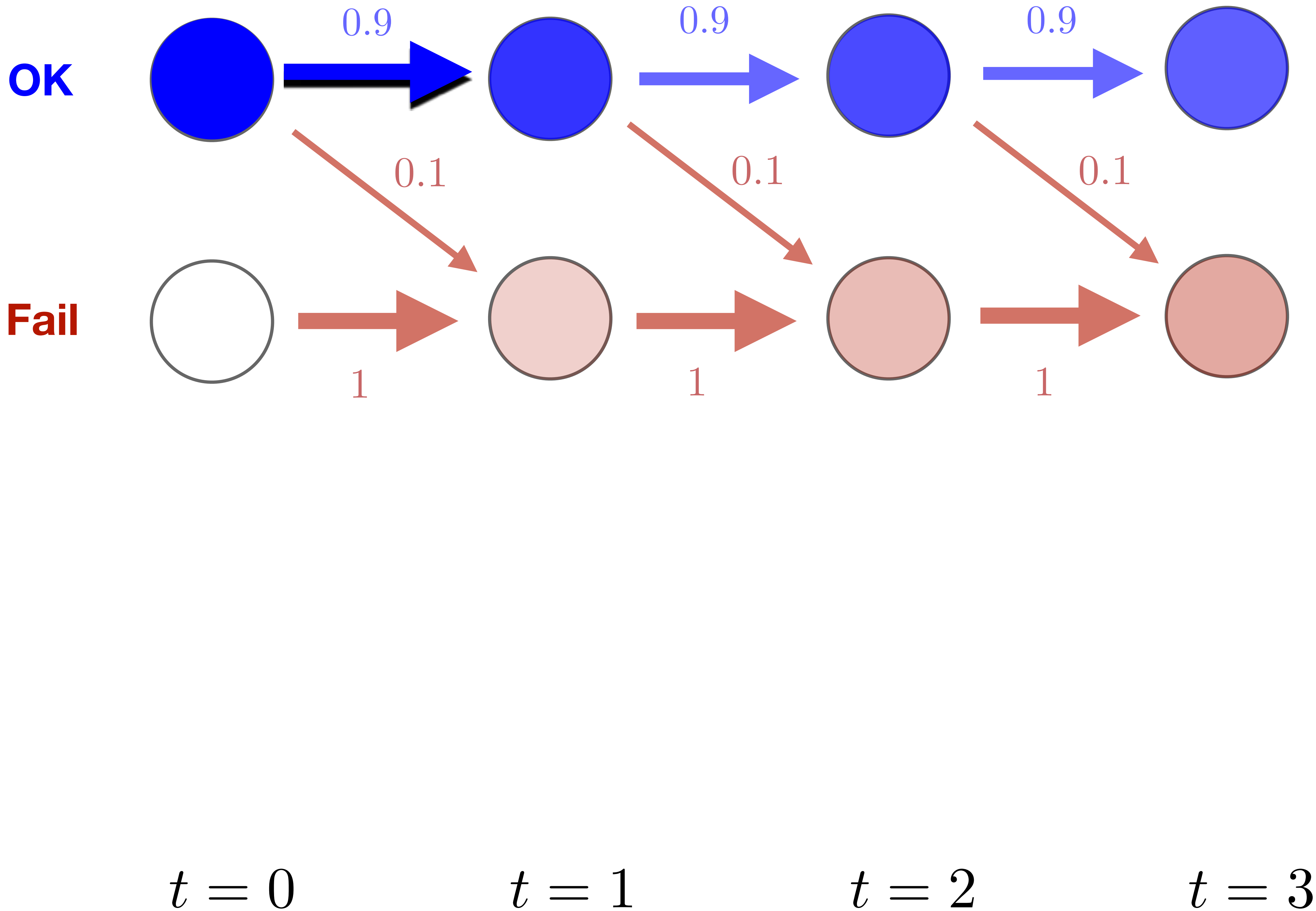
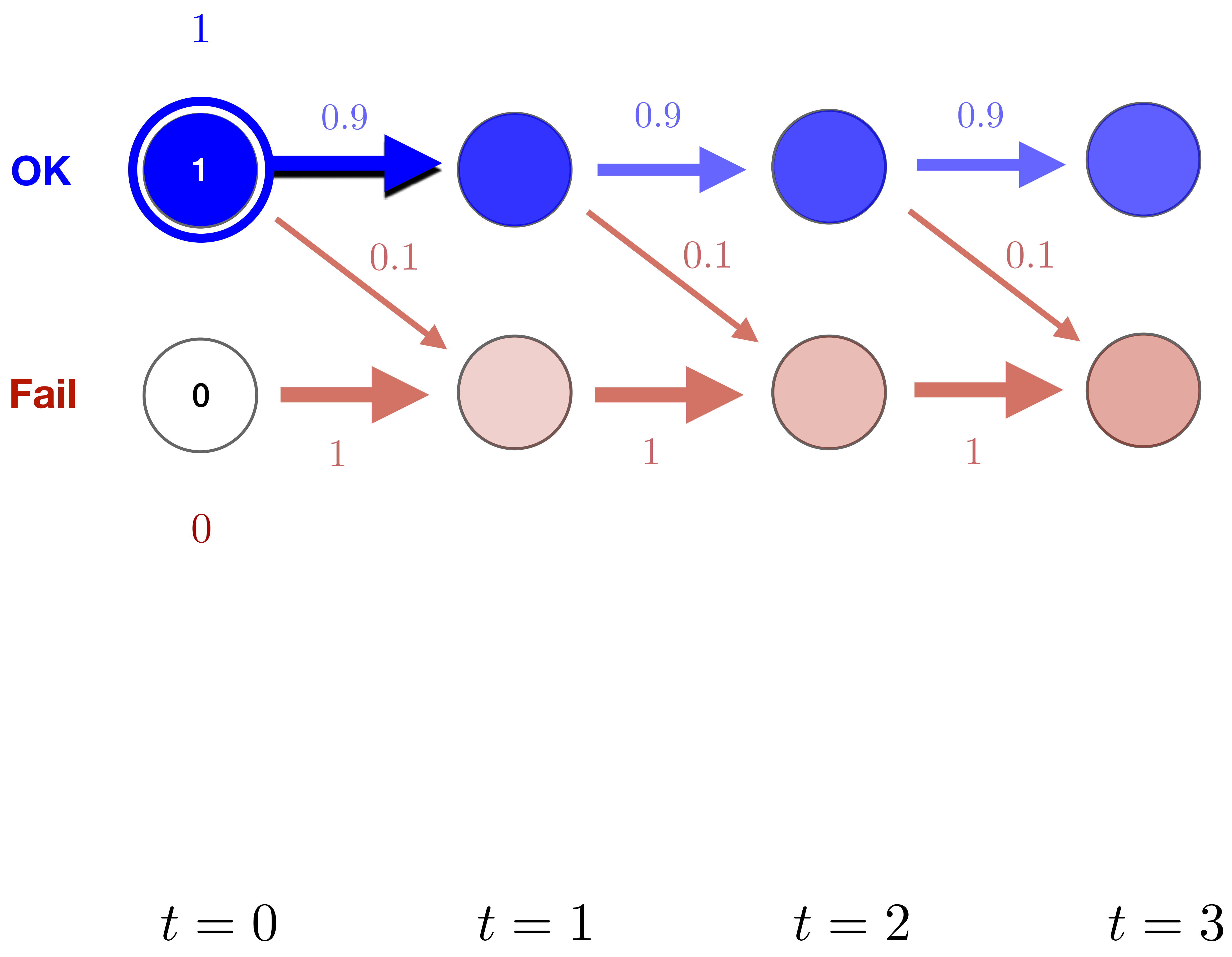
Markov Chain: Examples

