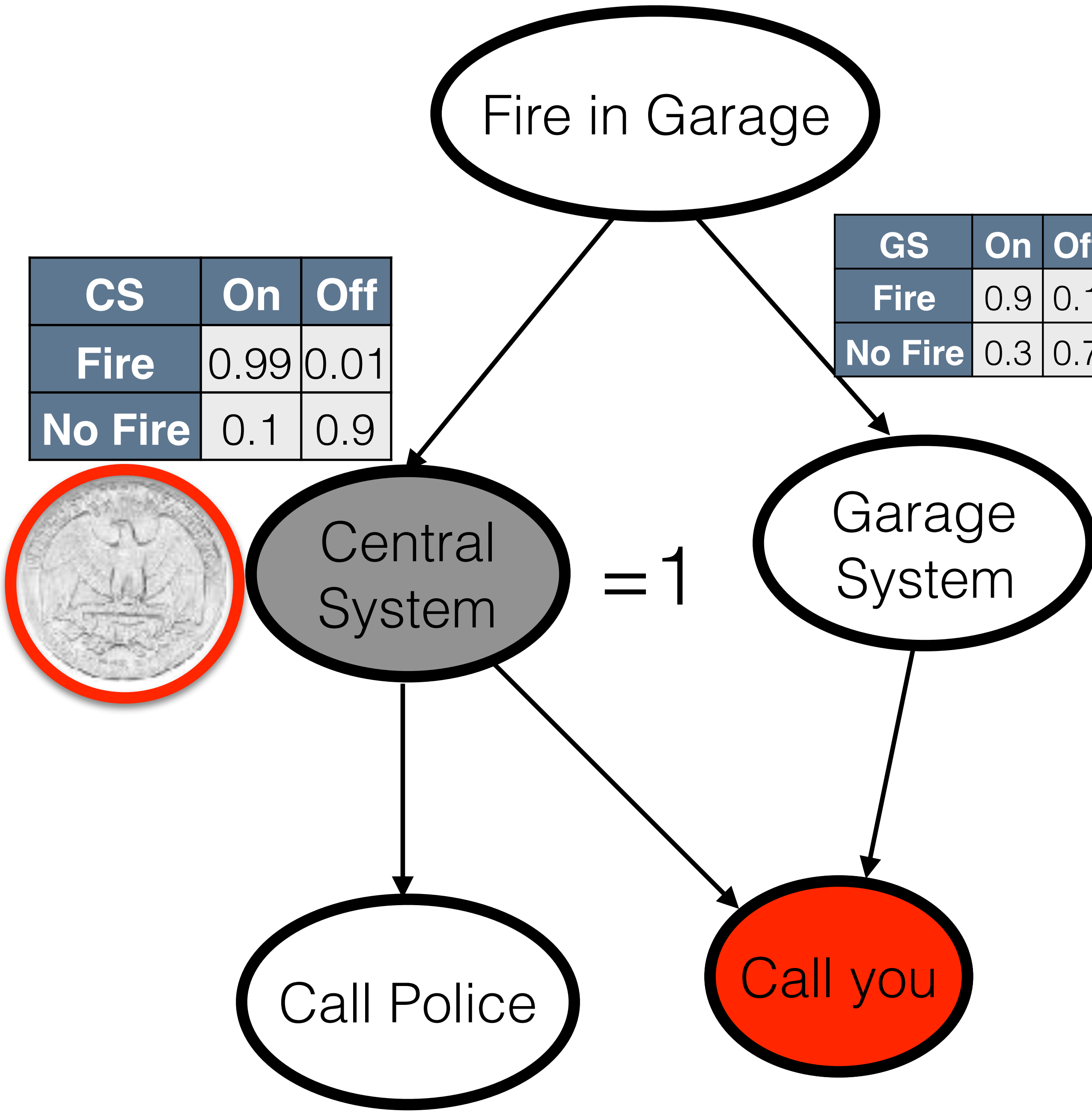
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

	F	CS	GS	P	Y
1	0	1	0	1	0
2	0	1	1	1	1
3	1	1	1	1	1
4	0	1	0	1	1
5	0	1	0	1	1
6	0	1	0	1	1
7	0	1	1	1	1
8	0	1	0	1	0
9	0	1	0	1	1
10	0	1	1	1	1
11	0	1	0	1	1
12	0	1	1	1	0
13	0	1	0	1	1
14	0	1	1	1	1
15	0	1	0	1	1
16	0	1	0	1	1
17	0	1	0	1	1
18	0	1	0	1	1
19	0	1	1	1	1
20	0	1	0	1	1
21	0	1	0	1	1
22	0	1	0	1	1
23	0	1	1	1	1
24	0	1	0	1	1
25	0	1	1	1	1

Importance Sampling

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

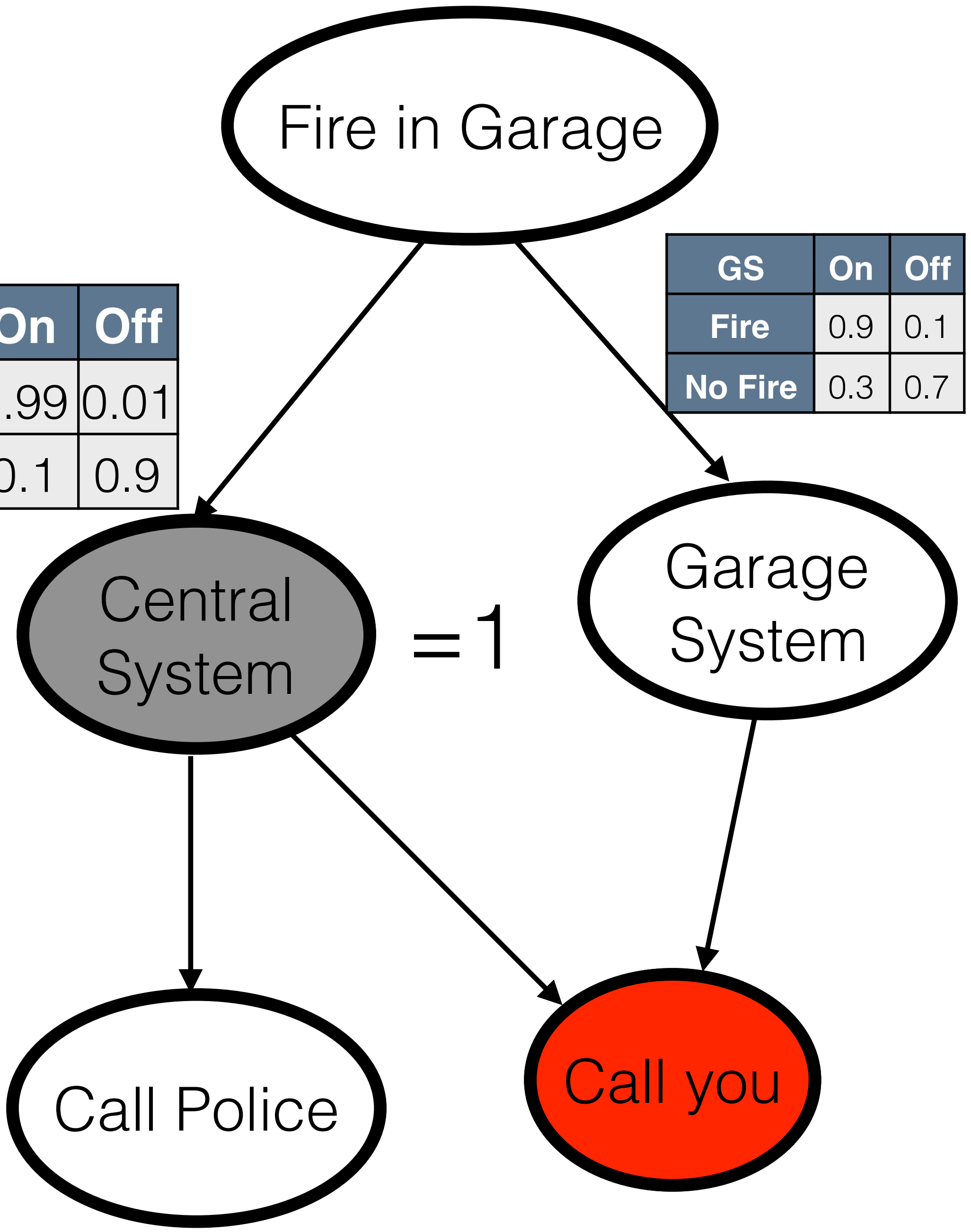
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no call
CS on	GS on	0.95	0.05
CS off	GS on	0.7	0.3
CS on	GS off	0.9	0.1
CS off	GS off	0.01	0.99

	F	CS	GS	P	Y
1	0	1	0	1	0
2	0	1	1	1	1
3	1	1	1	1	1
4	0	1	0	1	1
5	0	1	0	1	1
6	0	1	0	1	1
7	0	1	1	1	1
8	0	1	0	1	0
9	0	1	0	1	1
10	0	1	1	1	1
11	0	1	0	1	1
12	0	1	1	1	0
13	0	1	0	1	1
14	0	1	1	1	1
15	0	1	0	1	1
16	0	1	0	1	1
17	0	1	0	1	1
18	0	1	0	1	1
19	0	1	1	1	1
20	0	1	0	1	1
21	0	1	0	1	1
22	0	1	0	1	1
23	0	1	1	1	1
24	0	1	0	1	1
25	0	1	1	1	1

Importance Sampling

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

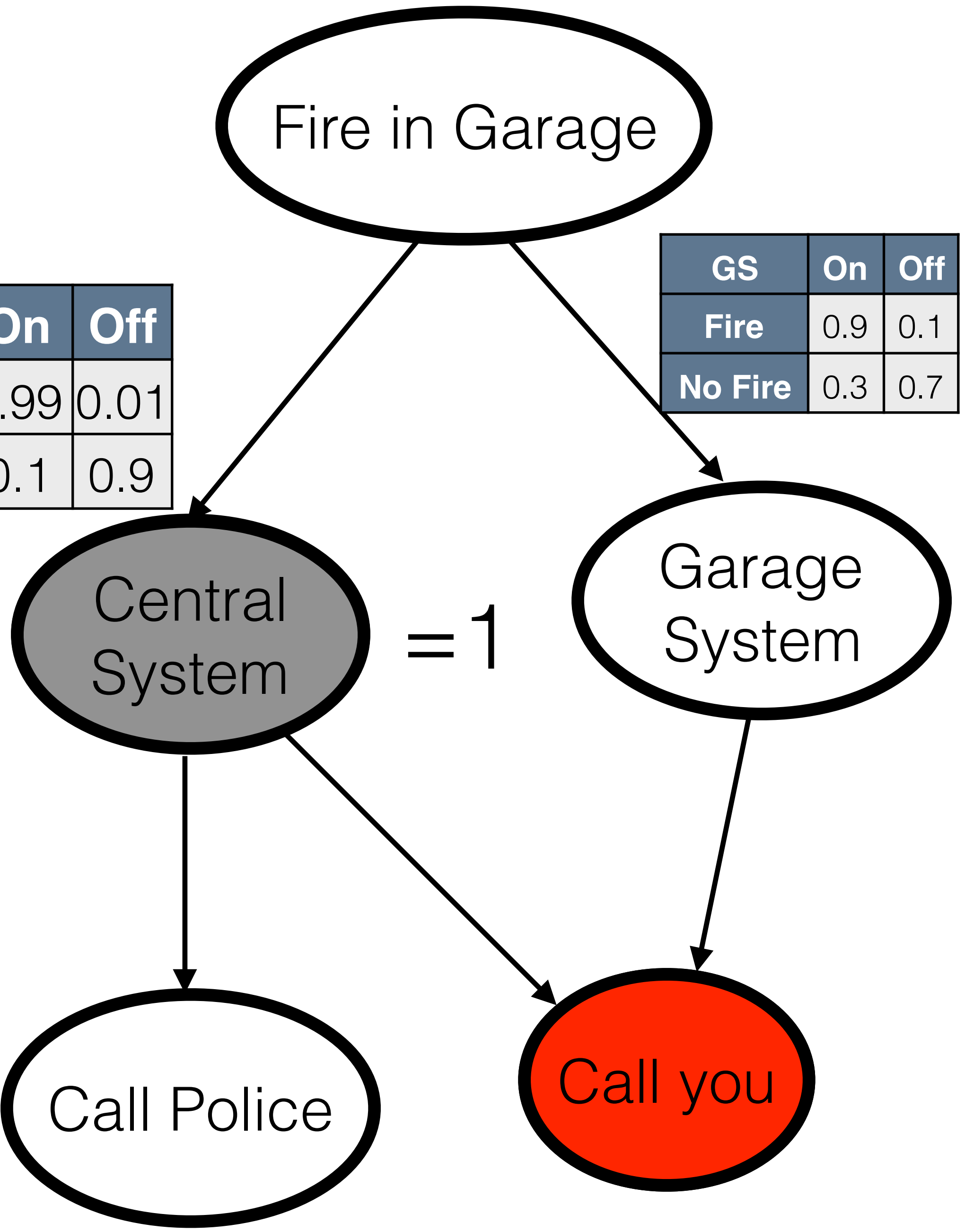
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no
CS	GS	0.95	0.05
CS	GS	0.7	0.3
CS	GS	0.9	0.1
CS	GS	0.01	0.99

	F	CS	GS	P	Y	Weight
1	0	1	0	1	0	
2	0	1	1	1	1	
3	1	1	1	1	1	
4	0	1	0	1	1	
5	0	1	0	1	1	
6	0	1	0	1	1	
7	0	1	1	1	1	
8	0	1	0	1	0	
9	0	1	0	1	1	
10	0	1	1	1	1	
11	0	1	0	1	1	
12	0	1	1	1	0	
13	0	1	0	1	1	
14	0	1	1	1	1	
15	0	1	0	1	1	
16	0	1	0	1	1	
17	0	1	0	1	1	
18	0	1	0	1	1	
19	0	1	1	1	1	
20	0	1	0	1	1	
21	0	1	0	1	1	
22	0	1	0	1	1	
23	0	1	1	1	1	
24	0	1	0	1	1	
25	0	1	1	1	1	

Importance Sampling

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

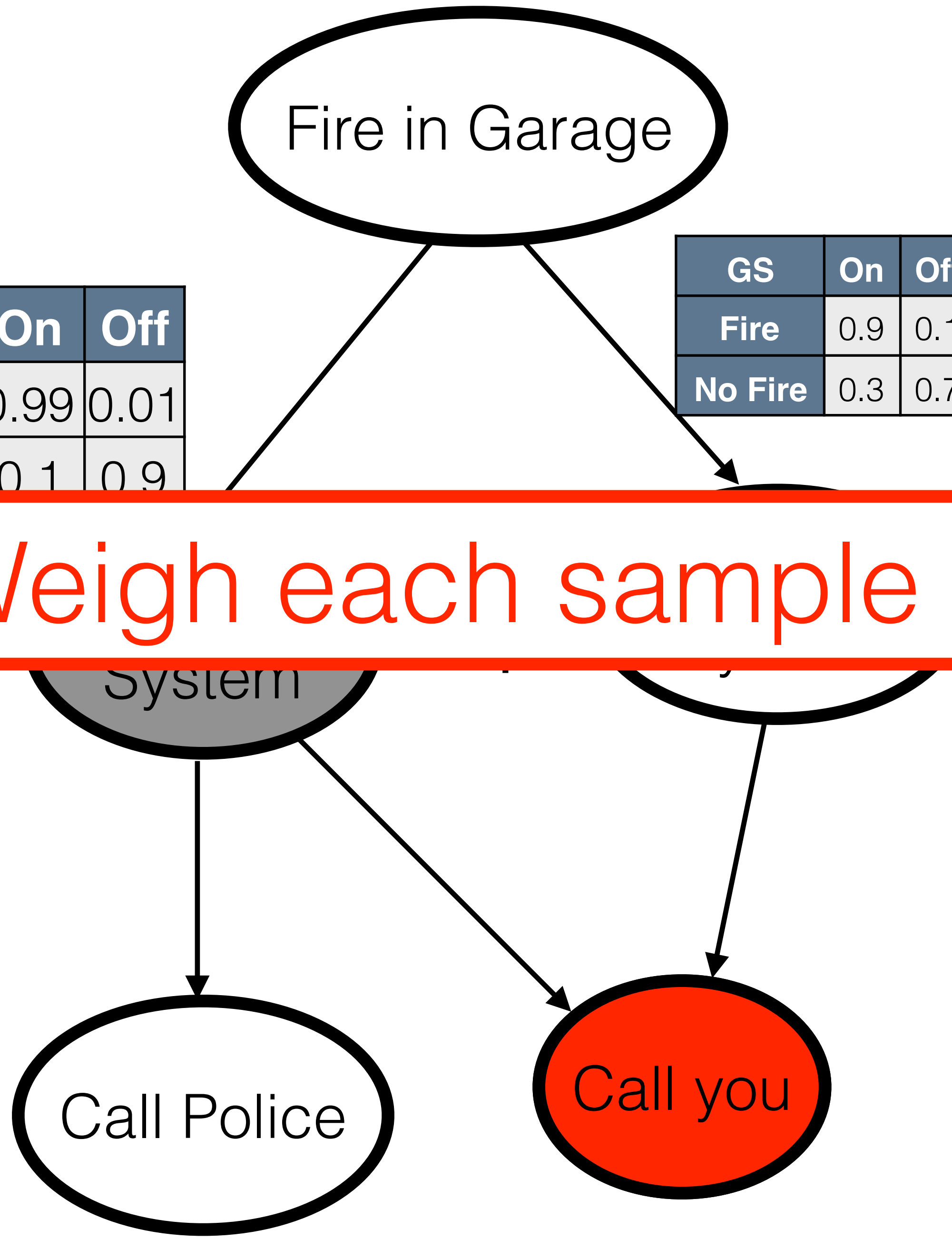
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no
CS	GS	0.95	0.05
CS	GS	0.7	0.3
CS	GS	0.9	0.1
CS	GS	0.01	0.99

	F	CS	GS	P	Y	Weight
1	0	1	0	1	0	
2	0	1	1	1	1	
3	1	1	1	1	1	
4	0	1	0	1	1	
5	0	1	0	1	1	
6	0	1	0	1	1	
7	0	1	1	1	1	
8	0	1	0	1	0	
9	0	1	0	1	1	
10	0	1	1	1	1	
11	0	1	0	1	1	
12	0	1	1	1	0	
13	0	1	0	1	1	
14	0	1	1	1	1	
15	0	1	0	1	1	
16	0	1	0	1	1	
17	0	1	0	1	1	
18	0	1	0	1	1	
19	0	1	1	1	1	
20	0	1	0	1	1	
21	0	1	0	1	1	
22	0	1	0	1	1	
23	0	1	1	1	1	
24	0	1	0	1	1	
25	0	1	1	1	1	

Importance Sampling

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

Weigh each sample differently!