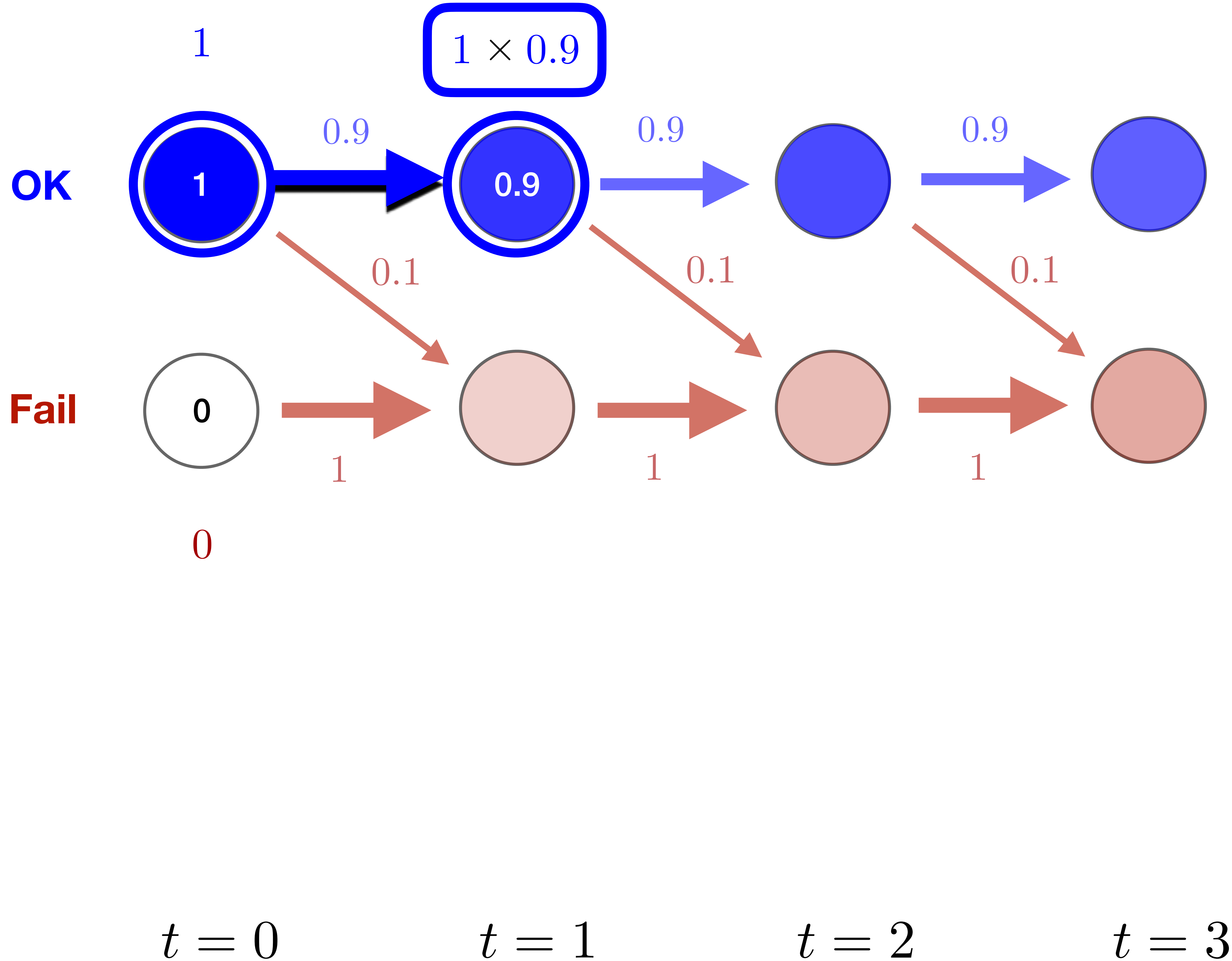
Stochastic Processes

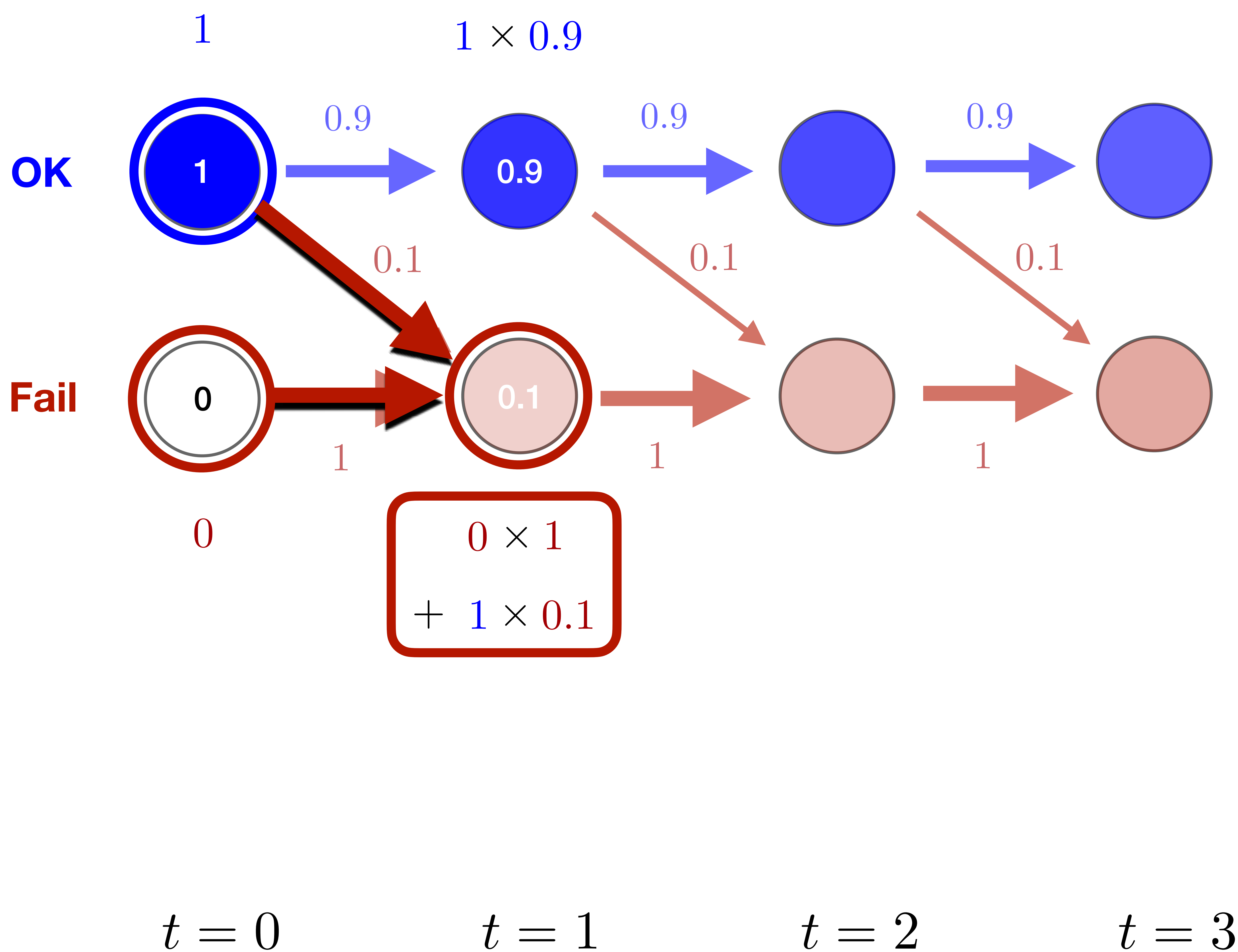
Major sources:

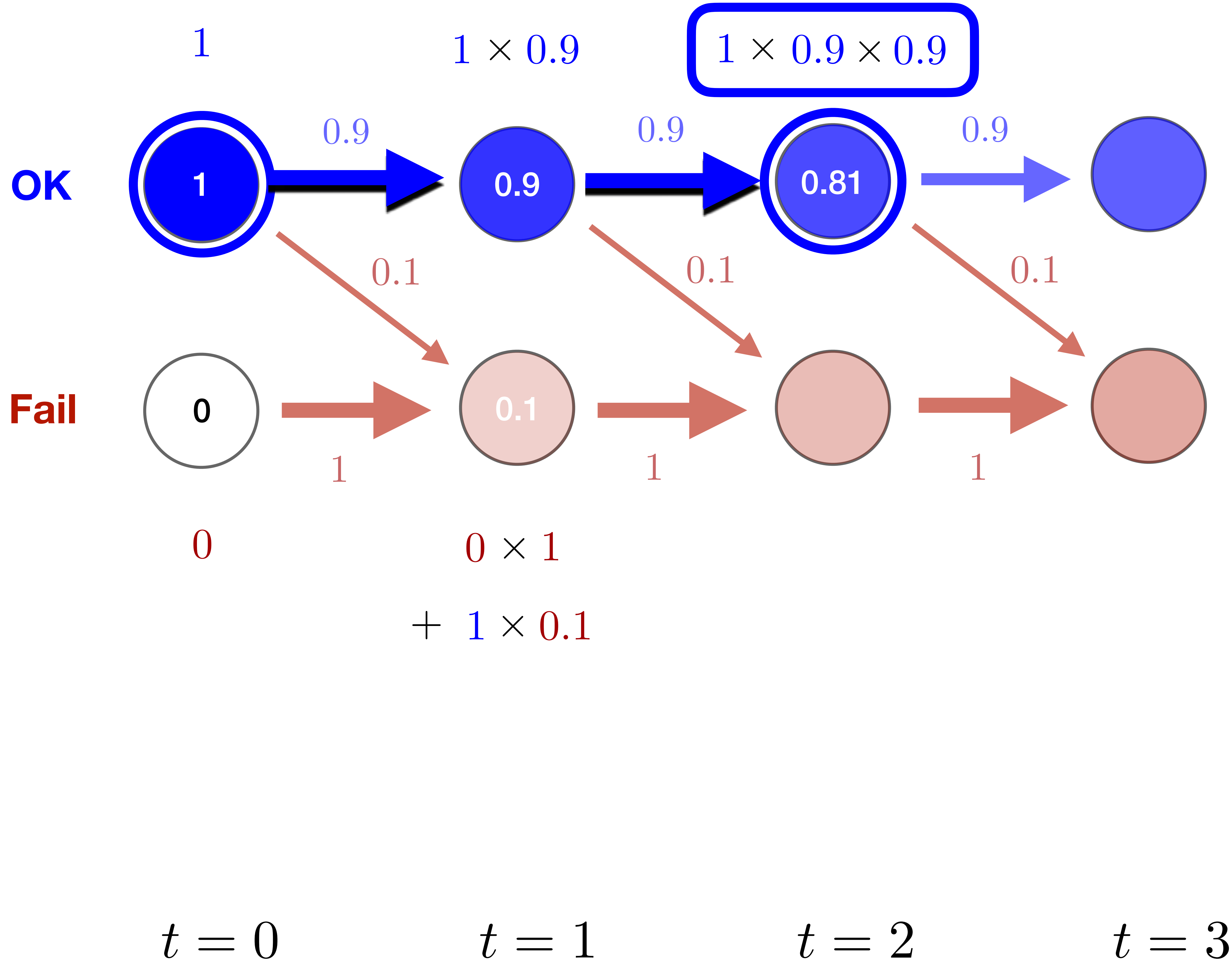
Dan Calderone - Summer 2024


OK
Fail
0.9
0.9
0.9
0.1
0.1
0.1
1
1
1
$t = 0$
$t = 1$
$t = 2$
$t = 3$

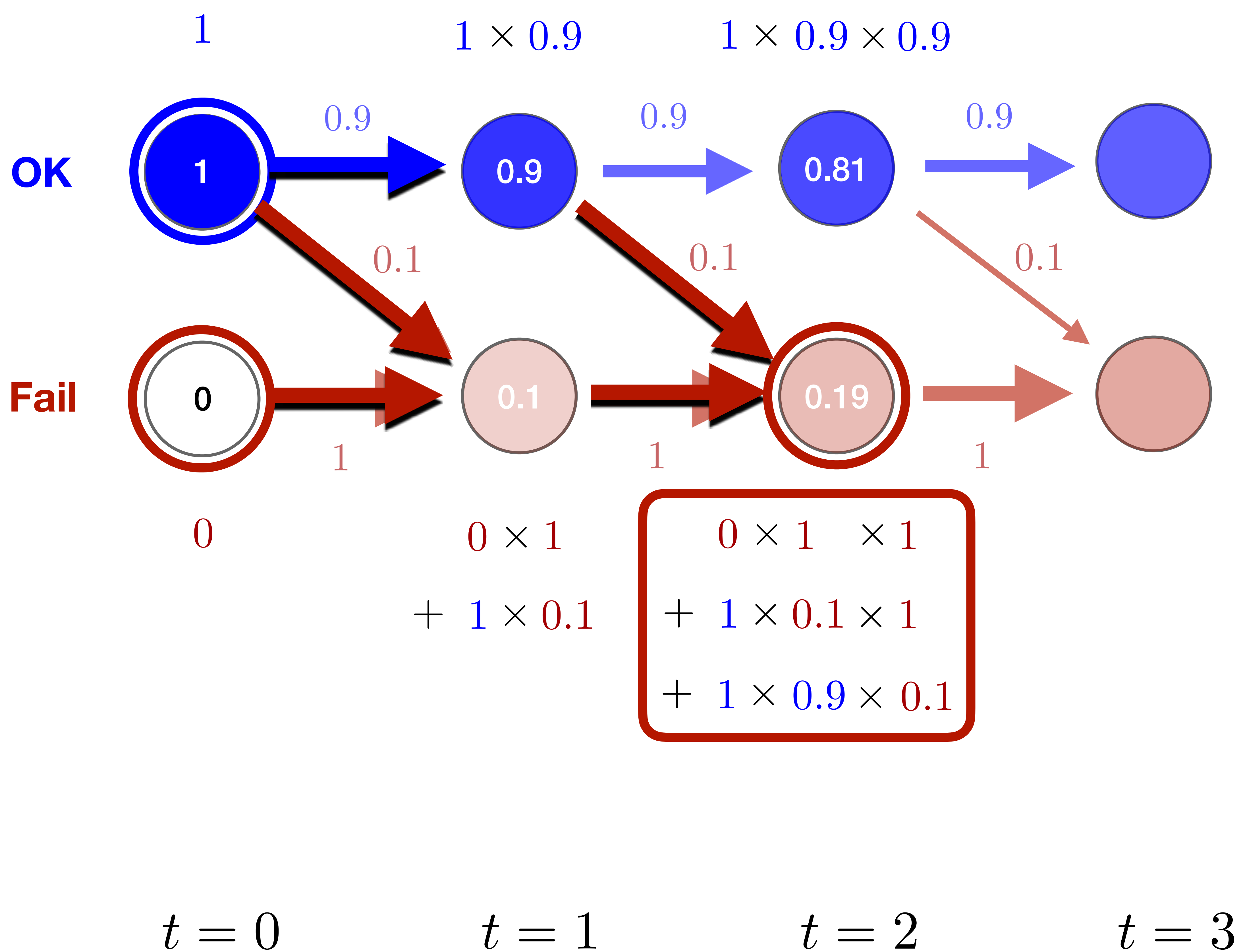

1
OK
1
0.9
0.9
0.9
0.1
0.1
0.1
Fail
0
1
1
1
0
$t = 0$
$t = 1$
$t = 2$