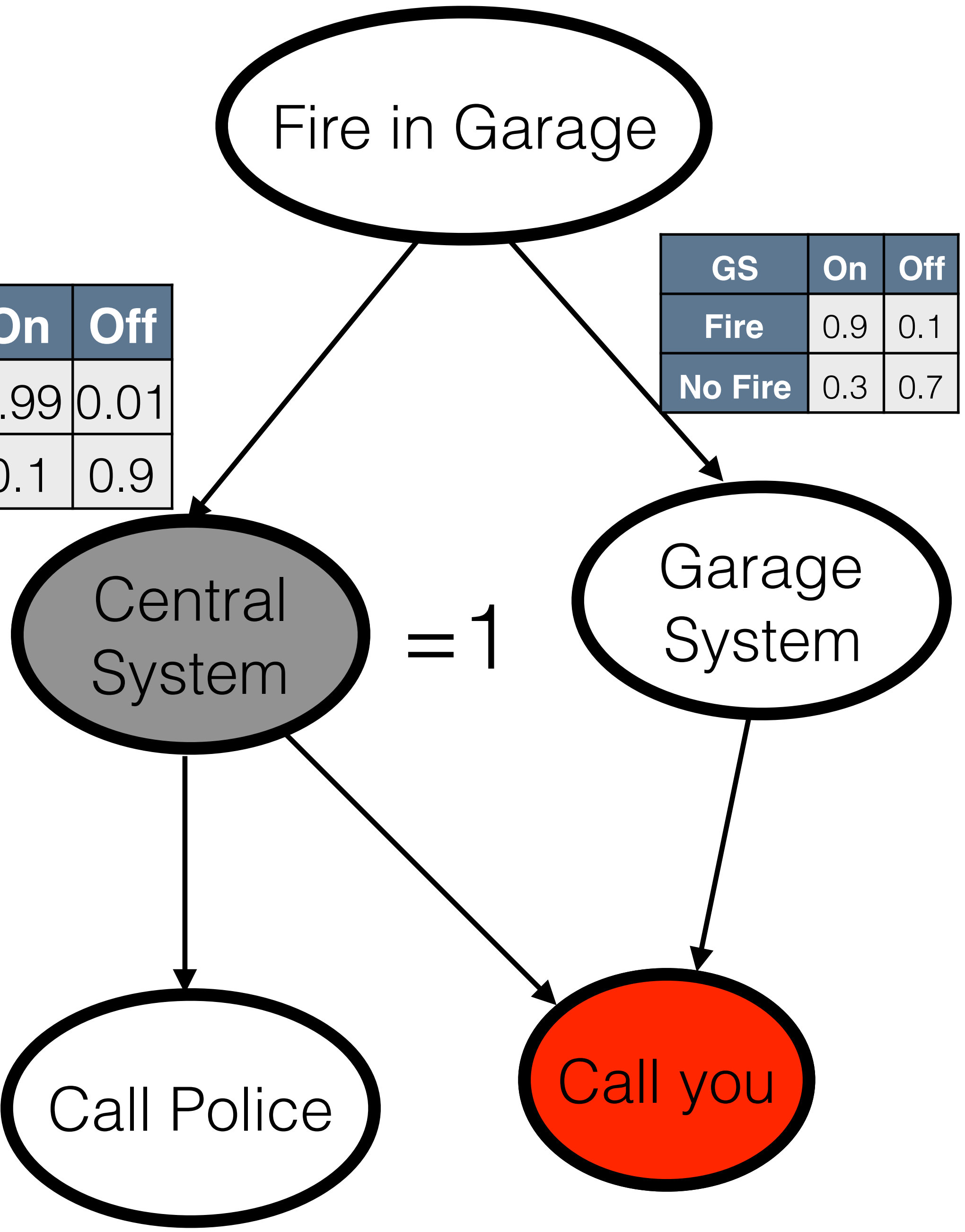
P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no
CS	GS	0.95	0.05
CS	GS	0.7	0.3
CS	GS	0.9	0.1
CS	GS	0.01	0.99

	F	CS	GS	P	Y	Weight
1	0	1	0	1	0	
2	0	1	1	1	1	
3	1	1	1	1	1	
4	0	1	0	1	1	
5	0	1	0	1	1	
6	0	1	0	1	1	
7	0	1	1	1	1	
8	0	1	0	1	0	
9	0	1	0	1	1	
[illegible]	0	1	1	1	1	
[illegible]	0	1	0	1	1	
[illegible]	0	1	1	1	0	
13	0	1	0	1	1	
14	0	1	1	1	1	
15	0	1	0	1	1	
16	0	1	0	1	1	
17	0	1	0	1	1	
18	0	1	0	1	1	
19	0	1	1	1	1	
20	0	1	0	1	1	
21	0	1	0	1	1	
22	0	1	0	1	1	
23	0	1	1	1	1	
24	0	1	0	1	1	
25	0	1	1	1	1	

Importance Sampling

Fire	No Fire
0.05	0.95



CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7

P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no
CS	GS	0.95	0.05
CS	GS	0.7	0.3
CS	GS	0.9	0.1
CS	GS	0.01	0.99

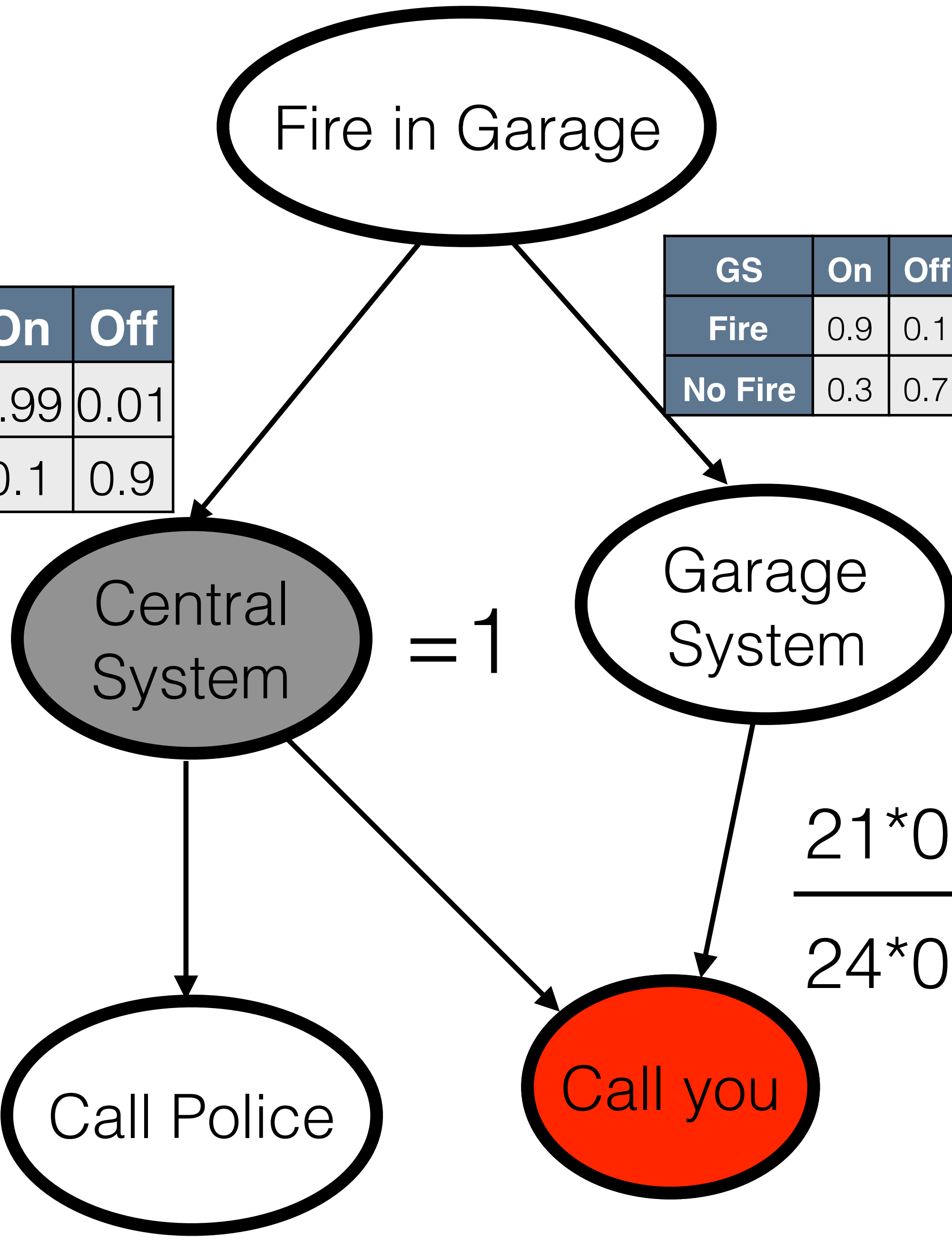
	F	CS	GS	P	Y	Weight
1	0	1	0	1	0	0.1
2	0	1	1	1	1	0.1
3	1	1	1	1	1	0.99
4	0	1	0	1	1	0.1
5	0	1	0	1	1	0.1
6	0	1	0	1	1	0.1
7	0	1	1	1	1	0.1
8	0	1	0	1	0	0.1
9	0	1	0	1	1	0.1
10	0	1	1	1	1	0.1
11	0	1	0	1	1	0.1
12	0	1	1	1	0	0.1
13	0	1	0	1	1	0.1
14	0	1	1	1	1	0.1
15	0	1	0	1	1	0.1
16	0	1	0	1	1	0.1
17	0	1	0	1	1	0.1
18	0	1	0	1	1	0.1
19	0	1	1	1	1	0.1
20	0	1	0	1	1	0.1
21	0	1	0	1	1	0.1
22	0	1	0	1	1	0.1
23	0	1	1	1	1	0.1
24	0	1	0	1	1	0.1
25	0	1	1	1	1	0.1

Importance Sampling

Fire	No Fire
0.05	0.95

CS	On	Off
Fire	0.99	0.01
No Fire	0.1	0.9

GS	On	Off
Fire	0.9	0.1
No Fire	0.3	0.7



$$\frac{21*0.1 + 1\text{x}0.99}{24*0.1 + 1\text{x}0.99}$$

P	Call	no call
CS on	1	0
CS off	0.01	0.99

Y		Call	no
CS	GS	0.95	0.05
CS	GS	0.7	0.3
CS	GS	0.9	0.1
CS	GS	0.01	0.99

	F	CS	GS	P	Y	Weight
1	0	1	0	1	0	0.1
2	0	1	1	1	1	0.1
3	1	1	1	1	1	0.99
4	0	1	0	1	1	0.1
5	0	1	0	1	1	0.1
6	0	1	0	1	1	0.1
7	0	1	1	1	1	0.1
8	0	1	0	1	0	0.1
9	0	1	0	1	1	0.1
10	0	1	1	1	1	0.1
11	0	1	0	1	1	0.1
12	0	1	1	1	0	0.1
13	0	1	0	1	1	0.1
14	0	1	1	1	1	0.1
15	0	1	0	1	1	0.1
16	0	1	0	1	1	0.1
17	0	1	0	1	1	0.1
18	0	1	0	1	1	0.1
19	0	1	1	1	1	0.1
20	0	1	0	1	1	0.1
21	0	1	0	1	1	0.1
22	0	1	0	1	1	0.1
23	0	1	1	1	1	0.1
24	0	1	0	1	1	0.1
25	0	1	1	1	1	0.1

Importance Sampling

- We really want to draw from distribution P.
- But we can only draw from distribution Q easily
- Trick:
  - Draw $x_1, \ldots, x_n \sim Q$
  - Re-weight each sample x_t by $P(X = x_t)/Q(X = x_t)$

Importance Sampling

- Why does it work?

$$\mathbb{E}_{X\sim P}[f(X)] = \sum_x P(X=x)f(x)$$

Importance Sampling

- Why does it work?

$$\begin{aligned}
\mathbb{E}_{X\sim P}[f(X)] &= \sum_x P(X=x)f(x) \\
&= \sum_x Q(X=x)\left(\frac{P(X=x)}{Q(X=x)}f(x)\right)
\end{aligned}$$

Importance Sampling

- Why does it work?

$$\begin{aligned}\mathbb{E}_{X\sim P}[f(X)] &= \sum_x P(X=x)f(x) \\ &= \sum_x Q(X=x)\left(\frac{P(X=x)}{Q(X=x)}f(x)\right) \\ &= \mathbb{E}_{X\sim Q}\left[\frac{P(X)}{Q(X)}f(X)\right]\end{aligned}$$

Importance Sampling

- Why does it work?

$$
\begin{aligned}
\mathbb{E}_{X\sim P}[f(X)] &= \sum_x P(X=x)f(x) \\
&= \sum_x Q(X=x)\left(\frac{P(X=x)}{Q(X=x)}f(x)\right) \\
&= \mathbb{E}_{X\sim Q}\left[\frac{P(X)}{Q(X)}f(X)\right] \\
&\approx \frac{1}{n}\sum_{t=1}^{n}\frac{P(X=x_t)}{Q(X=x_t)}f(x_t)
\end{aligned}
$$

Importance Sampling

- Why does it work?

$$
\begin{aligned}
\mathbb{E}_{X\sim P}[f(X)] &= \sum_x P(X=x)f(x) \\
&= \sum_x Q(X=x)\left(\frac{P(X=x)}{Q(X=x)}f(x)\right) \\
&= \mathbb{E}_{X\sim Q}\left[\frac{P(X)}{Q(X)}f(X)\right] \\
&\approx \frac{1}{n}\sum_{t=1}^{n}\frac{P(X=x_t)}{Q(X=x_t)}f(x_t)
\end{aligned}
$$

- Example: $f(X) = \mathbf{1}\{X \in \mathsf{Set}\}$, then $\mathbb{E}_{X\sim P}[f(X)] = P(X \in \mathsf{Set})$

Importance Sampling

- Why does it work?

$$\begin{aligned}
\mathbb{E}_{X\sim P}[f(X)] &= \sum_x P(X=x)f(x) \\
&= \sum_x Q(X=x)\left(\frac{P(X=x)}{Q(X=x)}f(x)\right) \\
&= \mathbb{E}_{X\sim Q}\left[\frac{P(X)}{Q(X)}f(X)\right] \\
&\approx \frac{1}{n}\sum_{t=1}^{n}\frac{P(X=x_t)}{Q(X=x_t)}f(x_t)
\end{aligned}$$

- Example: $f(X) = \mathbf{1}\{X \in \text{Set}\}$, then $\mathbb{E}_{X\sim P}[f(X)] = P(X \in \text{Set})$
- Hence, using importance weighted sampling,

$$P(X \in \text{Set}) \approx \frac{1}{n}\sum_{t=1}^{n}\mathbf{1}\{x_t \in \text{Set}\}\frac{P(X=x_t)}{Q(X=x_t)}$$

Importance Sampling

Importance Sampling

Importance Sampling

$$P(1) = 0.9, \quad \forall j \neq 1 \ P(j) = 0.1/5$$

Importance Sampling

$P(1) = 0.9, \quad \forall j \neq 1 \; P(j) = 0.1/5$

Importance Sampling

$P(1) = 0.9, \quad \forall j \neq 1 \; P(j) = 0.1/5$

$\forall j \quad Q(j) = 1/6$

Importance Sampling

$P(1) = 0.9, \quad \forall j \neq 1 \; P(j) = 0.1/5$

$\forall j \quad Q(j) = 1/6$

$$\text{Set} = \{2, 4, 6\}$$

Importance Sampling

$$P(1) = 0.9, \quad \forall j \neq 1 \ P(j) = 0.1/5 \qquad\qquad \forall j \quad Q(j) = 1/6$$

$$\text{Set} = \{2, 4, 6\}$$

What is P(even)?

IMPORTANCE SAMPLING

$P(1) = 0.9, \quad \forall j \neq 1 \; P(j) = 0.1/5$ $\qquad \forall j \quad Q(j) = 1/6$

$$\text{Set} = \{2, 4, 6\}$$

What is P(even)?

$$\frac{1}{n}\sum_{t=1}^{n} \mathbf{1}\{x_t \in \{2,4,6\}\}\frac{P(x_t)}{Q(x_t)} = \frac{1}{n}\sum_{t=1}^{n} \mathbf{1}\{x_t \in \{2,4,6\}\}\frac{0.1/5}{1/6}$$

IMPORTANCE SAMPLING

$$P(1) = 0.9, \quad \forall j \neq 1 \; P(j) = 0.1/5 \qquad\qquad \forall j \quad Q(j) = 1/6$$

$$\text{Set} = \{2, 4, 6\}$$

What is P(even)?

$$\frac{1}{n}\sum_{t=1}^{n} \mathbf{1}\{x_t \in \{2,4,6\}\} \frac{P(x_t)}{Q(x_t)} = \frac{1}{n}\sum_{t=1}^{n} \mathbf{1}\{x_t \in \{2,4,6\}\} \frac{0.1/5}{1/6}$$

$$= 0.12 \times \frac{1}{n}\sum_{t=1}^{n} \mathbf{1}\{x_t \in \{2,4,6\}\} \approx 0.12 \times 0.5 = 0.06$$

Importance Sampling

Likelihood weighting:

Topologically sort variables (parents first children later)

For $t = 1$ to n (number of samples)

Set $w_t = 1$

For $i = 1$ to N (number of variables)

If X_i is observed,

Set $w_t \leftarrow w_t \cdot P(X_i = x_i | \text{Parents}(X_i) = \text{ already sampled})$

Set $x_i^t = x_i$ (the observed value)

Else, sample $x_i^t \sim P(X_i | \text{Parents}(X_i) = \text{ already sampled})$

End For

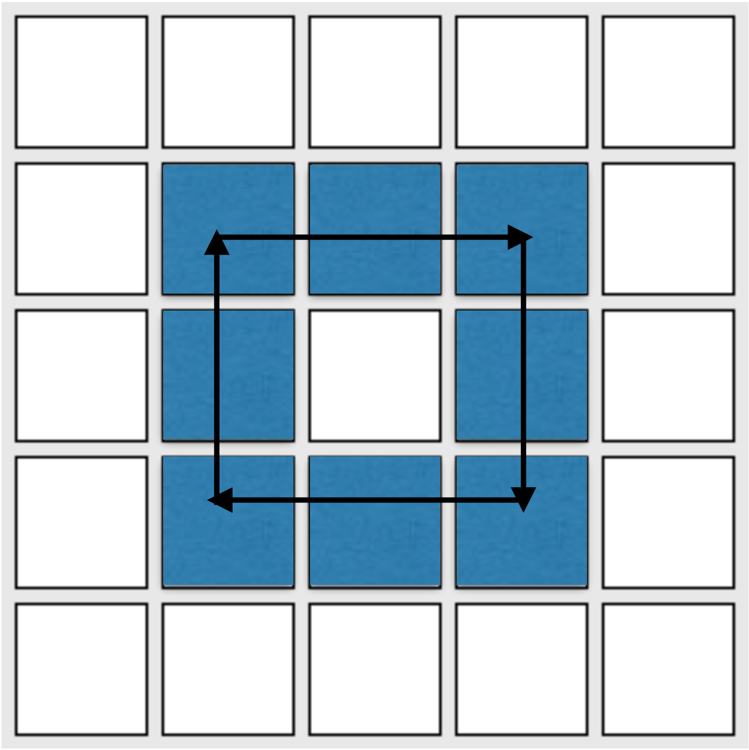
End For

Output,

$$P(\text{Variable} = \text{value} | \text{Observation}) = \frac{\sum_{t=1}^{n} w_t \mathbf{1}\{\text{Variable} = \text{value}\}}{\sum_{t=1}^{n} w_t}$$

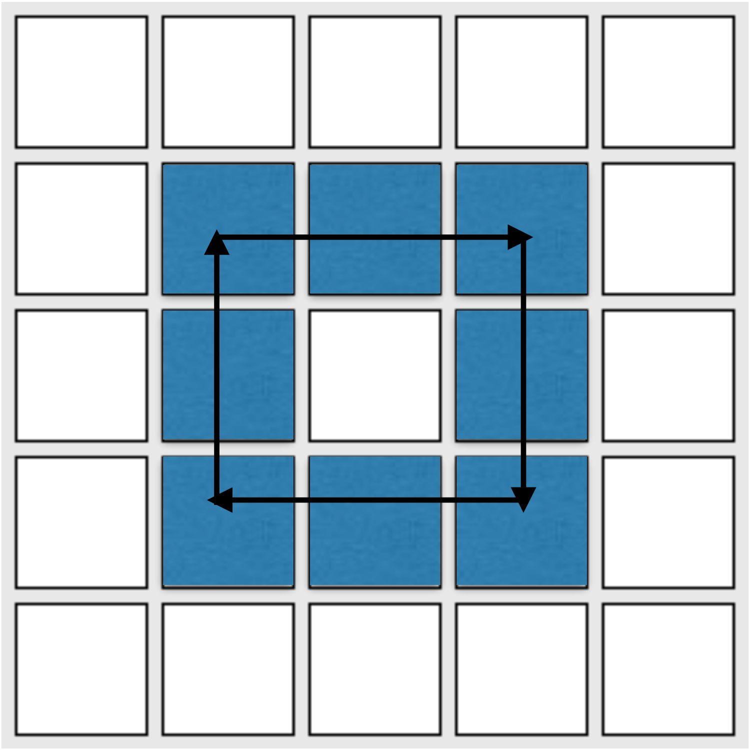
Hidden Markov Model (HMM)

Hidden Markov Model (HMM)



Hidden Markov Model (HMM)

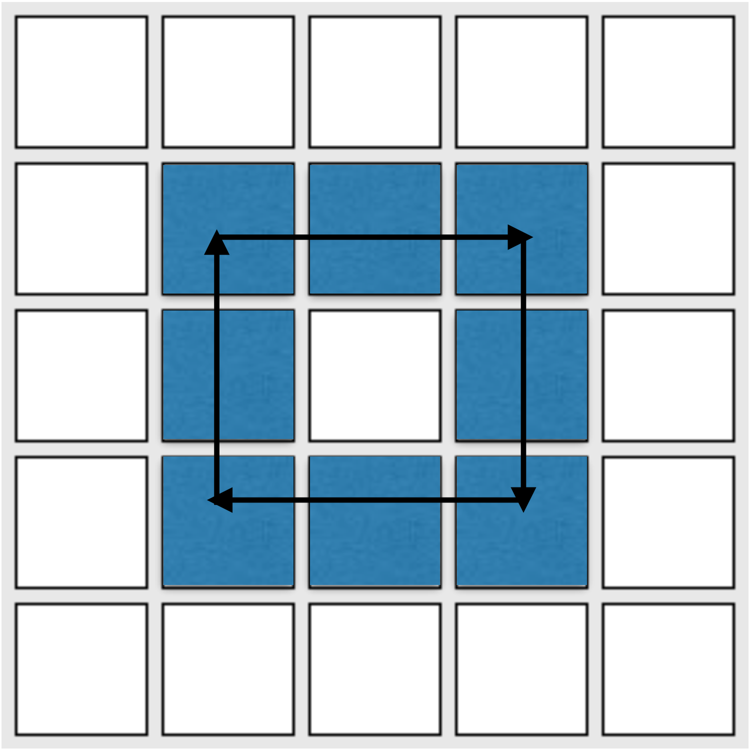
Example:



Hidden Markov Model (HMM)

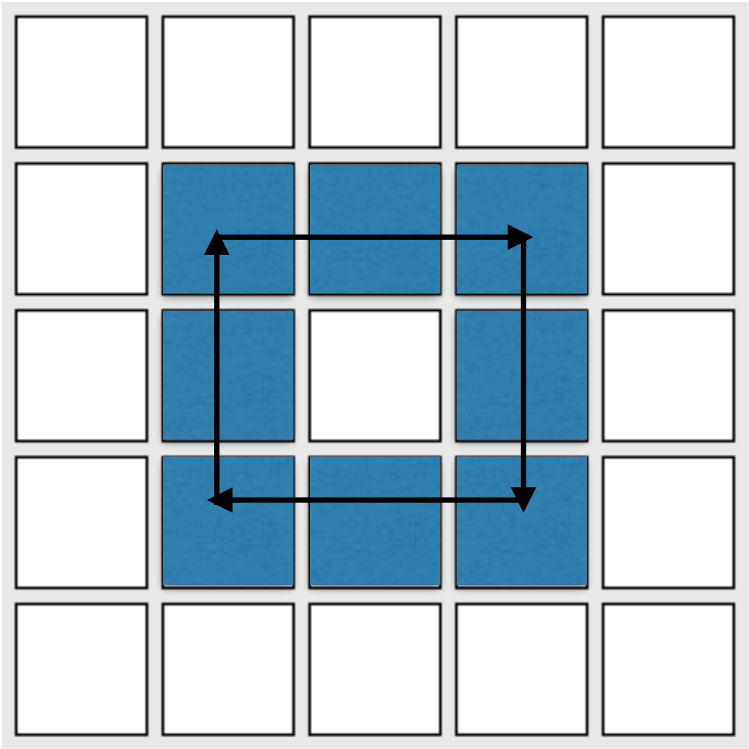
Example:

But you don't observe location (dark room)



Hidden Markov Model (HMM)

Example:

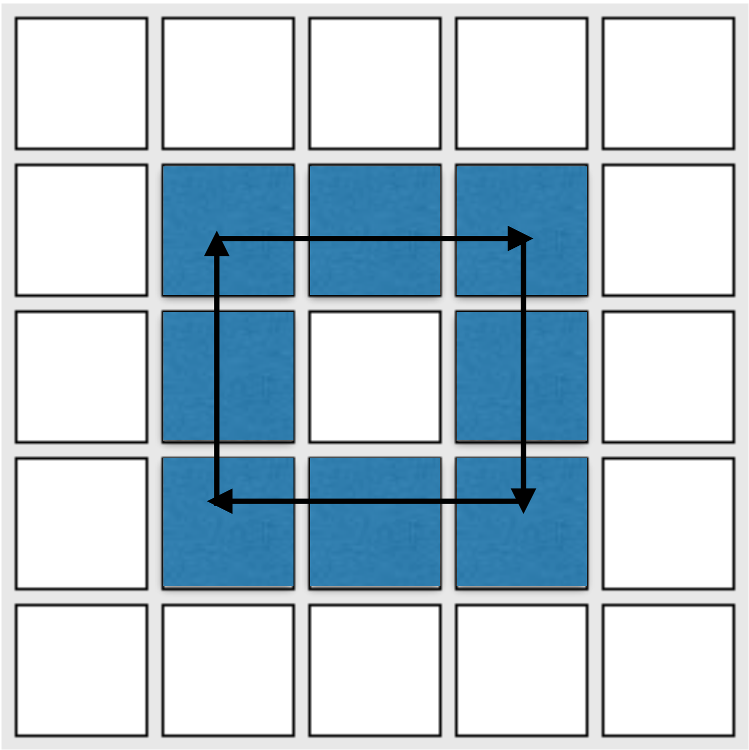


But you don't observe location (dark room)

You hear how close the bot is!

HIDDEN MARKOV MODEL (HMM)

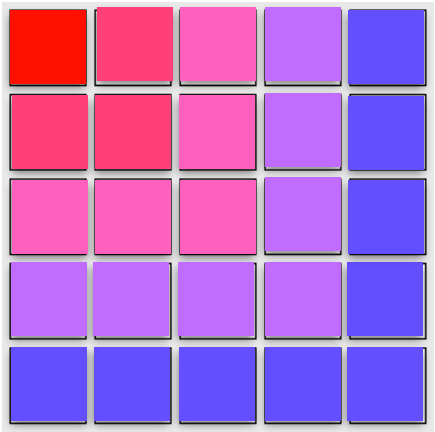
Example:



But you don't observe location (dark room)

You hear how close the bot is!

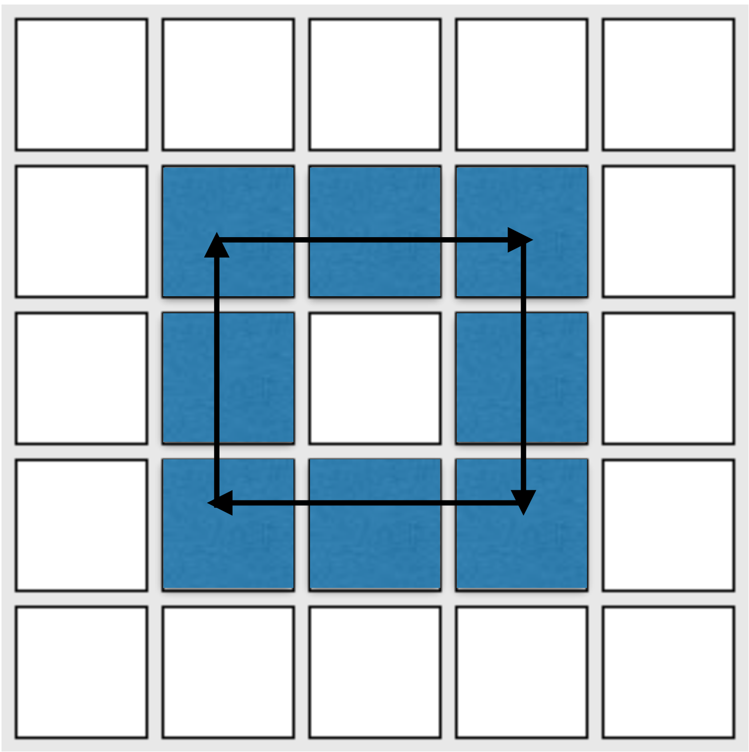
What you hear:



+ noise

Hidden Markov Model (HMM)

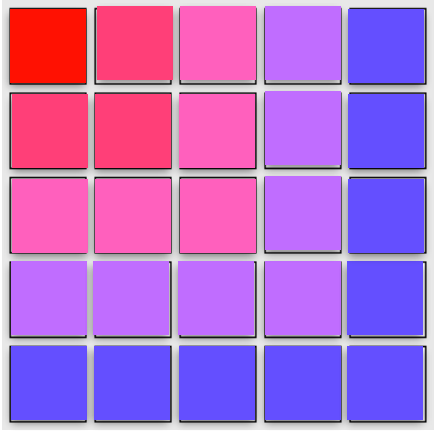
Example:



But you don't observe location (dark room)

You hear how close the bot is!

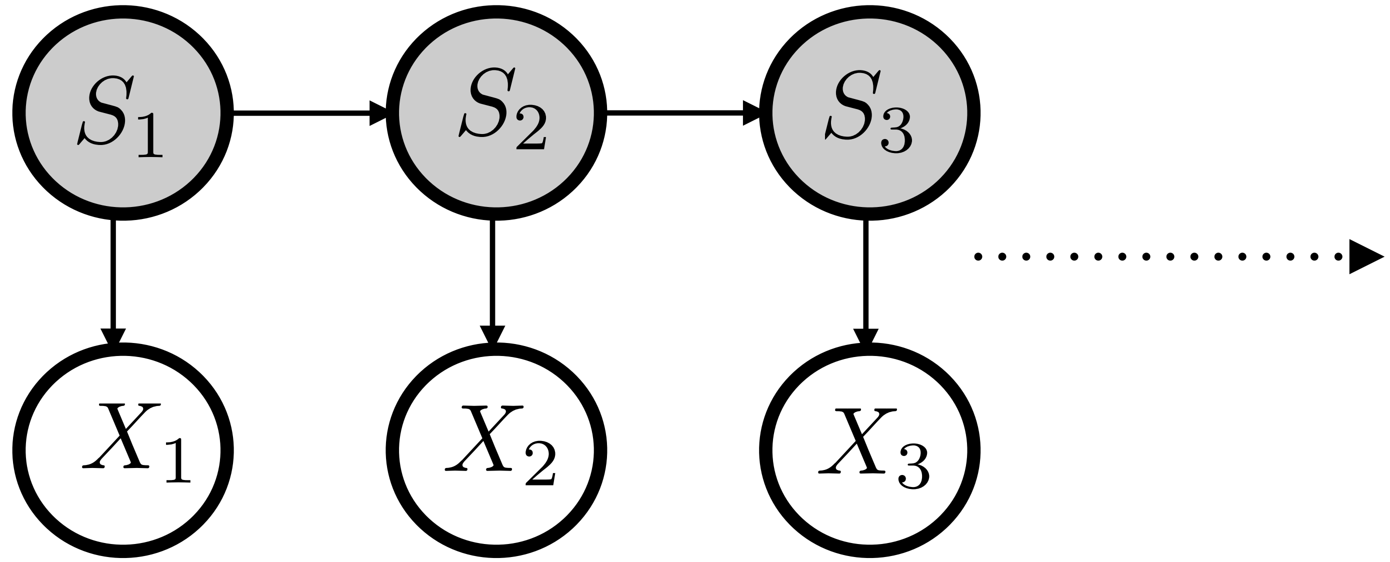
What you hear:



+ noise

Can you catch the Bot?

Hidden Markov Model (HMM)



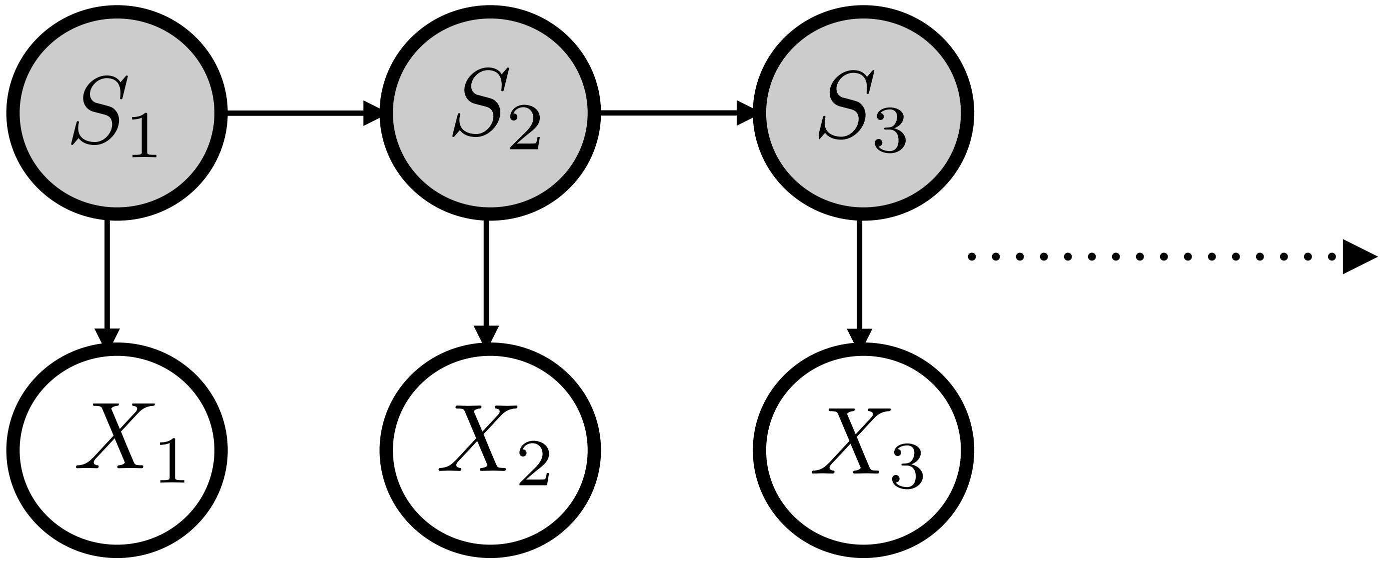
X_t's are what you hear (observation)

S_t's are the unseen locations (states)

Eg: for m x m grid we have, $K = m^2$ states

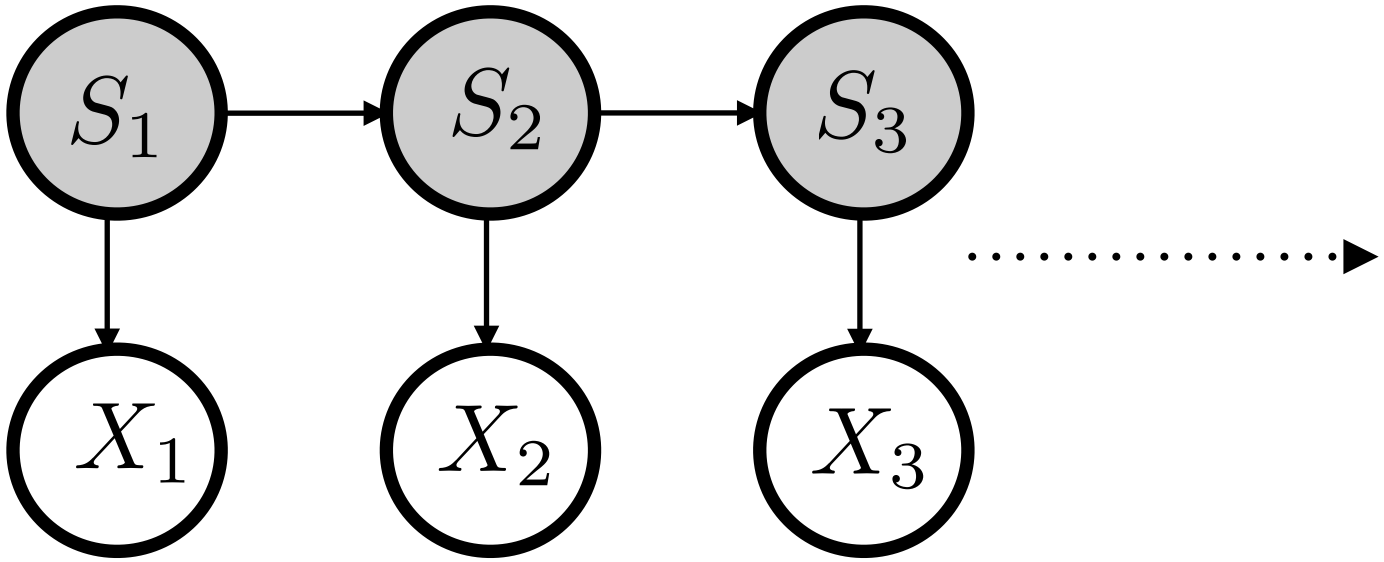
Number of alphabets = # colors you can observe

Hidden Markov Model (HMM)



Eg: for m x m grid we have, K = m^2 states

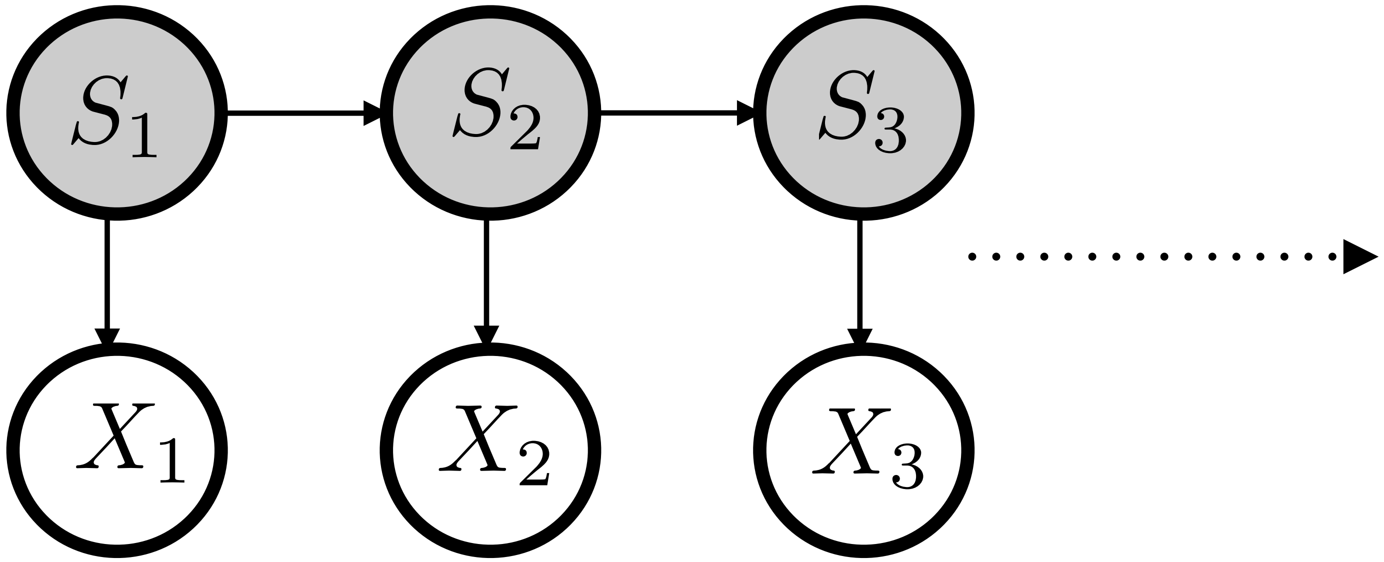
Hidden Markov Model (HMM)



Eg: for m x m grid we have, $K = m^2$ states

Transition matrix is K x K (too large)

Hidden Markov Model (HMM)



Eg: for m x m grid we have, $K = m^2$ states

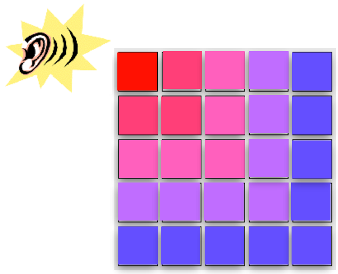
Transition matrix is K x K (too large)

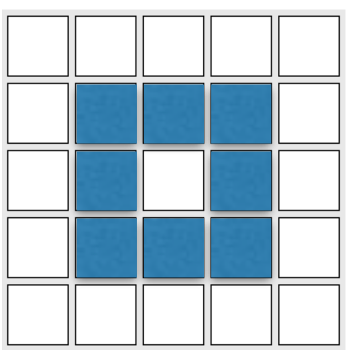
Use sampling to do approximate inference

Number of samples $n << m^4$

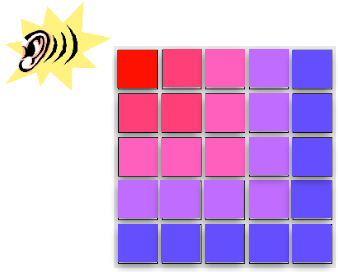
Hidden Markov Model (HMM)

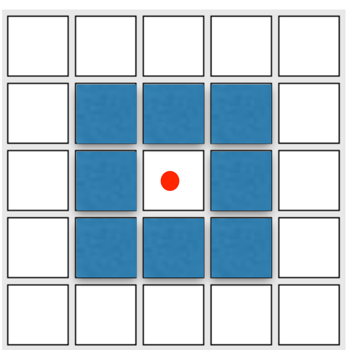
Hidden Markov Model (HMM)



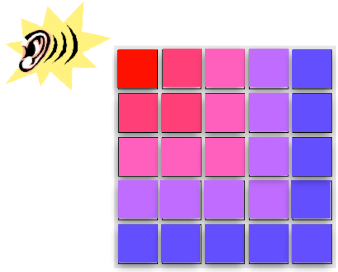


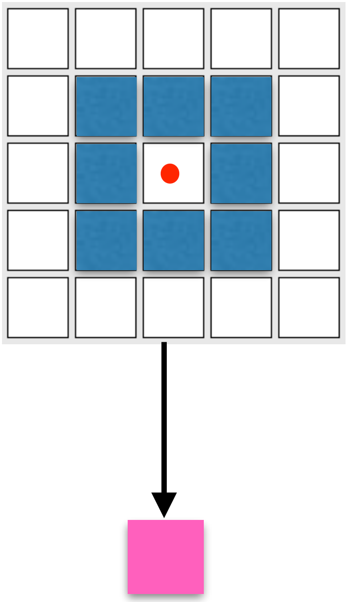
Hidden Markov Model (HMM)



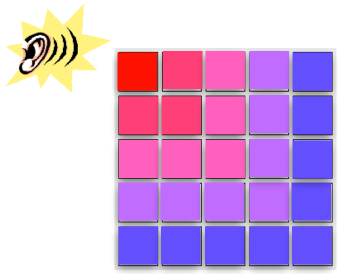


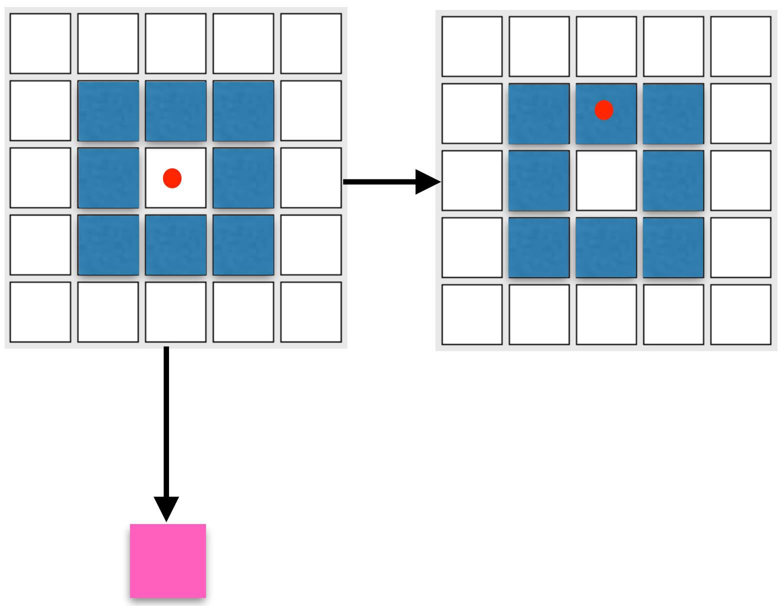
Hidden Markov Model (HMM)