$t = 3$


1
1 × 0.9
0.9
0.9
0.9
OK
1
0.9
0.1
0.1
0.1
Fail
0
1
1
1
0
t = 0
t = 1
t = 2
t = 3

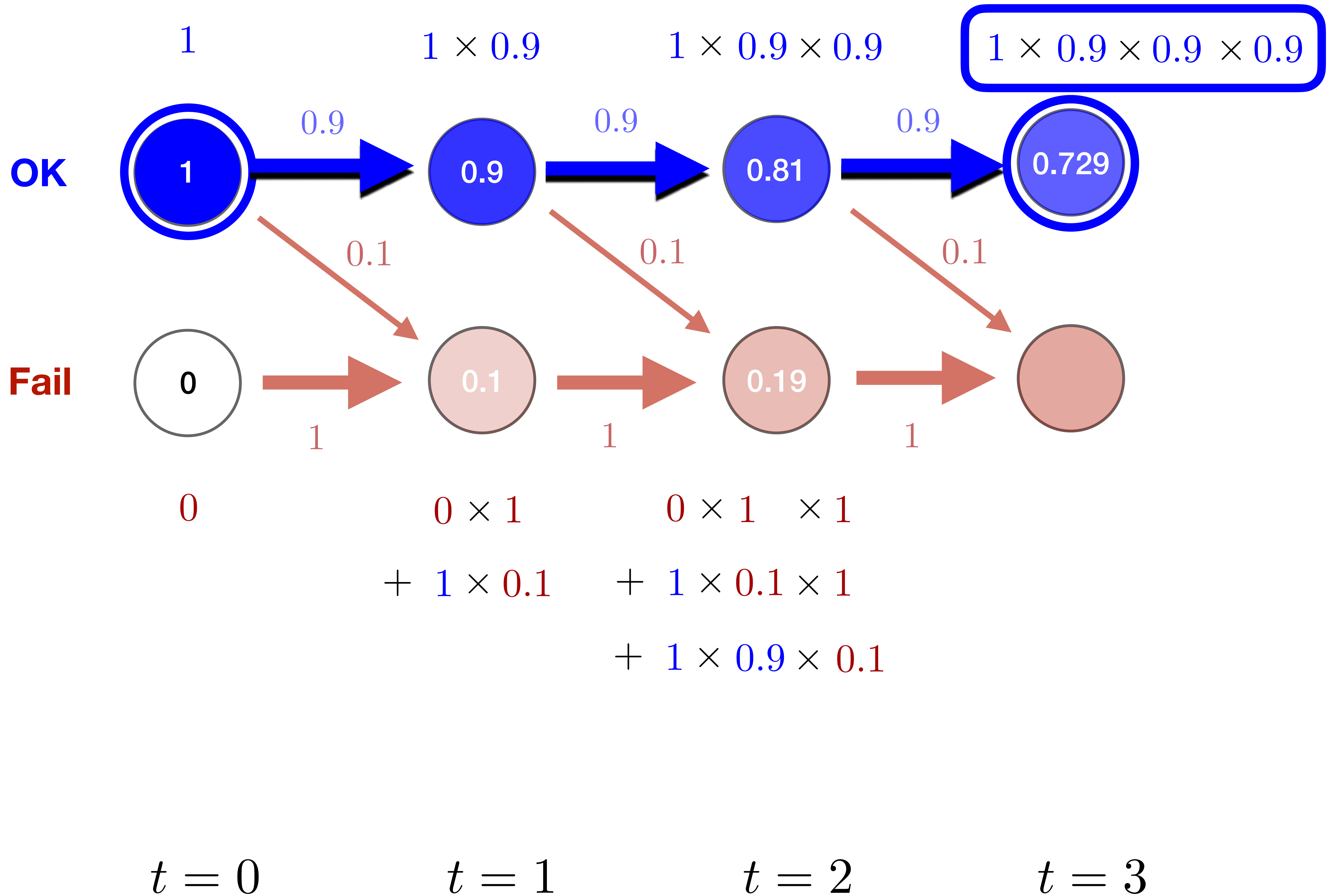

1
1×0.9
OK
Fail
1
0.9
0
0.1
0.9
0.9
0.9
0.1
0.1
0.1
1
1
1
0
0×1
$+ \; 1 \times 0.1$
$t = 0$
$t = 1$
$t = 2$
$t = 3$

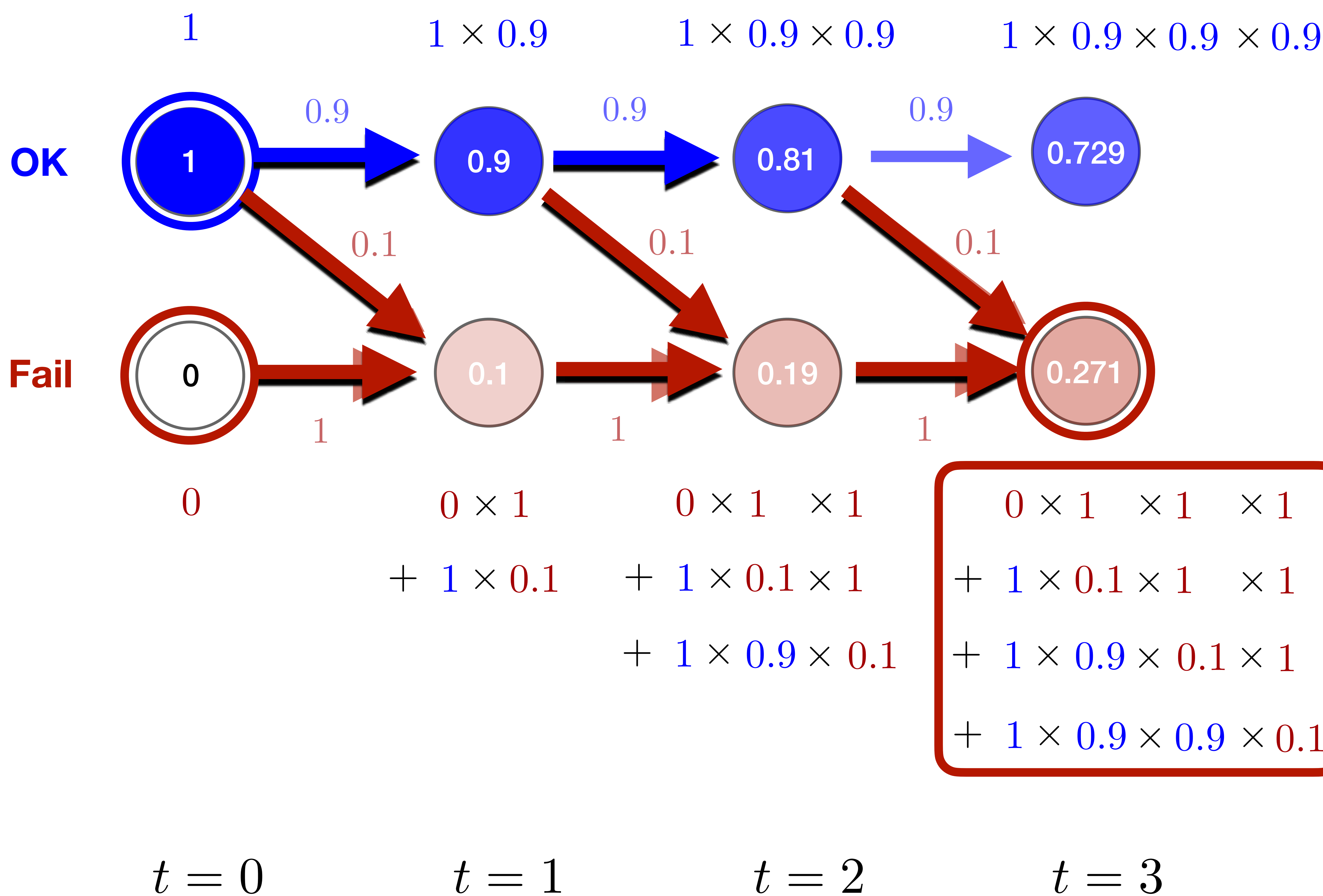


OK
Fail
1
1 × 0.9
1 × 0.9 × 0.9
0.9
0.9
0.9
1
0.9
0.81
0.1
0.1
0.1
0
0.1
1
1
1
0
0 × 1
+ 1 × 0.1
t = 0
t = 1
t = 2
t = 3

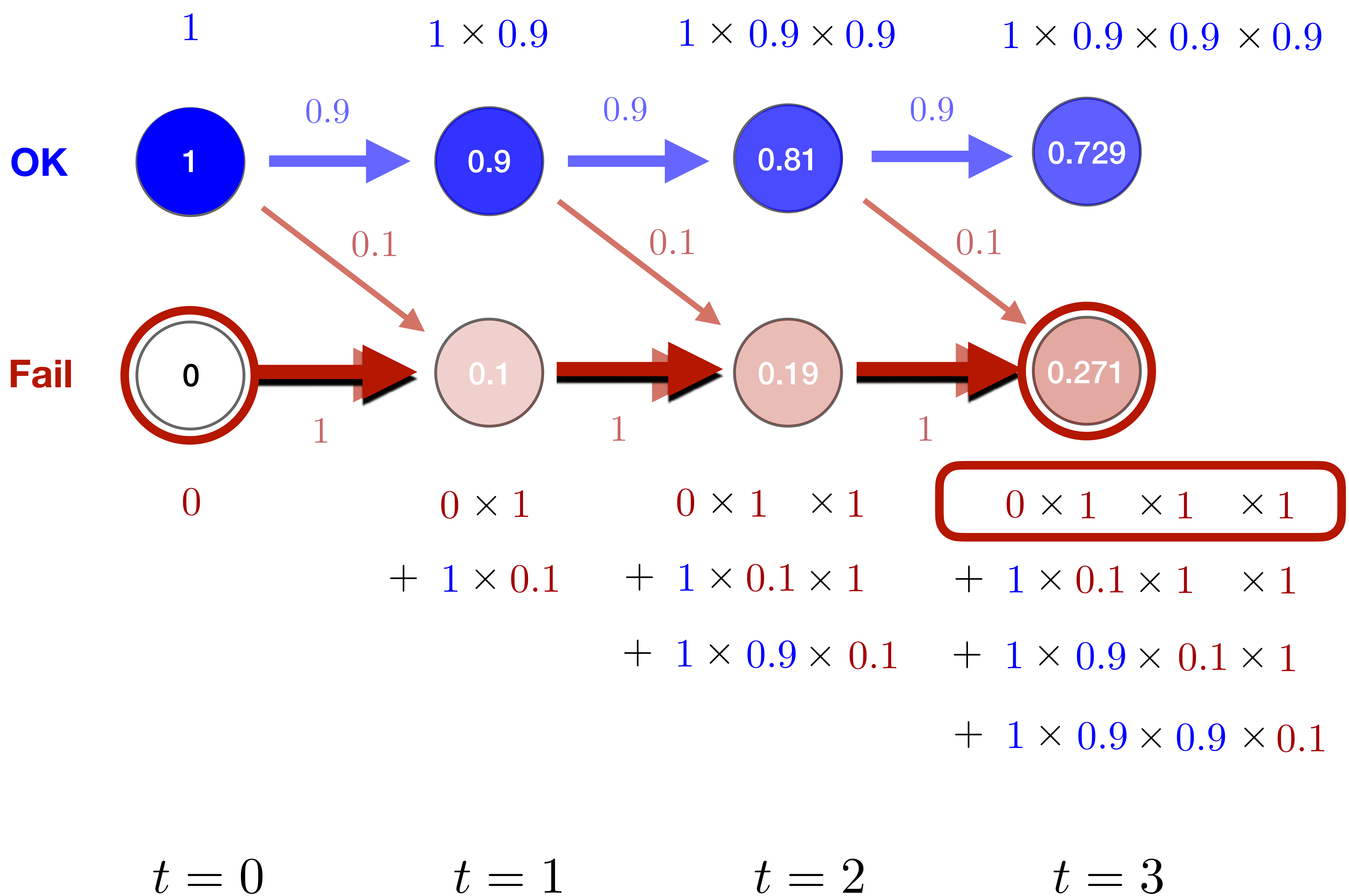


OK
Fail
1
1 × 0.9
1 × 0.9 × 0.9
0.9
0.9
0.9
0.1
0.1
0.1
1
1
1
1
0.9
0.81
0
0.1
0.19
0
0 × 1
+ 1 × 0.1
0 × 1 × 1
+ 1 × 0.1 × 1
+ 1 × 0.9 × 0.1
t = 0
t = 1
t = 2
t = 3