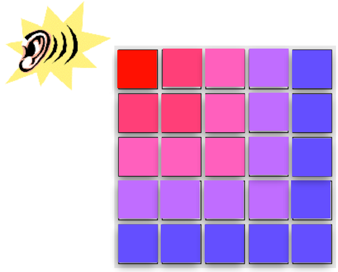


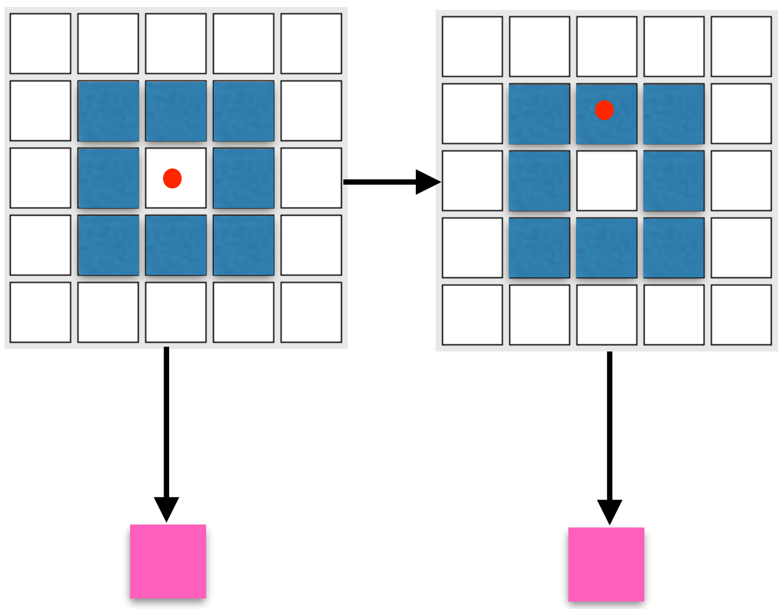
HIDDEN MARKOV MODEL (HMM)



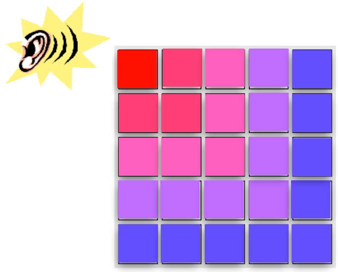


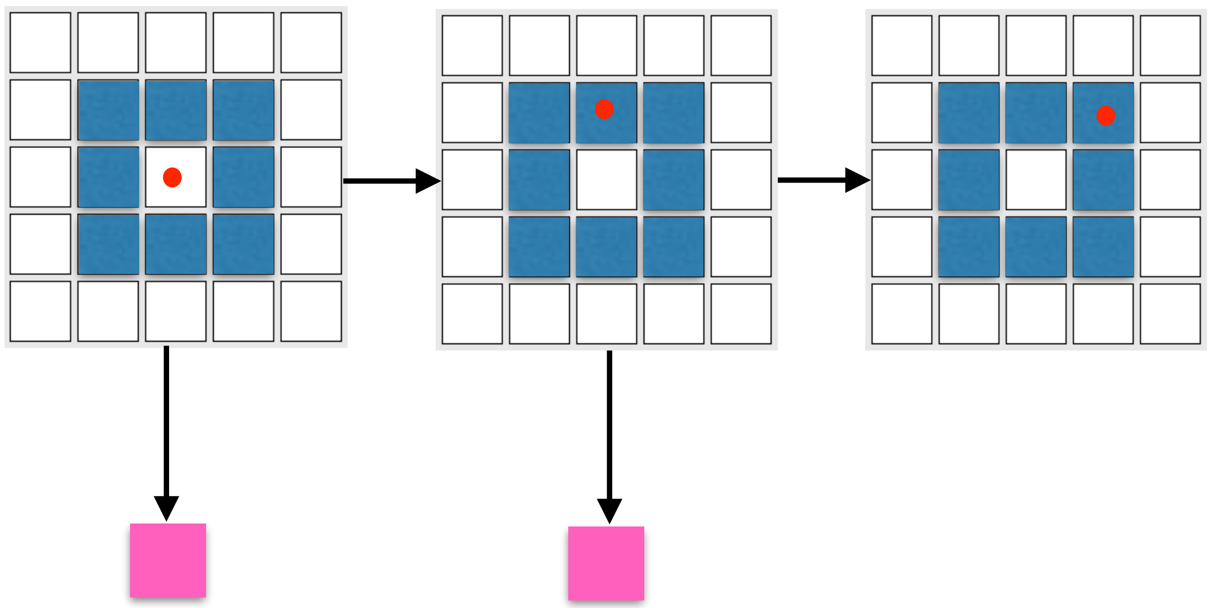
Hidden Markov Model (HMM)



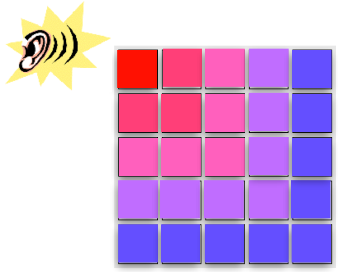


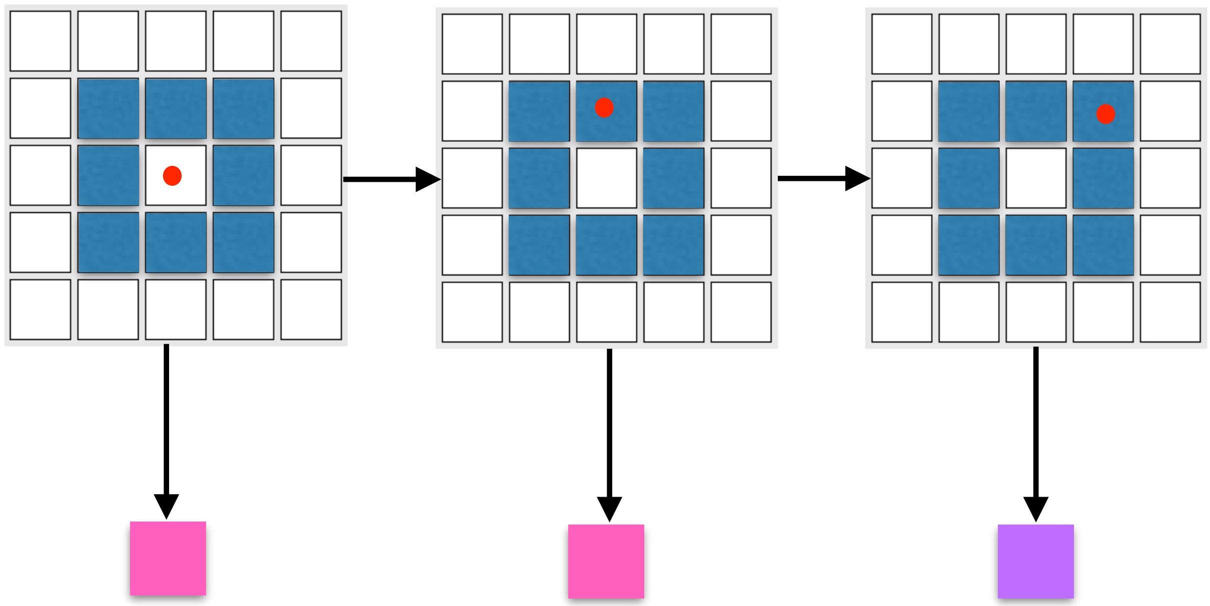
HIDDEN MARKOV MODEL (HMM)



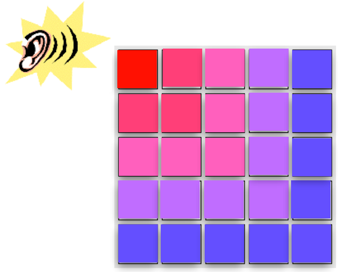


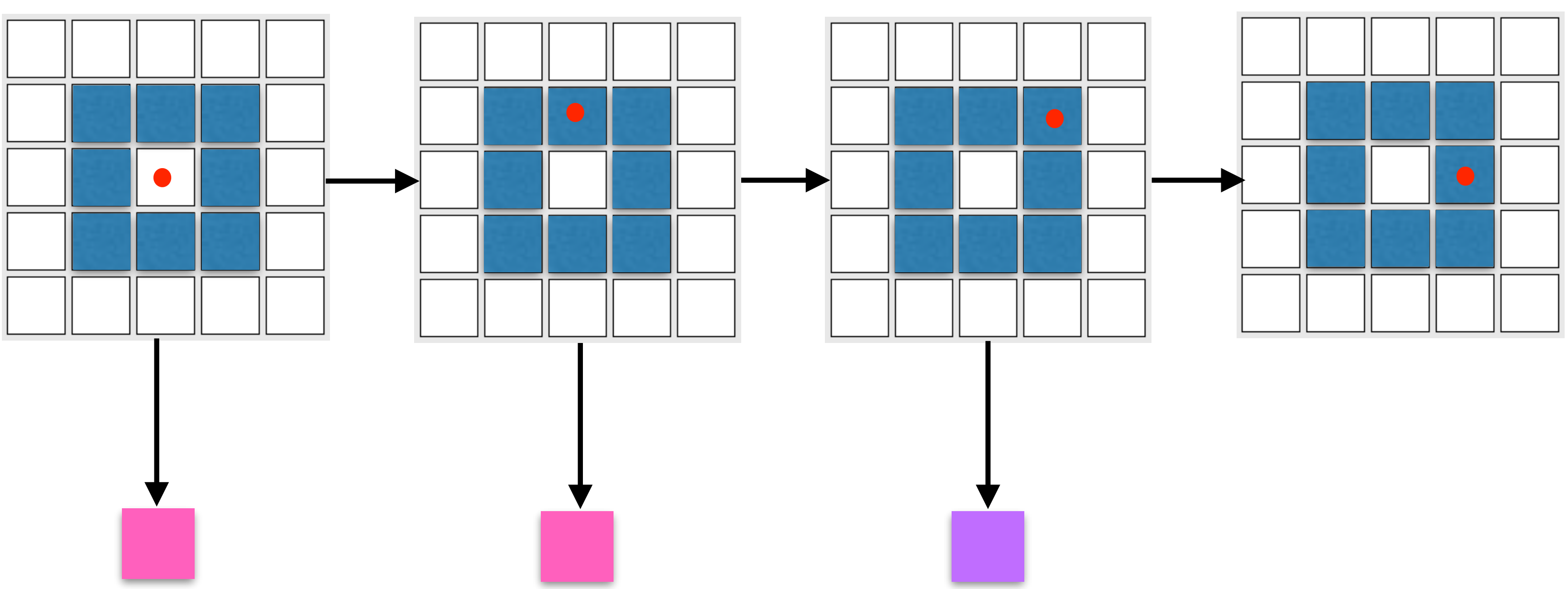
Hidden Markov Model (HMM)



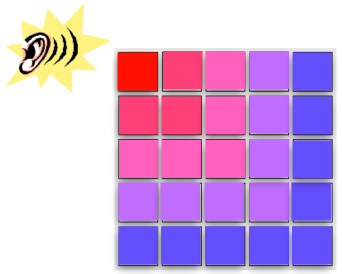


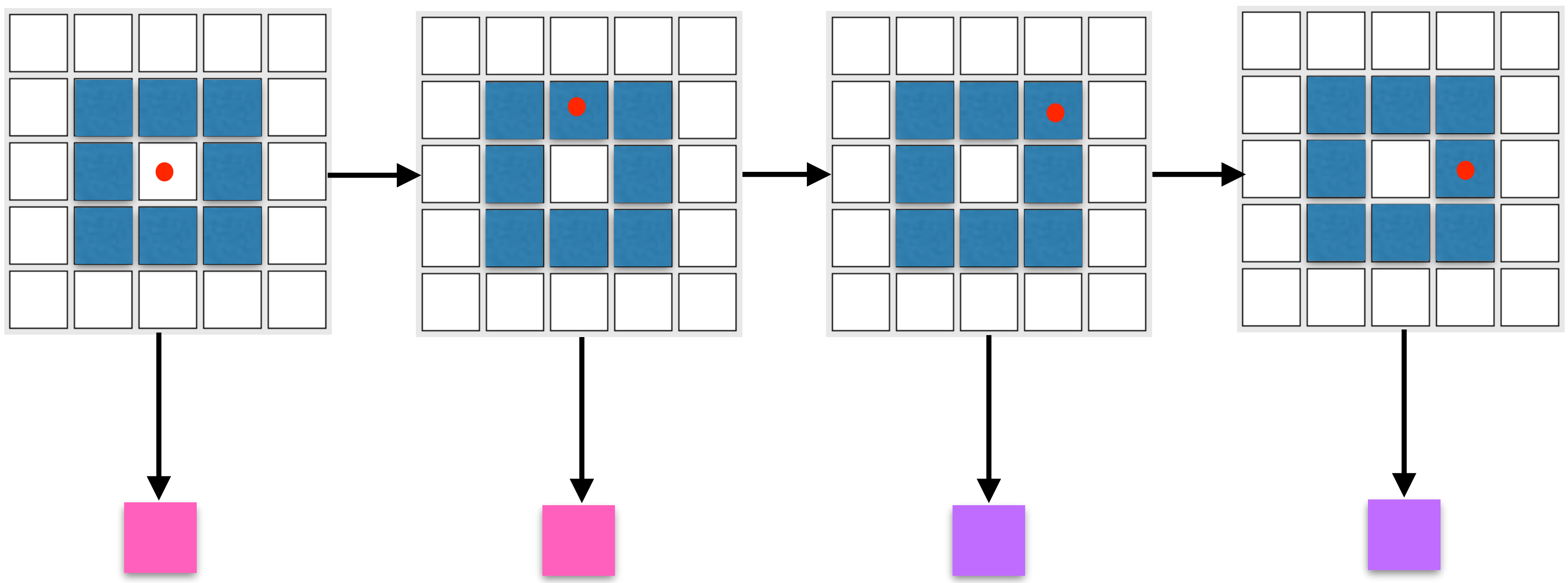
Hidden Markov Model (HMM)



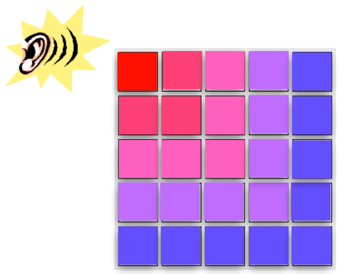


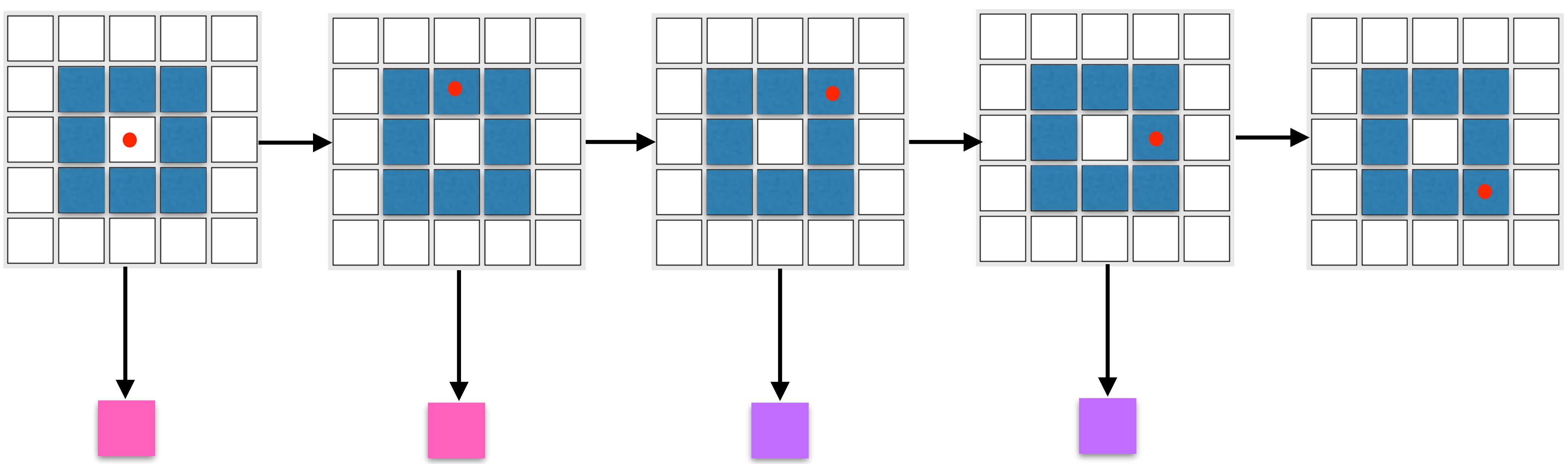
HIDDEN MARKOV MODEL (HMM)



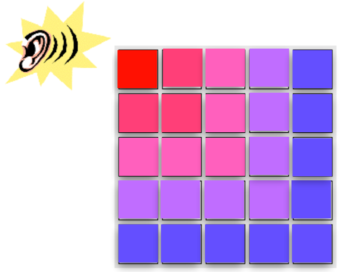


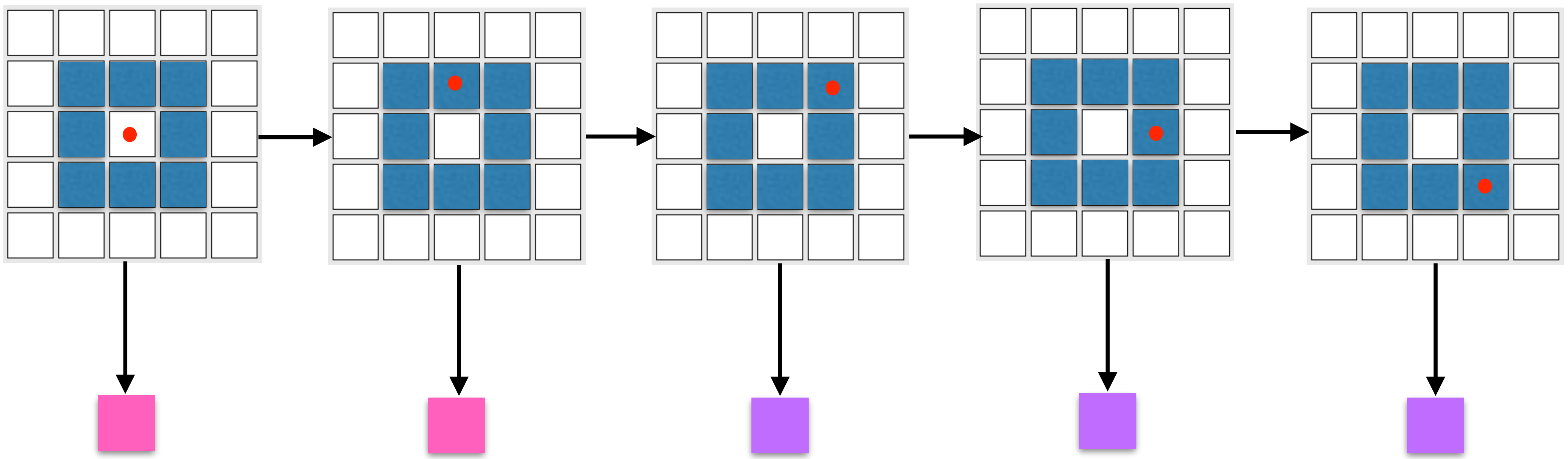
HIDDEN MARKOV MODEL (HMM)