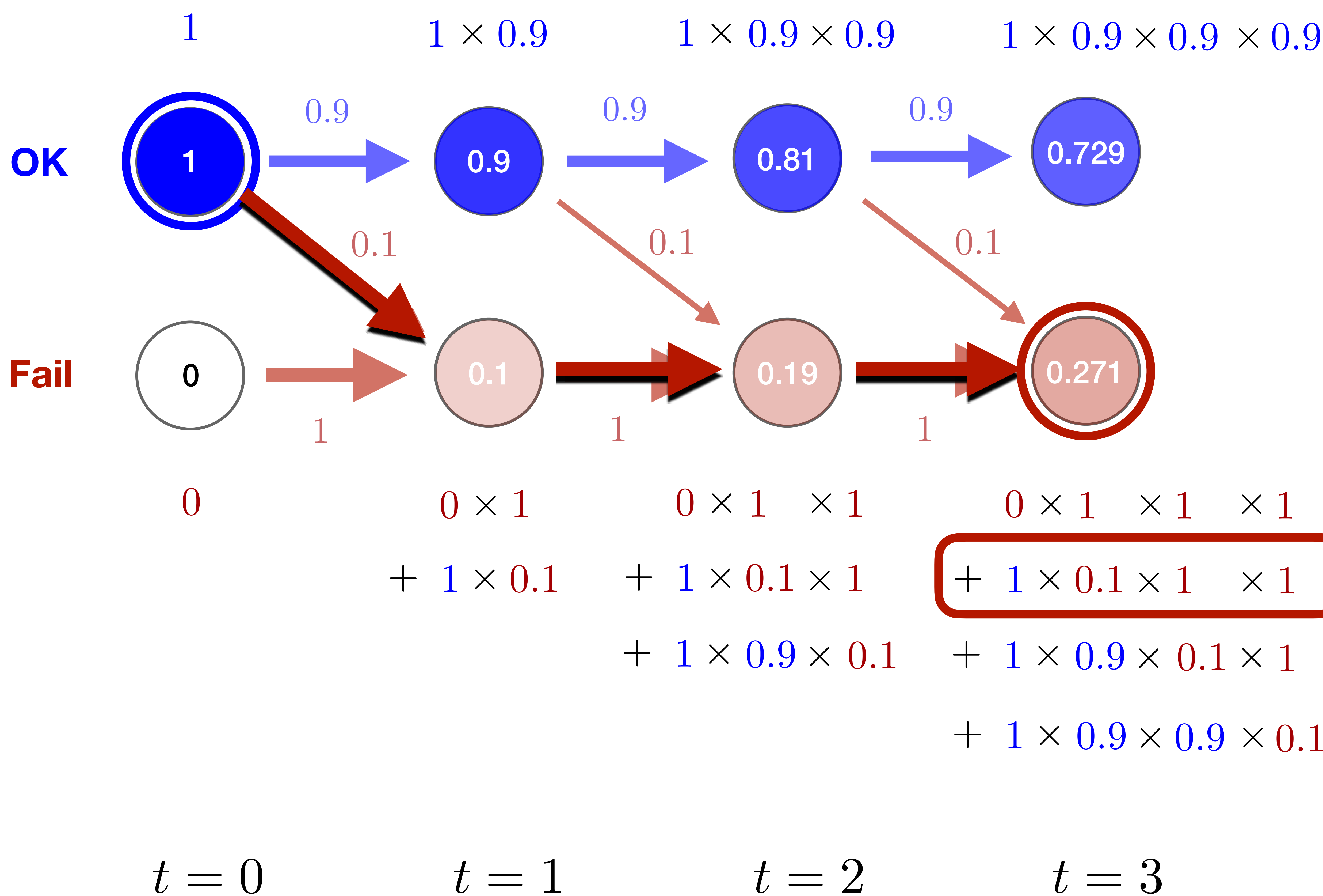

1
1 × 0.9
1 × 0.9 × 0.9
1 × 0.9 × 0.9 × 0.9
OK
1
0.9
0.9
0.9
0.81
0.9
0.729
0.1
0.1
0.1
Fail
0
0.1
0.19
1
1
1
0
0 × 1
0 × 1 × 1
+ 1 × 0.1
+ 1 × 0.1 × 1
+ 1 × 0.9 × 0.1
t = 0
t = 1
t = 2
t = 3



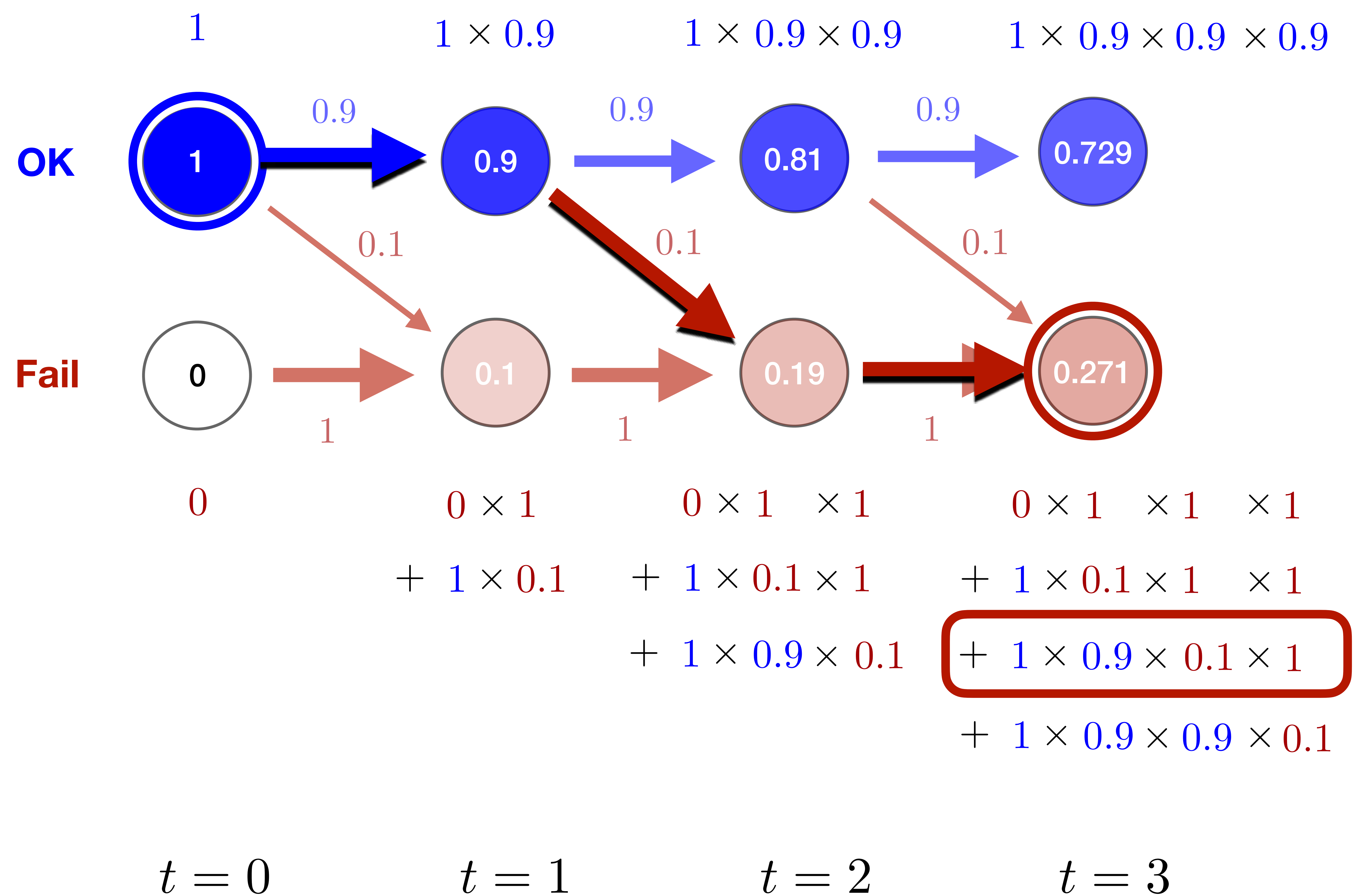
1
1 × 0.9
1 × 0.9 × 0.9
1 × 0.9 × 0.9 × 0.9
OK
Fail
0.9
0.9
0.9
0.1
0.1
0.1
1
1
1
1
0.9
0.81
0.729
0
0.1
0.19
0.271
0
0 × 1
+ 1 × 0.1
0 × 1 × 1
+ 1 × 0.1 × 1
+ 1 × 0.9 × 0.1
0 × 1 × 1 × 1
+ 1 × 0.1 × 1 × 1
+ 1 × 0.9 × 0.1 × 1
+ 1 × 0.9 × 0.9 × 0.1
t = 0
t = 1
t = 2
t = 3