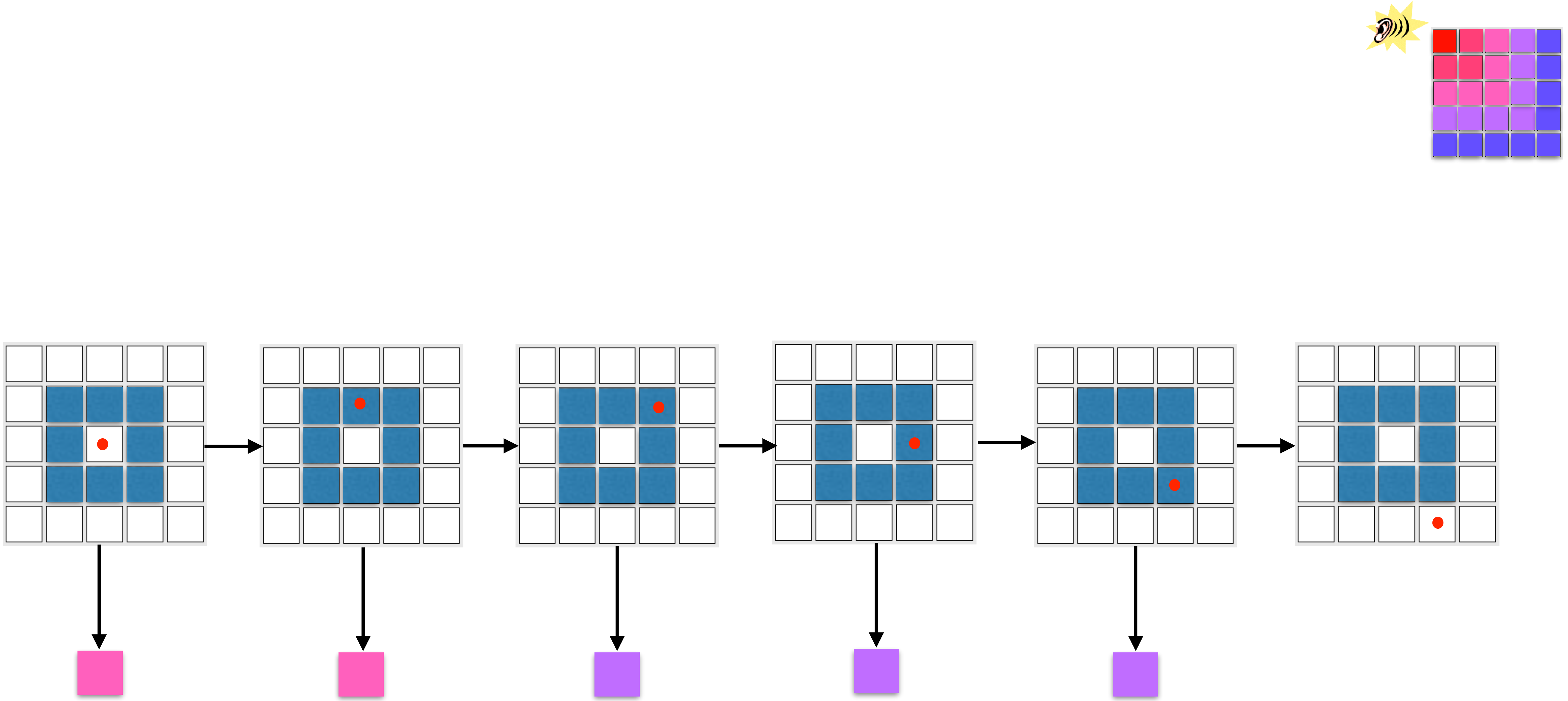


Hidden Markov Model (HMM)

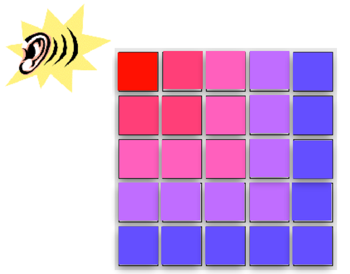


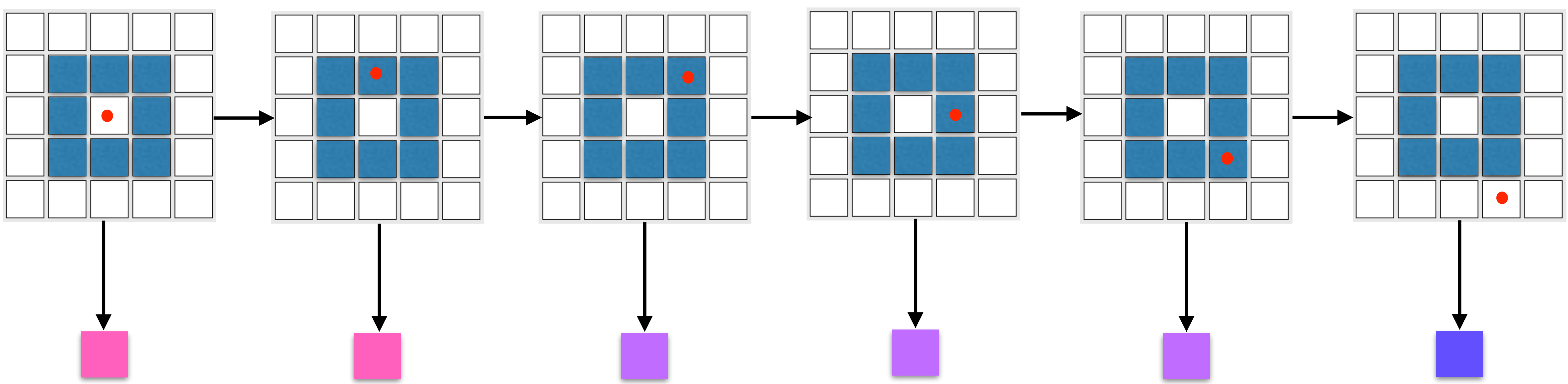


Hidden Markov Model (HMM)

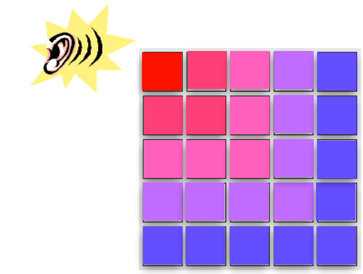


Hidden Markov Model (HMM)

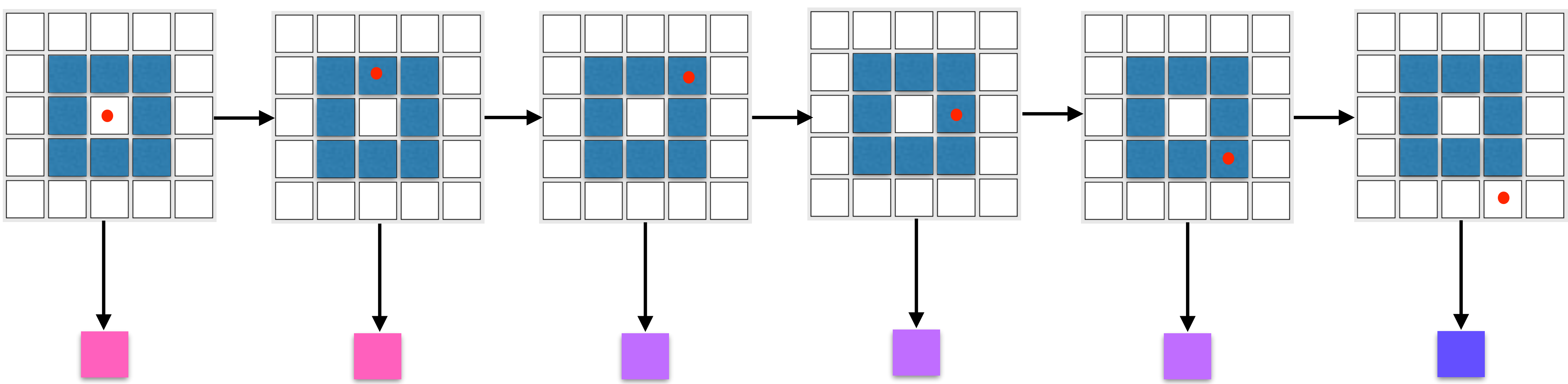




Hidden Markov Model (HMM)

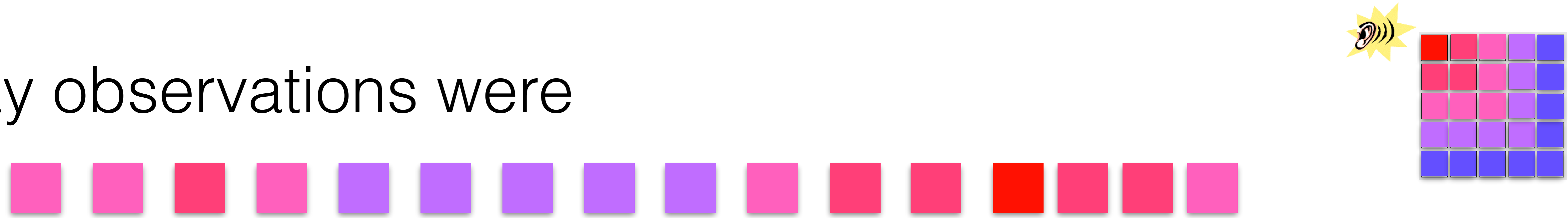


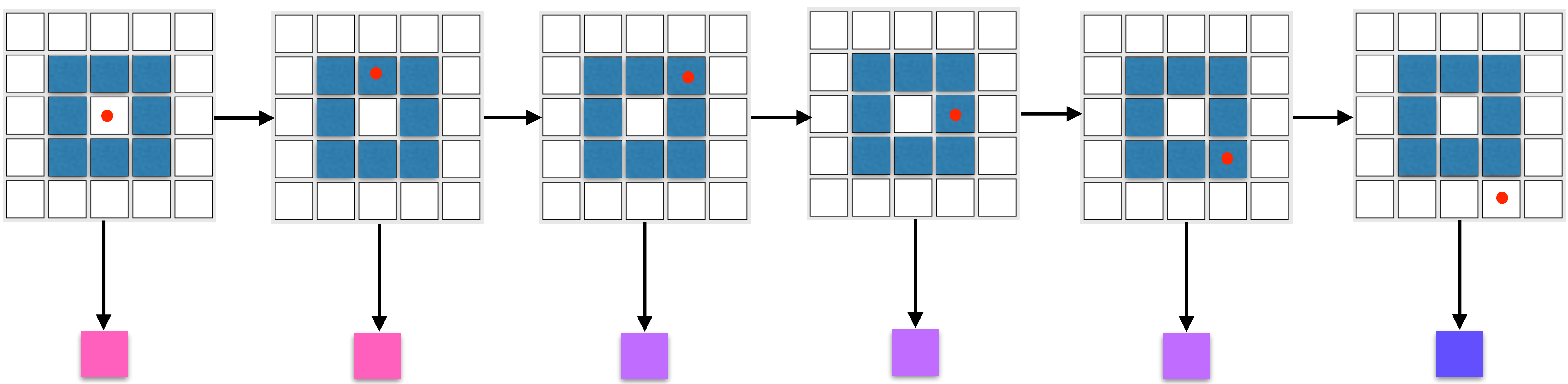
Eg: say observations were



Hidden Markov Model (HMM)

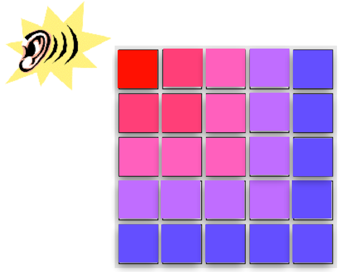
Eg: say observations were



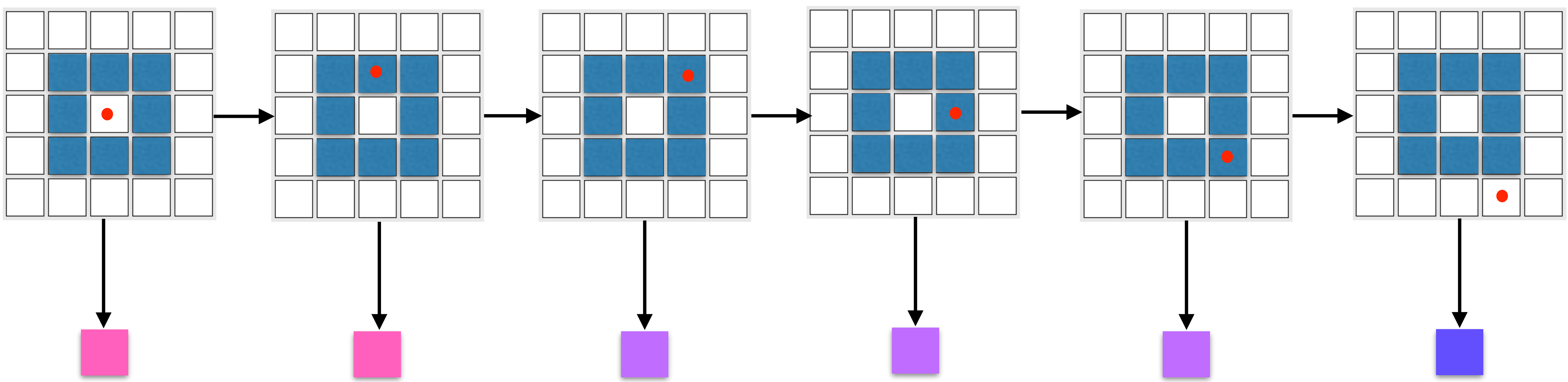


Rejection sampling: Reject samples that don't match observations

Hidden Markov Model (HMM)



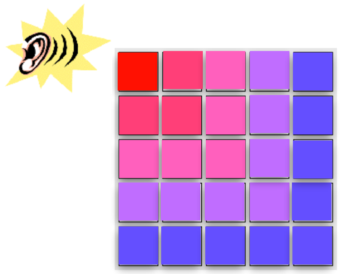
Eg: say observations were



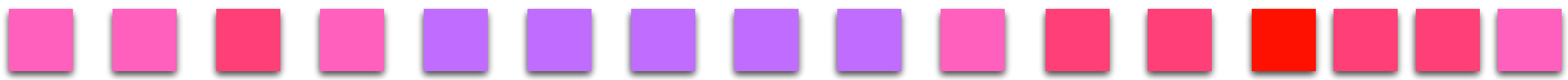
Rejection sampling: Reject samples that don't match observations

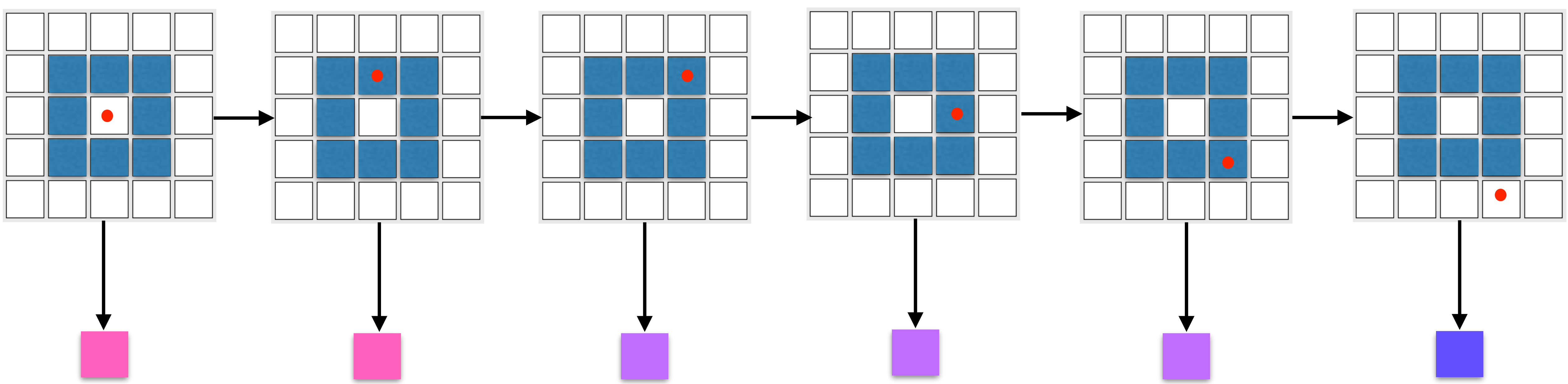
Most samples rejected

Hidden Markov Model (HMM)

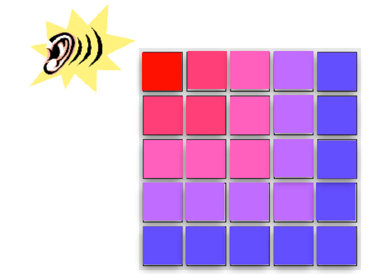


Eg: say observations were

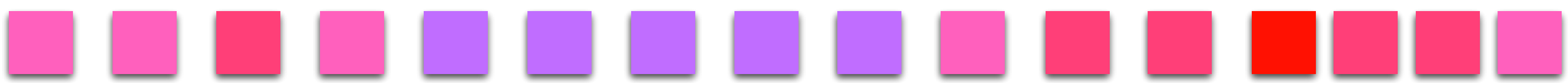


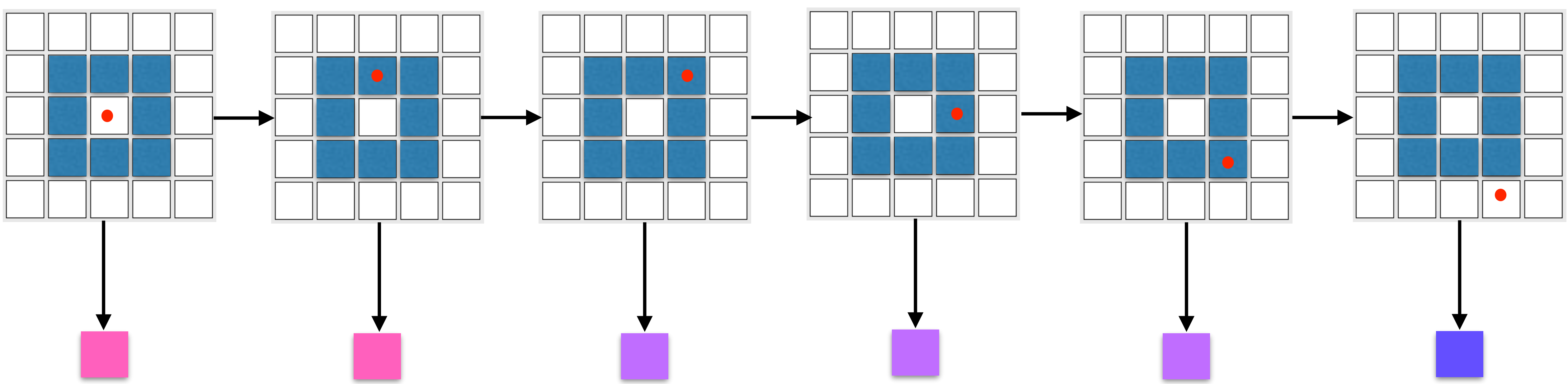


Hidden Markov Model (HMM)



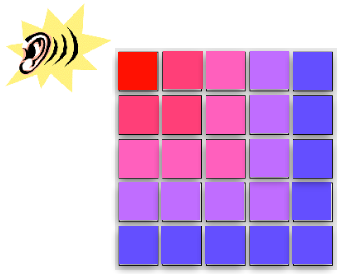
Eg: say observations were



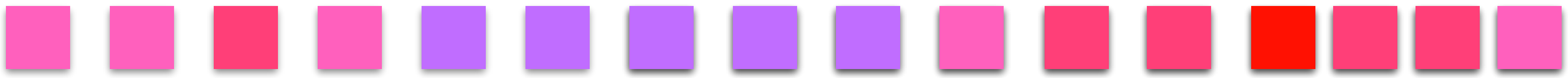


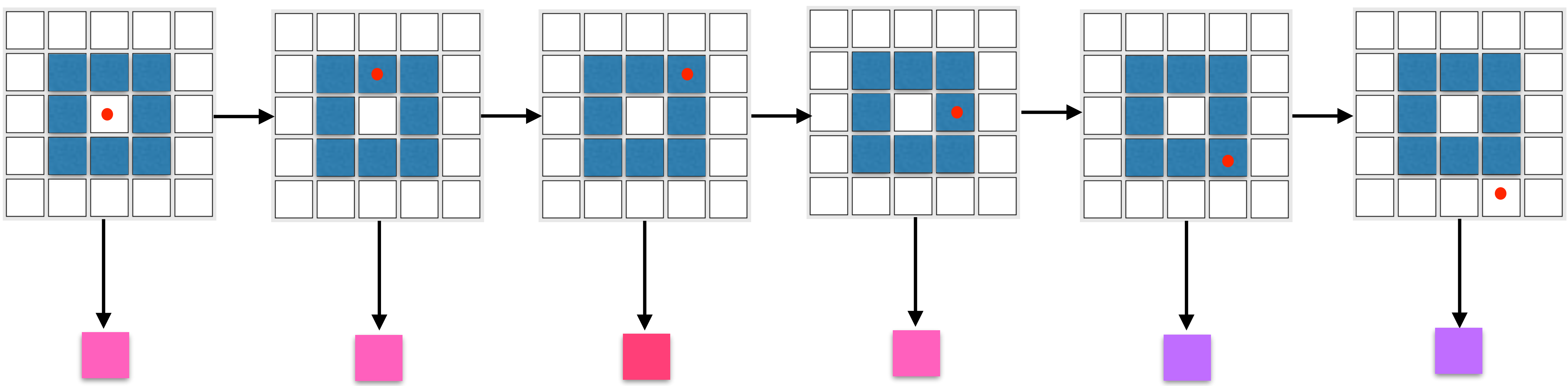
Importance weighting: weight samples

Hidden Markov Model (HMM)



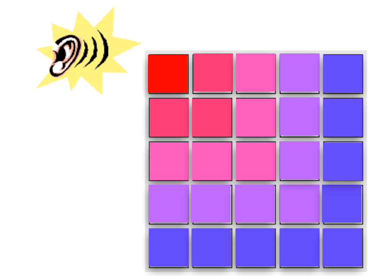
Eg: say observations were



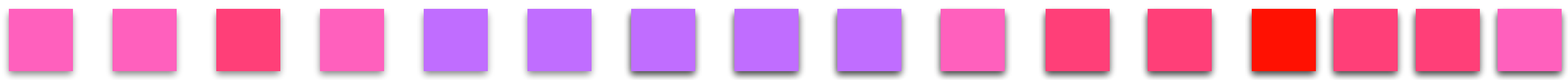


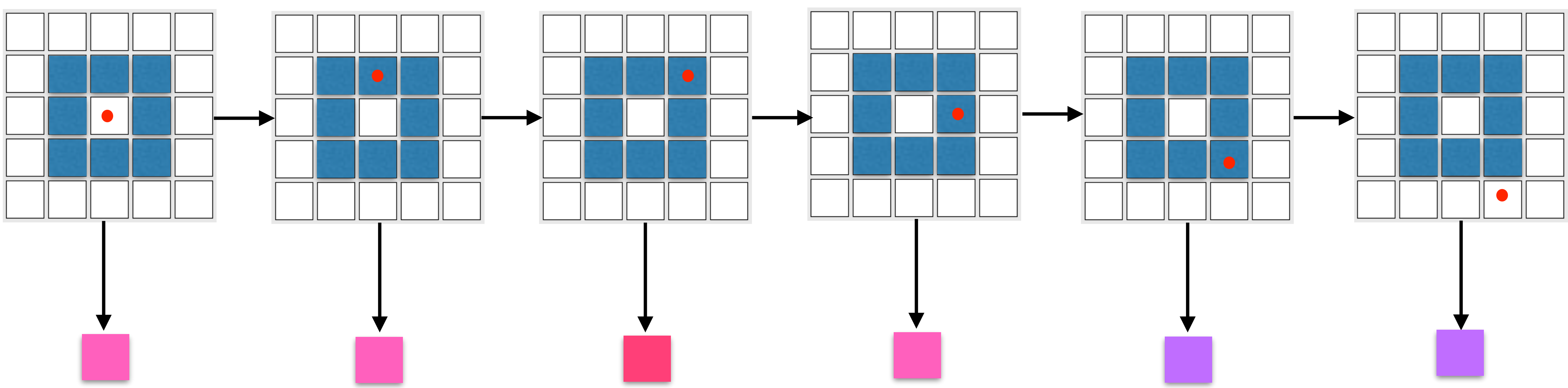
Importance weighting: weight samples

HIDDEN MARKOV MODEL (HMM)



Eg: say observations were

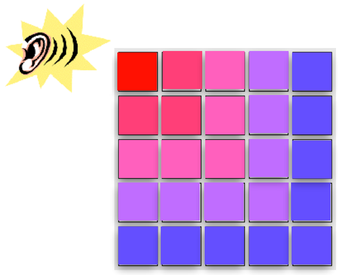




$P(\ \mid X_1=13) \times P(\ \mid X_2=8) \times P(\ \mid X_3=9) \times P(\ \mid X_5=24) \times P(\ \mid X_5=19) \times P(\ \mid X_4=14)$

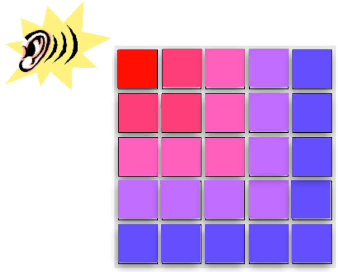
Importance weighting: weight samples

HMM Particle Filter

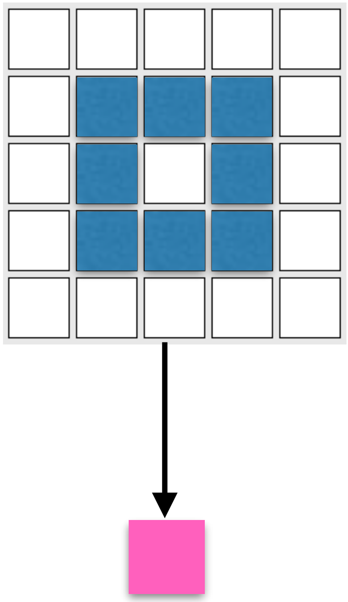


- Use multiple samples and track each ones weights.

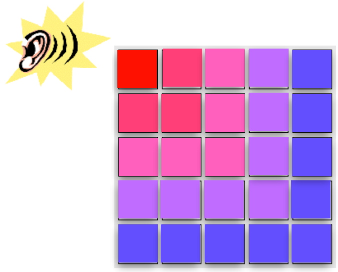
HMM Particle Filter



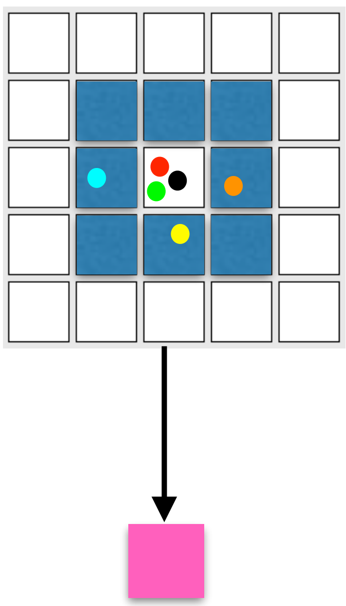
- Use multiple samples and track each ones weights.



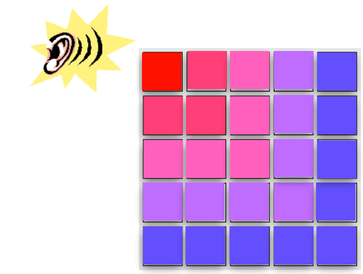
HMM Particle Filter



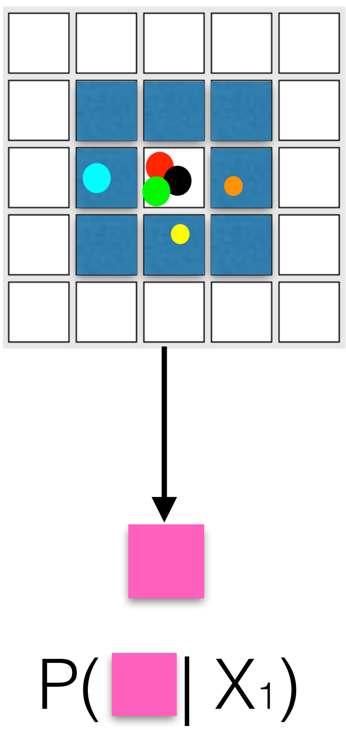
- Use multiple samples and track each ones weights.



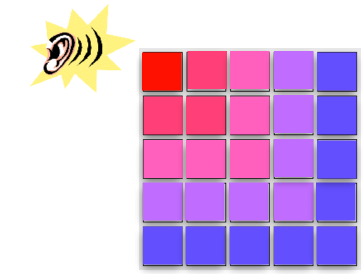
HMM Particle Filter



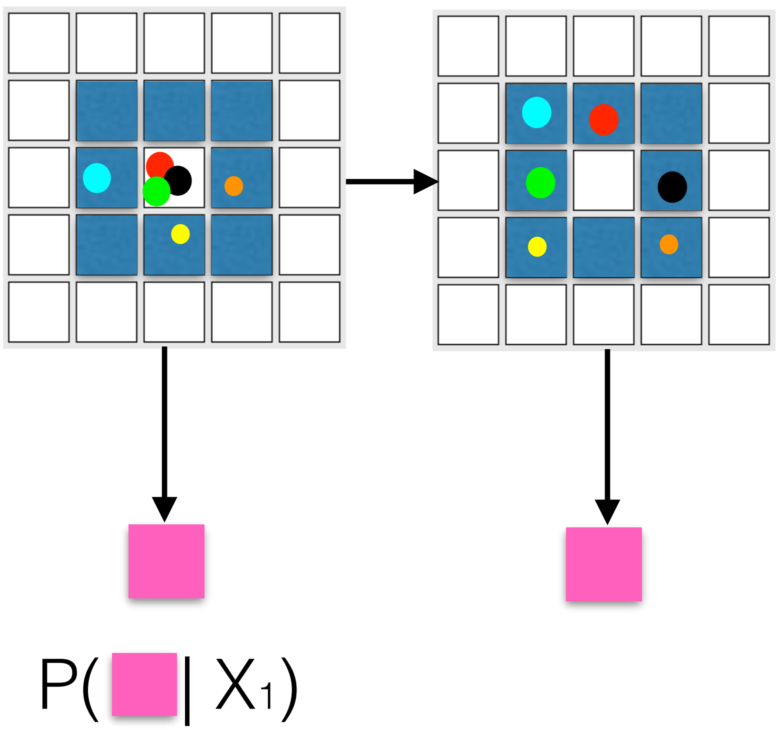
- Use multiple samples and track each ones weights.



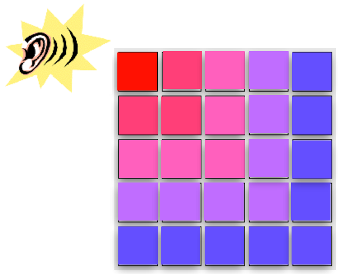
HMM Particle Filter



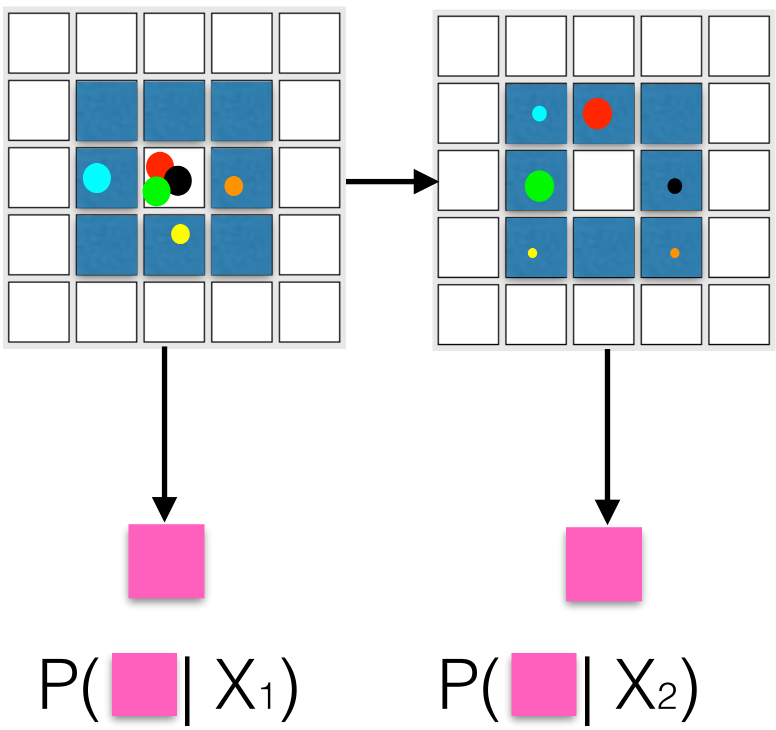
- Use multiple samples and track each ones weights.



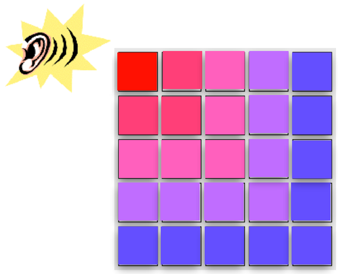
HMM Particle Filter



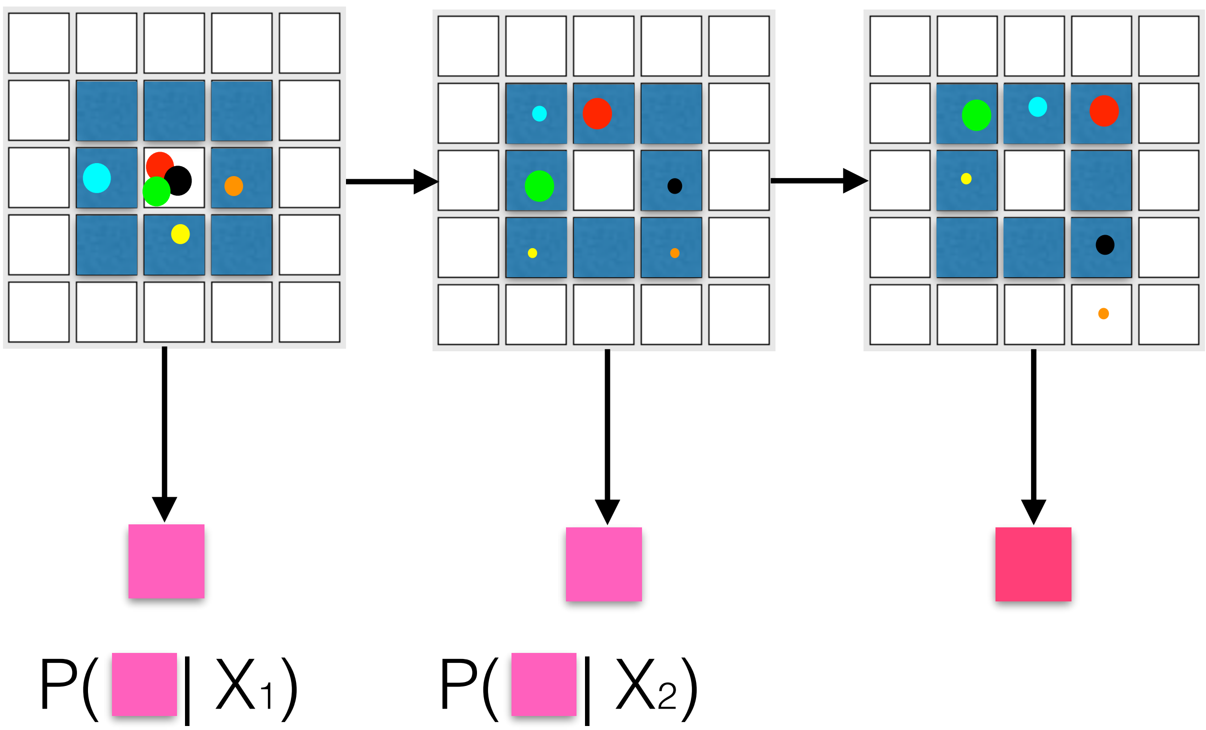
- Use multiple samples and track each ones weights.



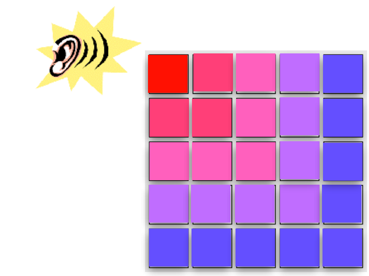
HMM Particle Filter



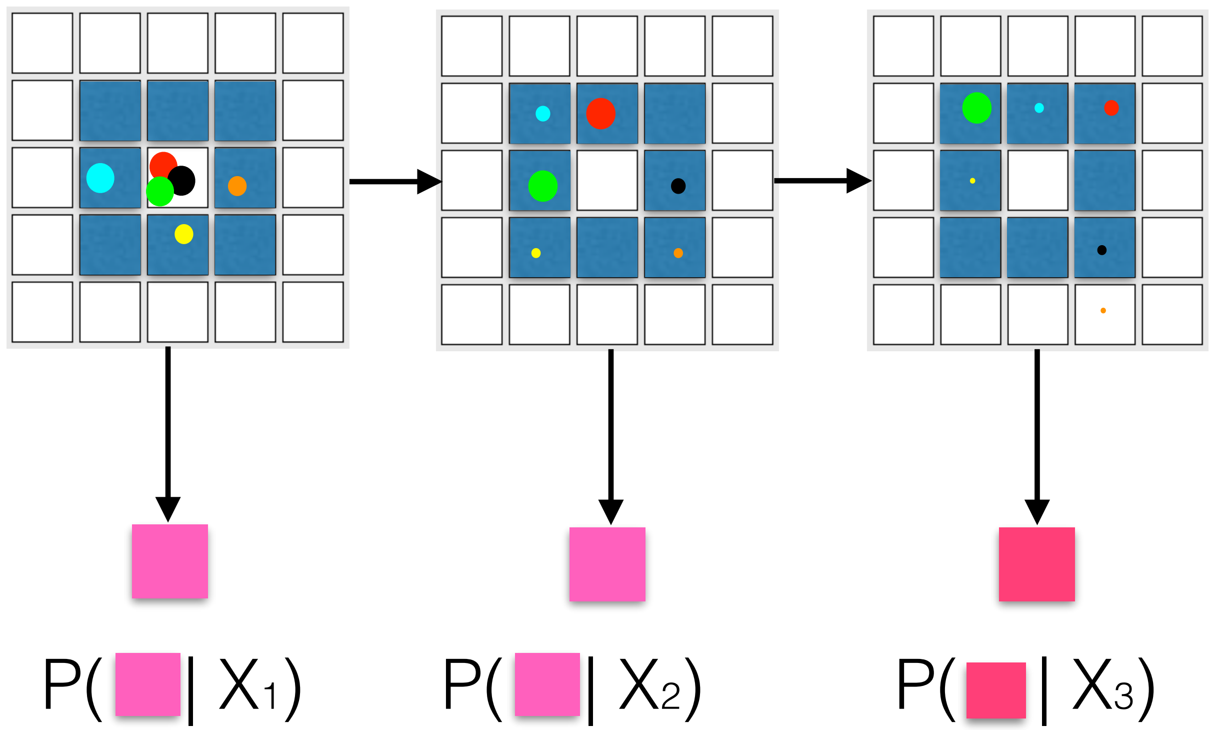
- Use multiple samples and track each ones weights.



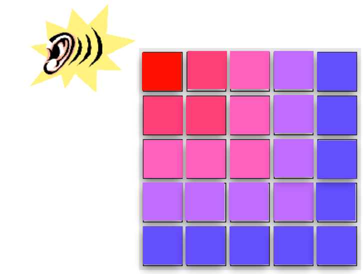
HMM Particle Filter



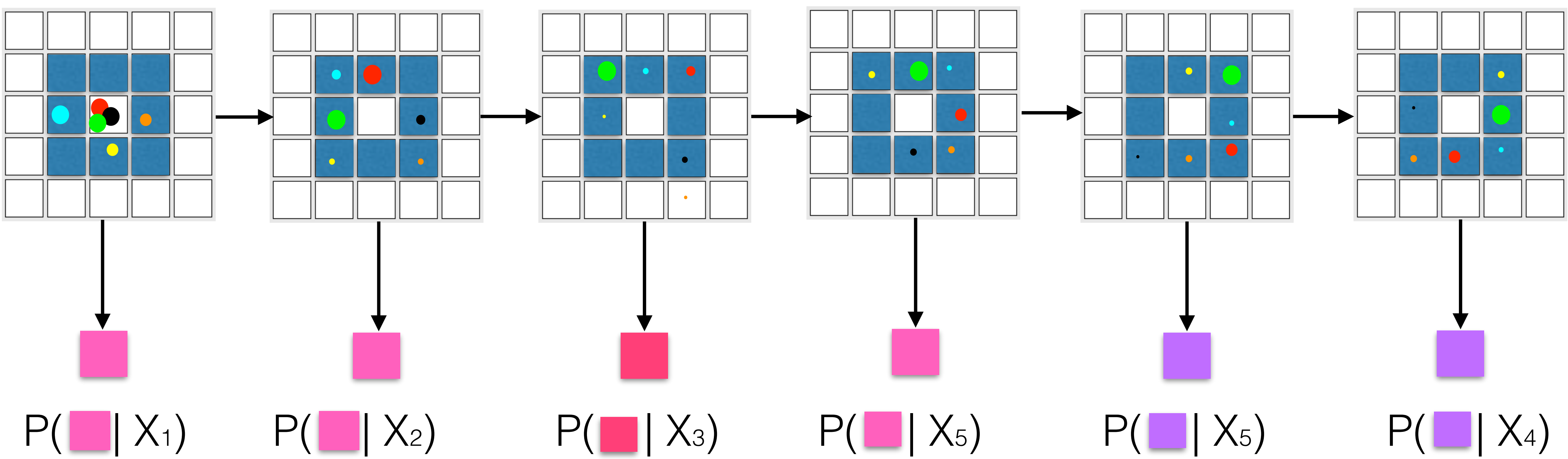
- Use multiple samples and track each ones weights.