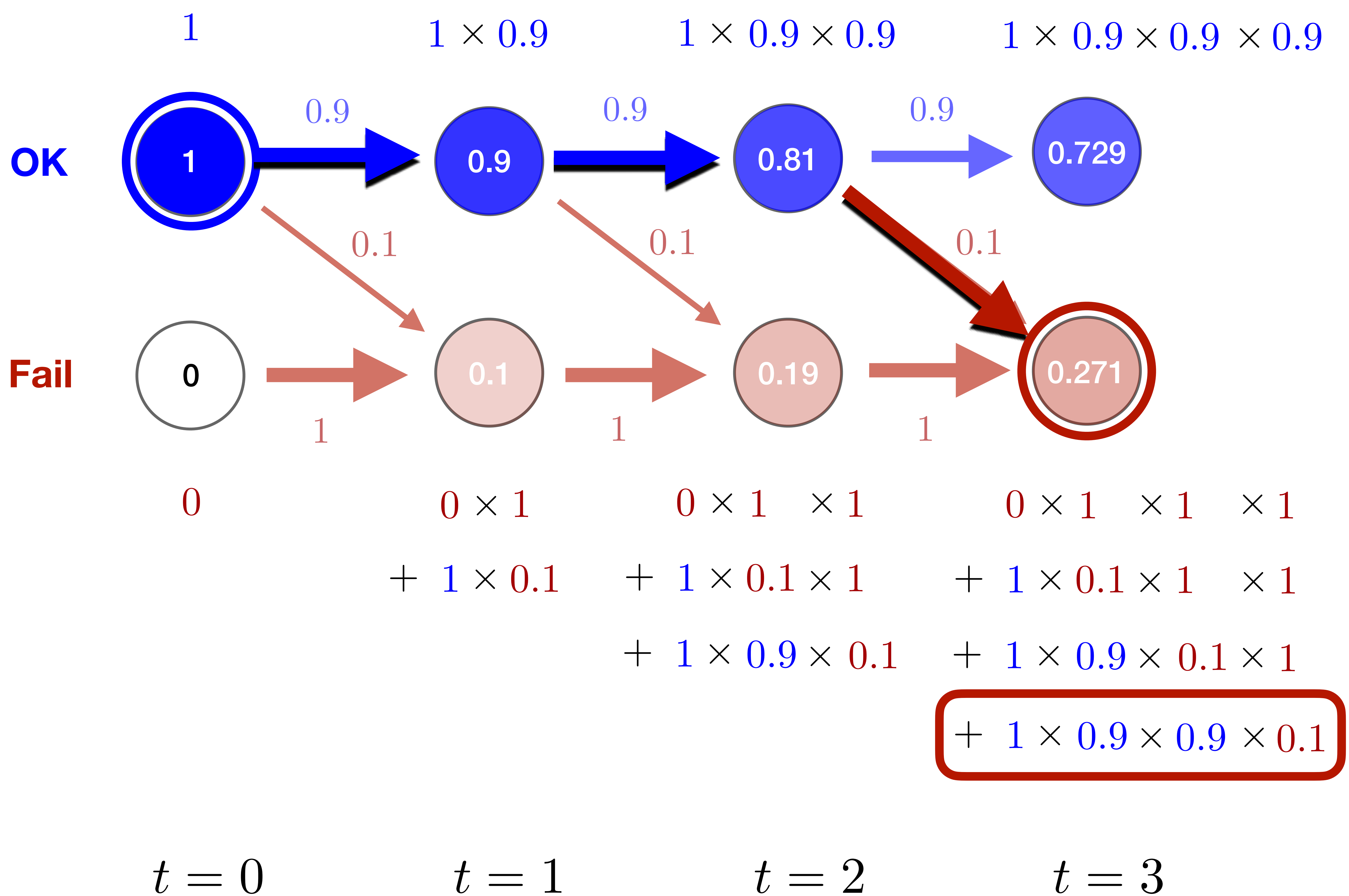
OK
Fail
1
1 × 0.9
1 × 0.9 × 0.9
1 × 0.9 × 0.9 × 0.9
0.9
0.9
0.9
0.1
0.1
0.1
1
0.9
0.81
0.729
0
0.1
0.19
0.271
1
1
1
0
0 × 1
+ 1 × 0.1
0 × 1 × 1
+ 1 × 0.1 × 1
+ 1 × 0.9 × 0.1
0 × 1 × 1 × 1
+ 1 × 0.1 × 1 × 1
+ 1 × 0.9 × 0.1 × 1
+ 1 × 0.9 × 0.9 × 0.1
t = 0
t = 1
t = 2
t = 3