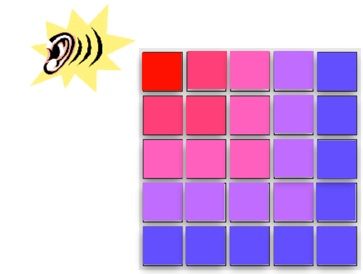
HMM Particle Filter



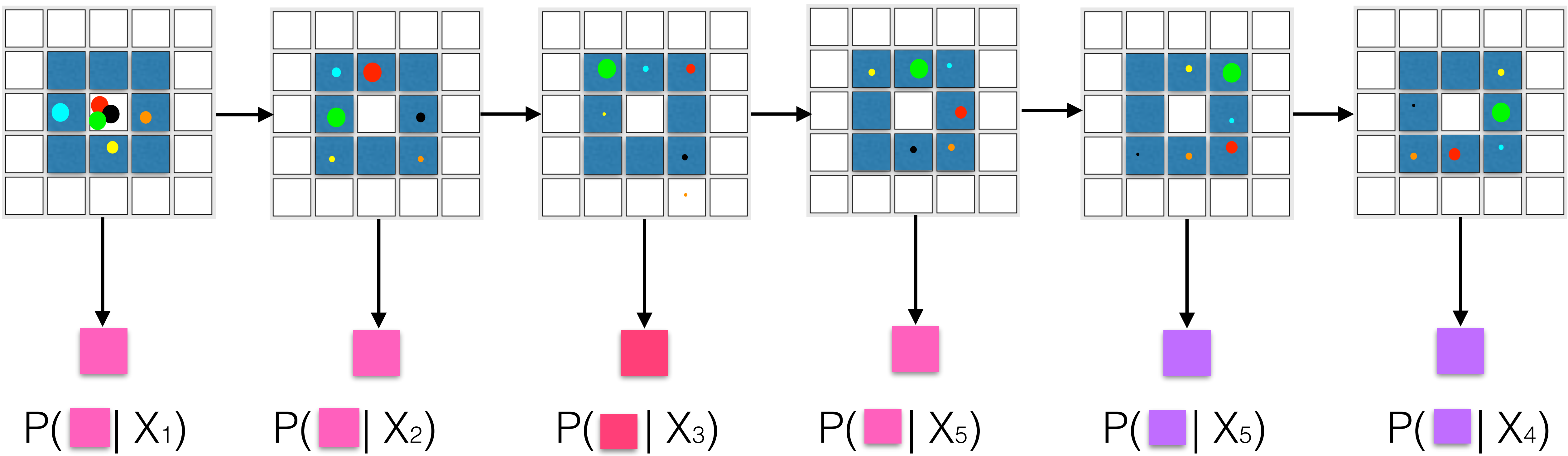
- Use multiple samples and track each ones weights.



HMM Particle Filter

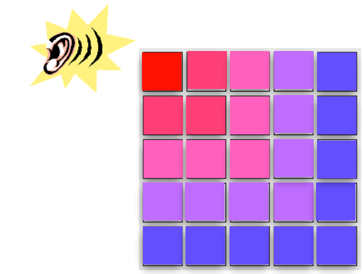


- Use multiple samples and track each ones weights.

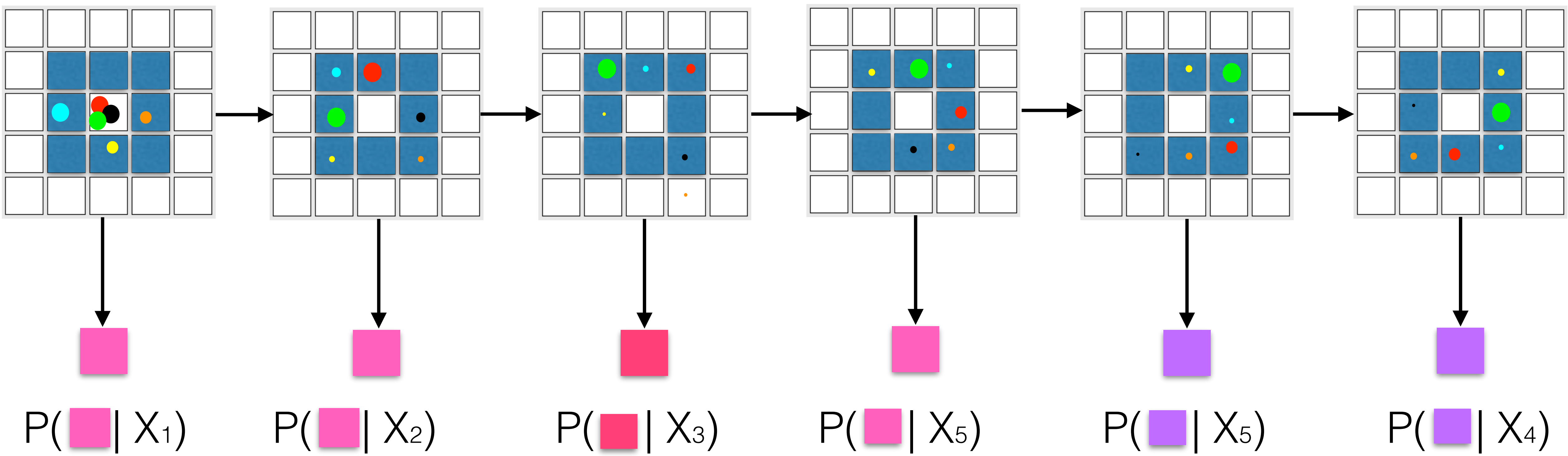


- This is same as 6 separate samples

HMM Particle Filter



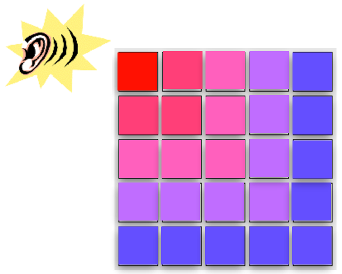
- Use multiple samples and track each ones weights.



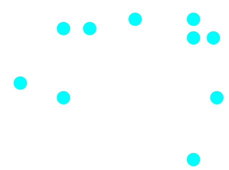
- This is same as 6 separate samples

- Instead of tracking each sample's weight, resample according to weights

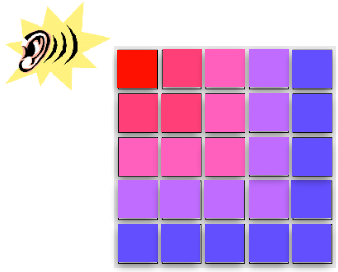
HMM Particle Filter



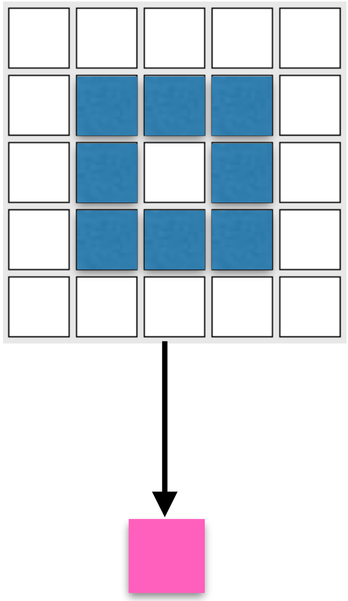
- Use multiple samples and track each ones weights.



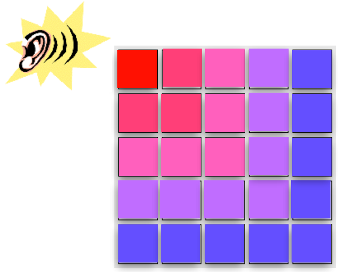
HMM Particle Filter



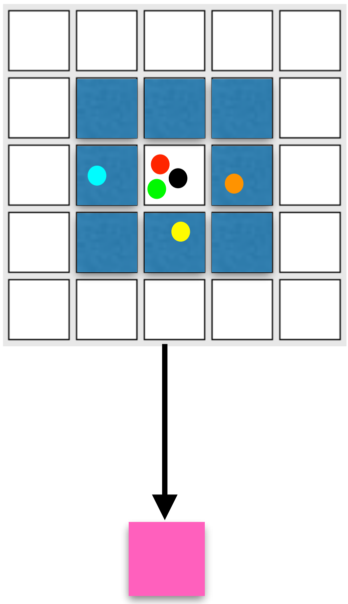
- Use multiple samples and track each ones weights.

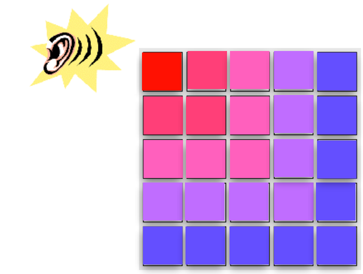


HMM Particle Filter

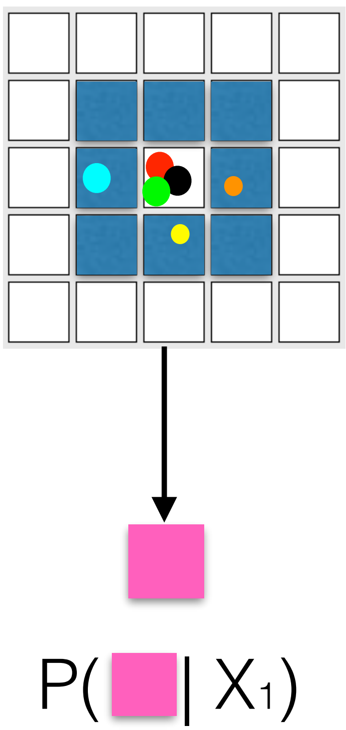


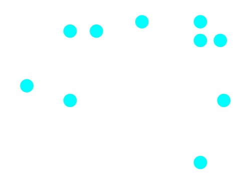
- Use multiple samples and track each ones weights.



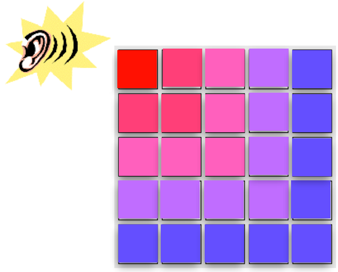


- Use multiple samples and track each ones weights.

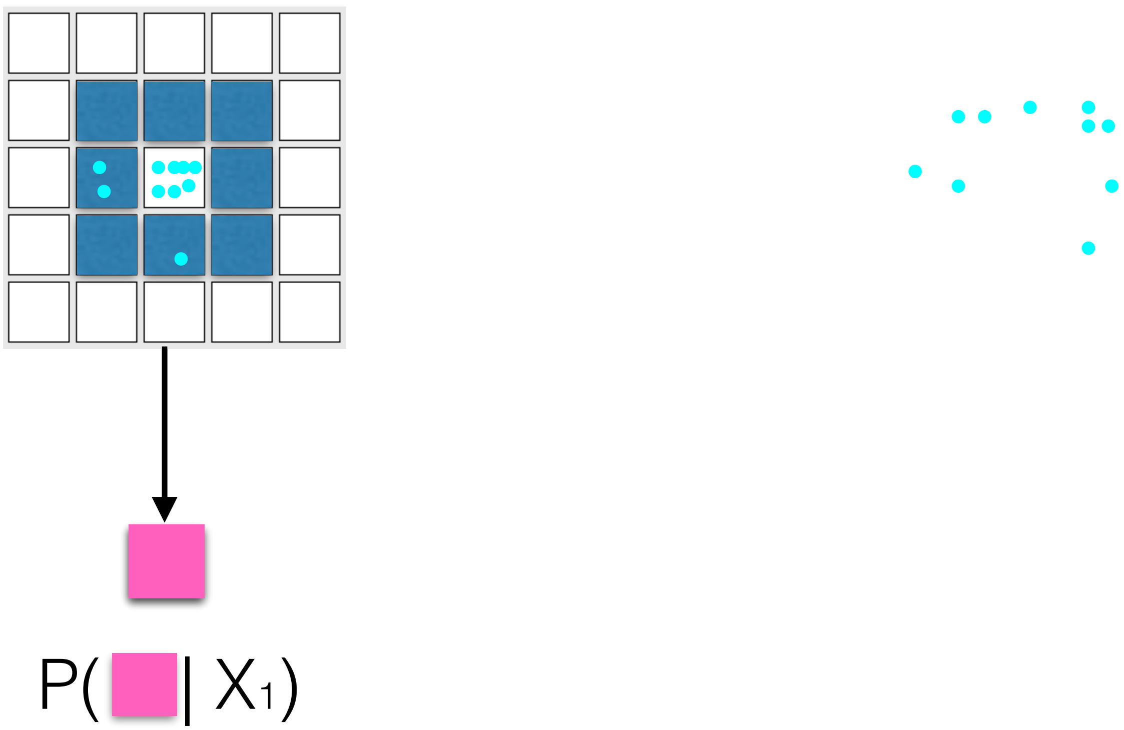




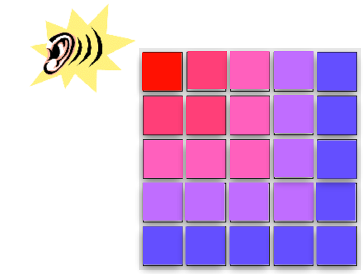
HMM Particle Filter



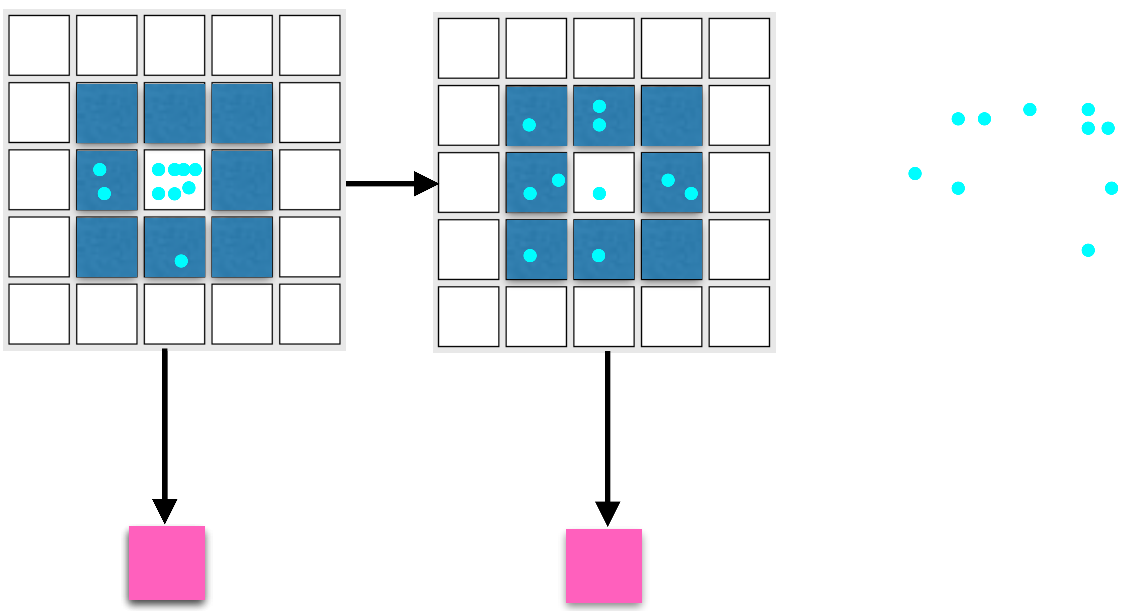
- Use multiple samples and track each ones weights.



HMM Particle Filter

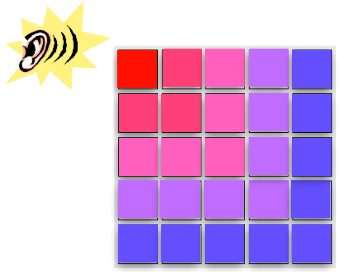


- Use multiple samples and track each ones weights.

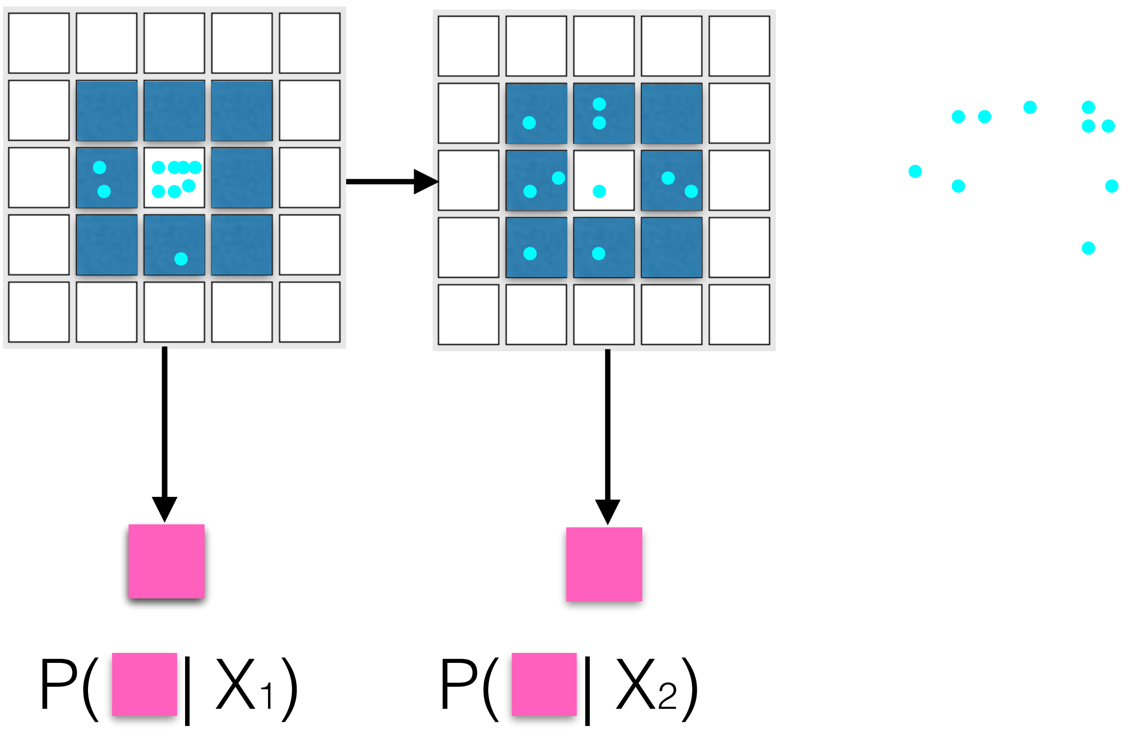


P(| X_1)

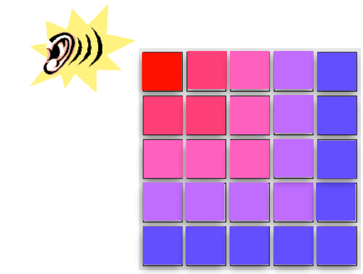
HMM Particle Filter



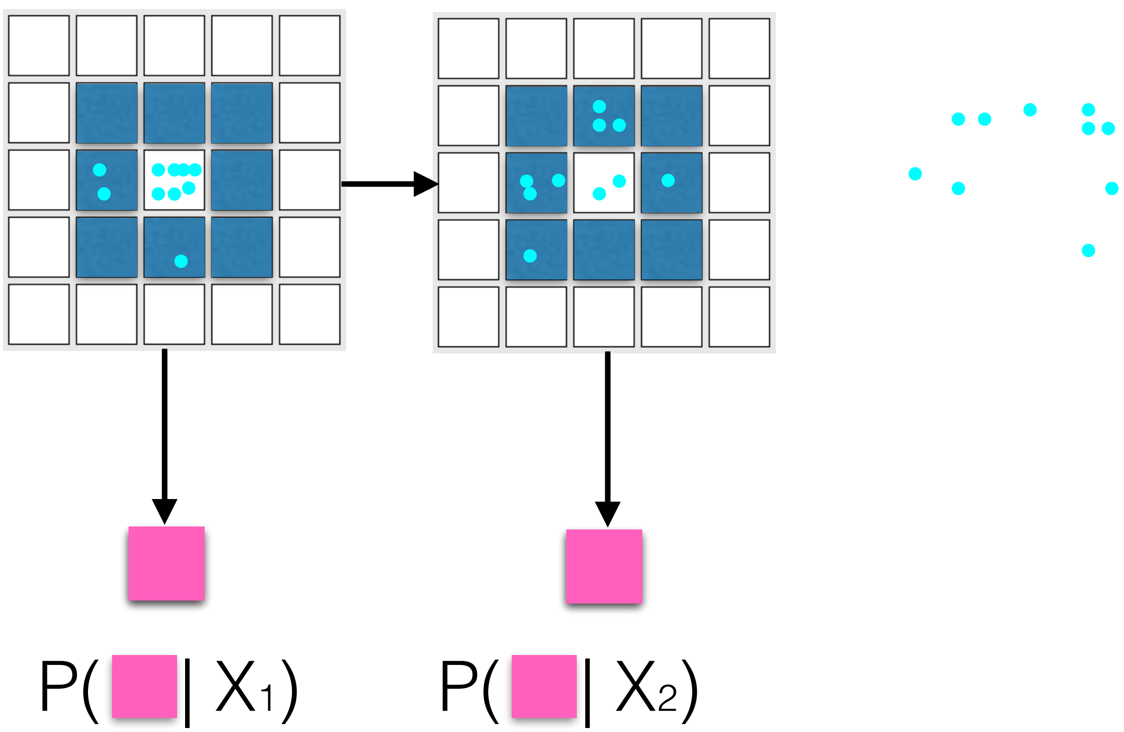
- Use multiple samples and track each ones weights.



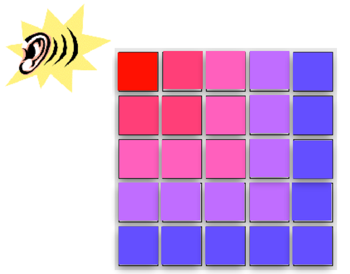
HMM Particle Filter



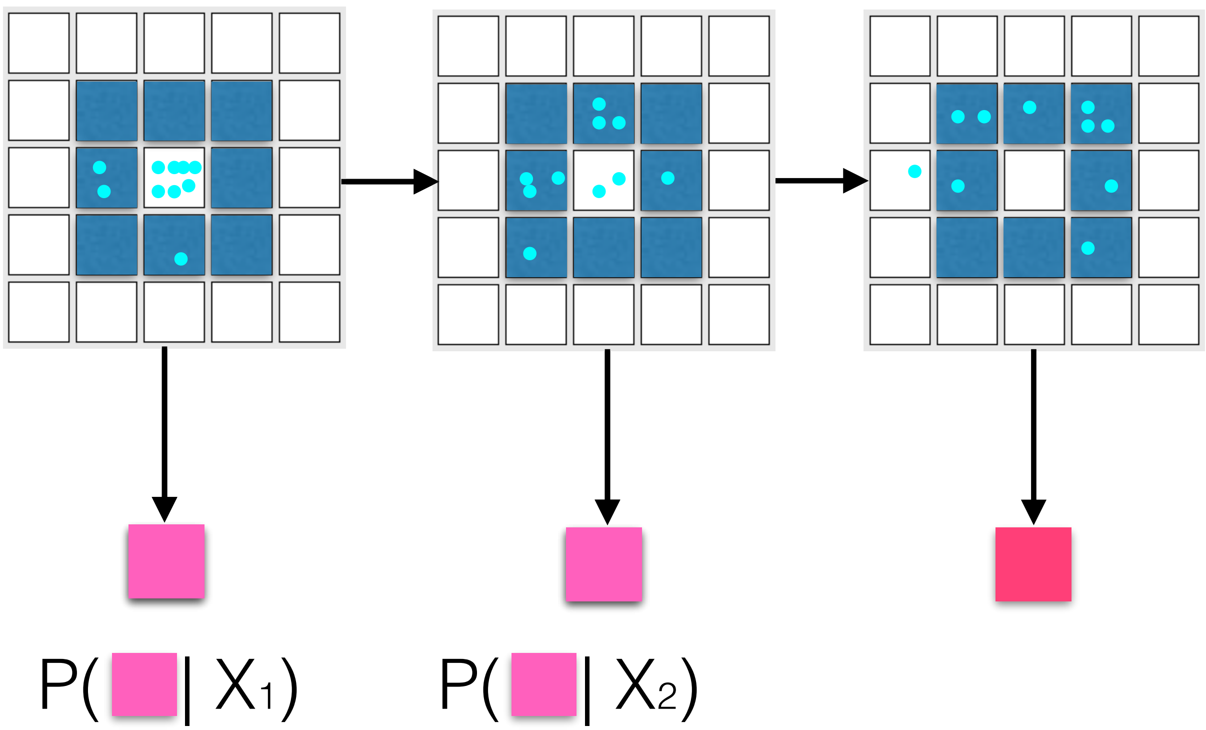
- Use multiple samples and track each ones weights.



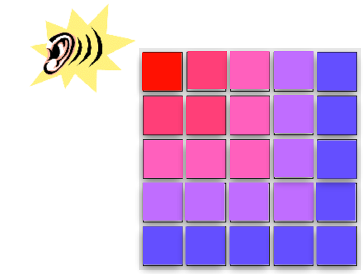
HMM Particle Filter



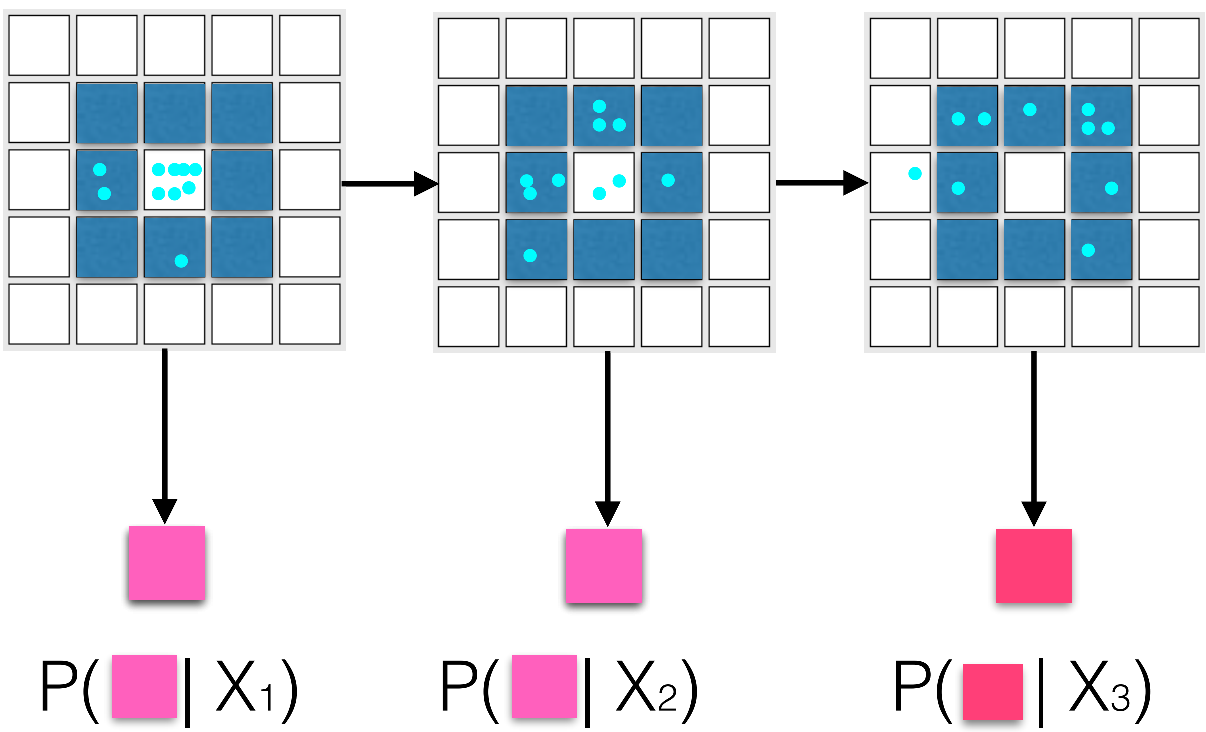
- Use multiple samples and track each ones weights.



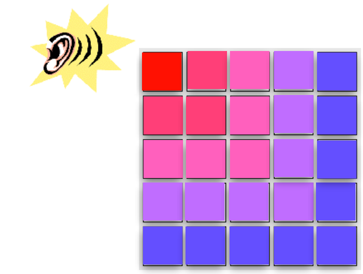
HMM Particle Filter



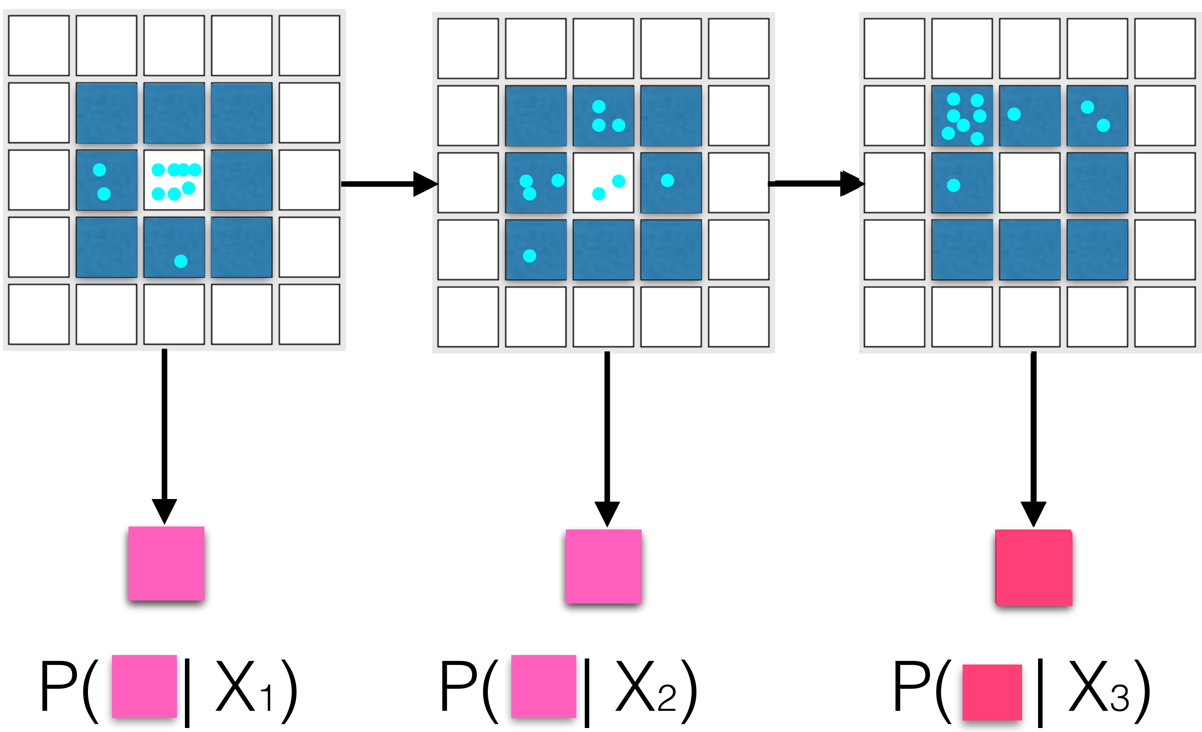
- Use multiple samples and track each ones weights.



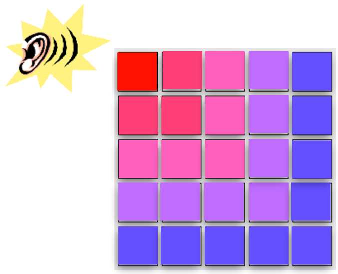
HMM Particle Filter



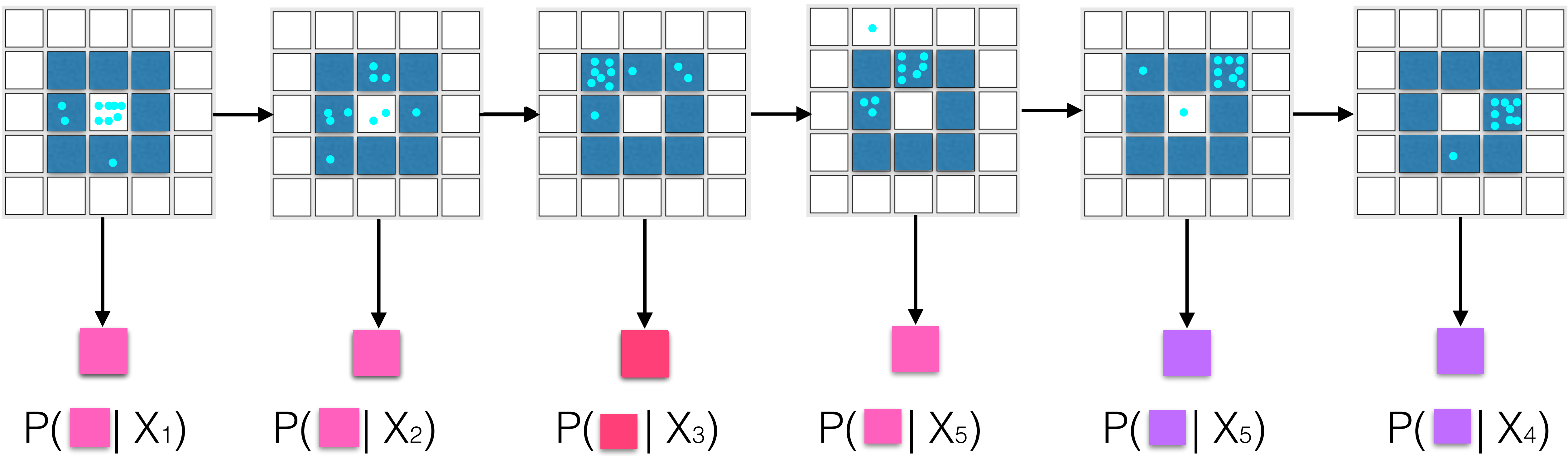
- Use multiple samples and track each ones weights.



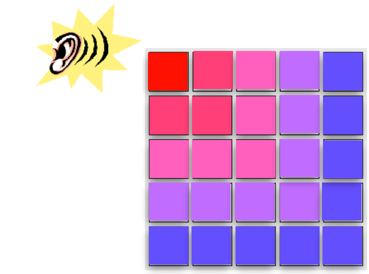
HMM Particle Filter



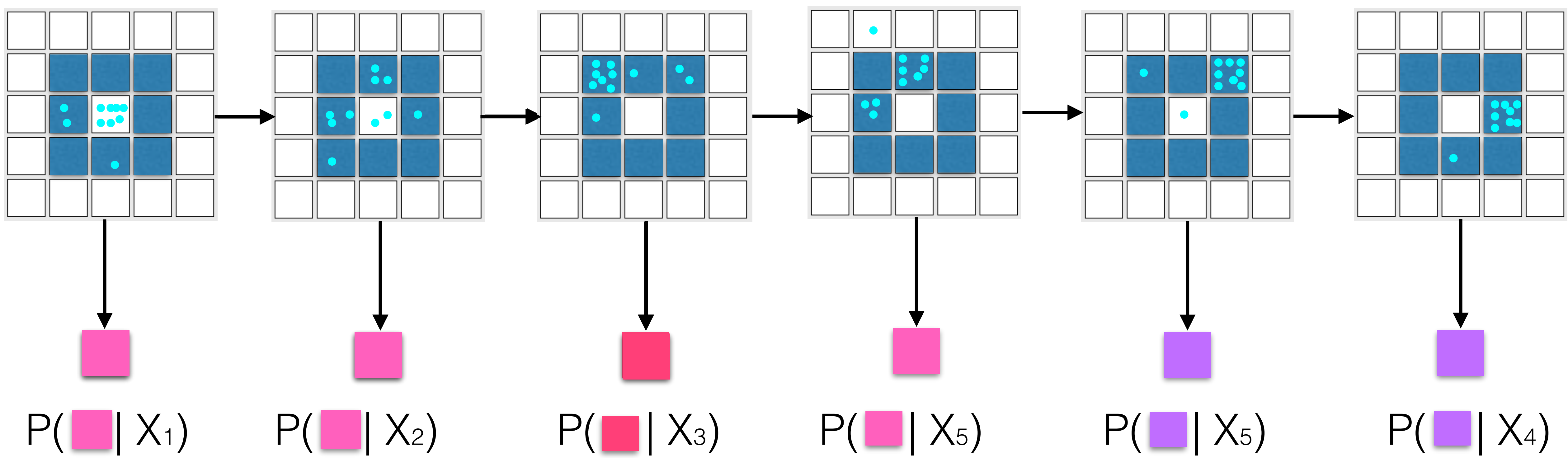
- Use multiple samples and track each ones weights.



HMM Particle Filter

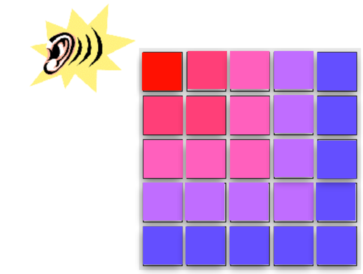


- Use multiple samples and track each ones weights.

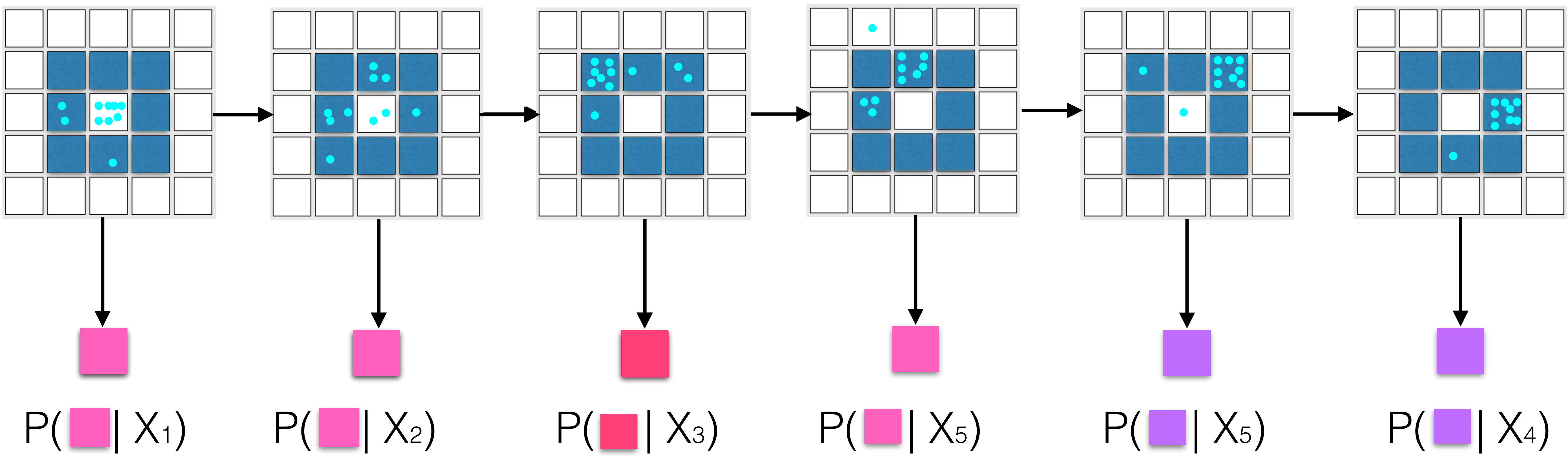


- On every round, transfer particles from previous states according to transition probability

HMM Particle Filter



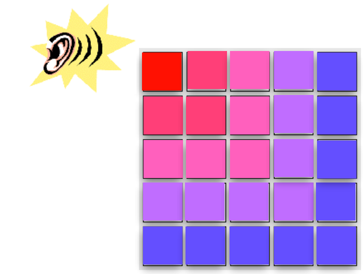
- Use multiple samples and track each ones weights.



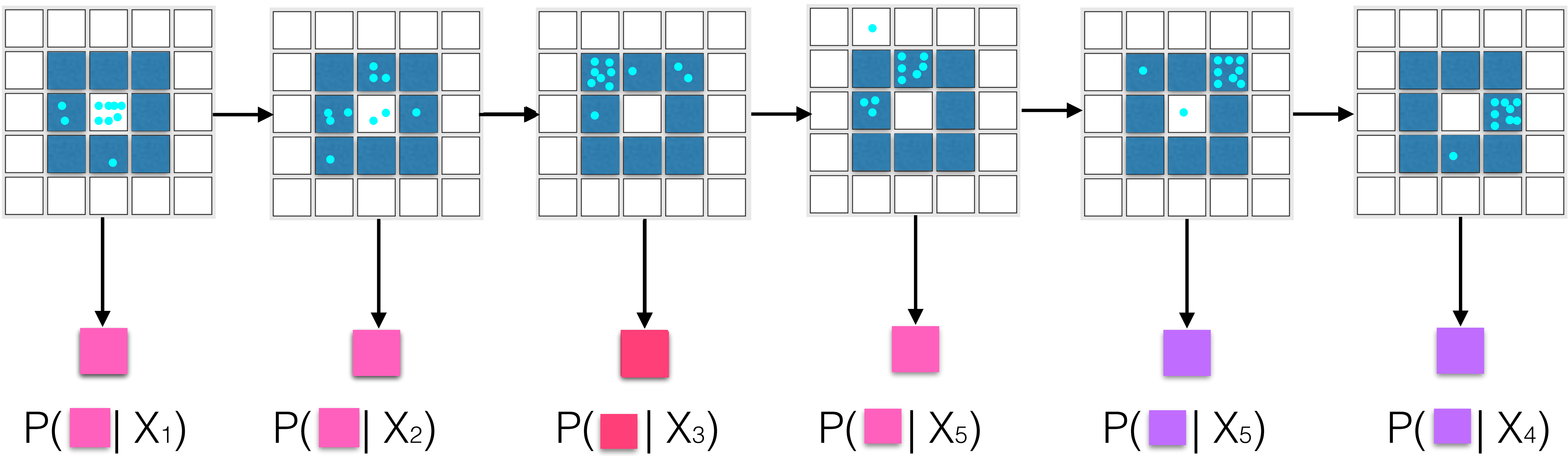
- On every round, transfer particles from previous states according to transition probability

- Resample particles according to P(observation|state)

HMM Particle Filter



- Use multiple samples and track each ones weights.



- On every round, transfer particles from previous states according to transition probability
- Resample particles according to P(observation|state)
- Use new particles to proceed

HMM Particle Filter

- Inference time only depends on number of samples
- Of course more the samples the better accuracy
- Often we don't need too many samples. Why ?

Gibbs Sampling

- Repeat n times for, n samples,
  - Start with arbitrary value for variables
  - Replace each variable by new sample from P(Variable| all other variables)
  - Go over all variables multiple times
  - Return final sample of the N variables

Variational Inference

- Basic idea: we want to infer $P(\text{Unobserved}|\text{Observed})$
  We create a new parametric distribution $Q_\theta(\text{Unobserved})$ where θ is picked based on Obervations
- We pick θ such that, Q_θ is close to $P(\text{Unobserved}|\text{Observed})$
- Closeness measured using KL divergence
- Mean-field approximation,

$$Q_\theta(X_1, \ldots, X_m) = \prod_{j=1}^{m} Q_{\theta_j}(X_j)$$