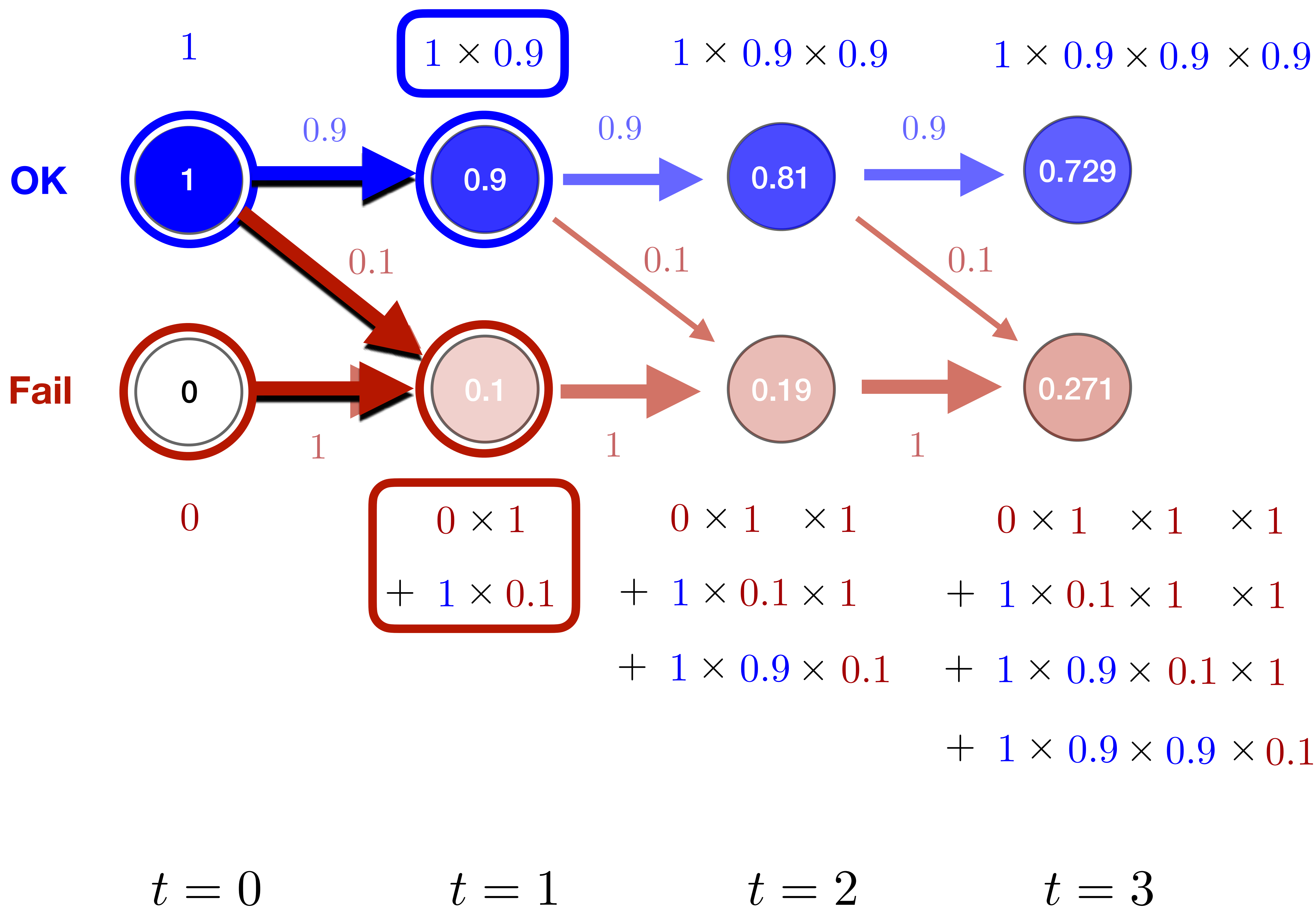
OK
Fail
1
1 × 0.9
1 × 0.9 × 0.9
1 × 0.9 × 0.9 × 0.9
0.9
0.9
0.9
0.1
0.1
0.1
1
1
1
1
0.9
0.81
0.729
0
0.1
0.19
0.271
0
0 × 1
+ 1 × 0.1
0 × 1 × 1
+ 1 × 0.1 × 1
+ 1 × 0.9 × 0.1
0 × 1 × 1 × 1
+ 1 × 0.1 × 1 × 1
+ 1 × 0.9 × 0.1 × 1
+ 1 × 0.9 × 0.9 × 0.1
t = 0
t = 1
t = 2
t = 3



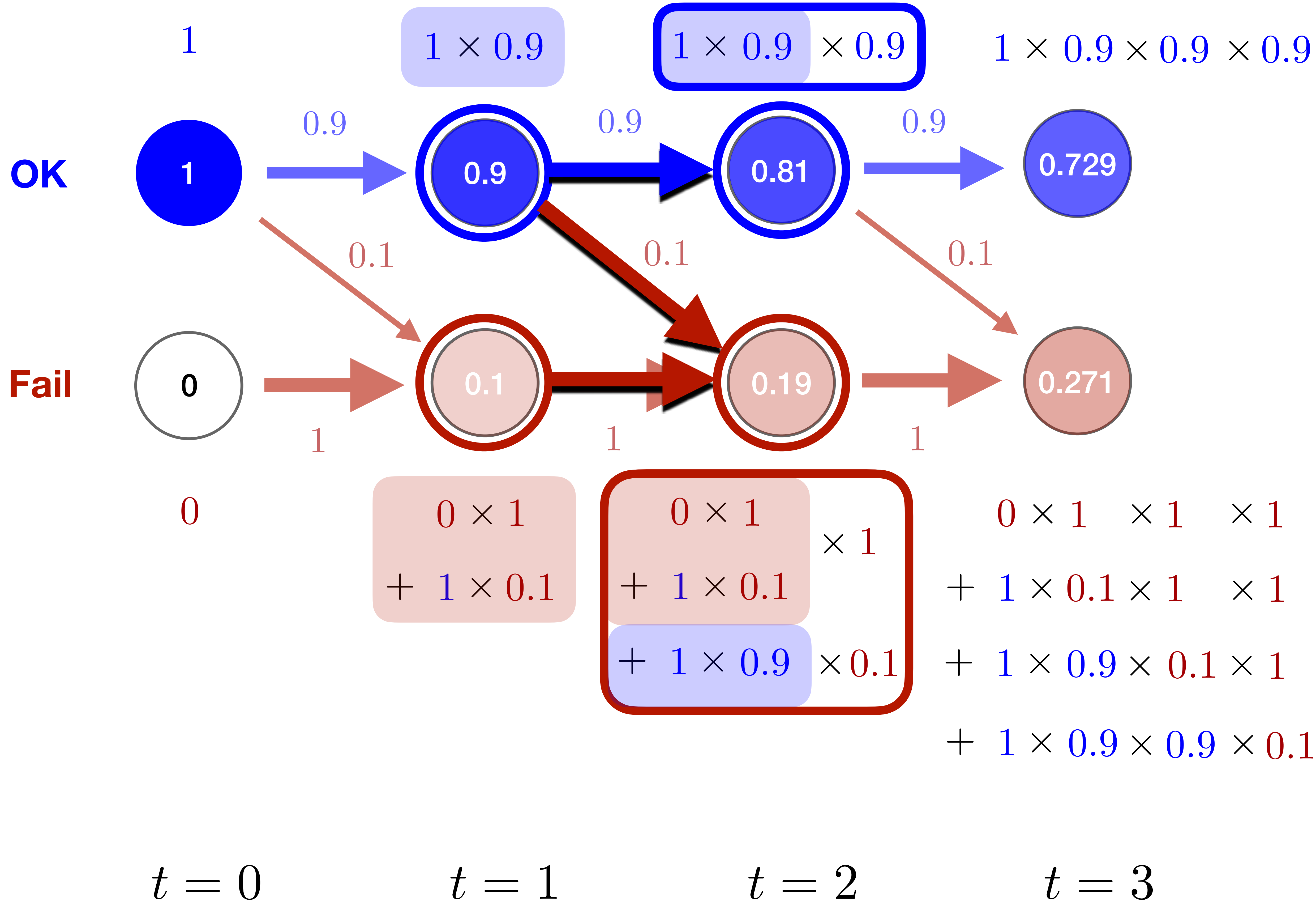
OK
Fail
1
1 × 0.9
1 × 0.9 × 0.9
1 × 0.9 × 0.9 × 0.9
1
0.9
0.81
0.729
0
0.1
0.19
0.271
0.9
0.9
0.9
0.1
0.1
0.1
1
1
1
0
0 × 1
+ 1 × 0.1
0 × 1 × 1
+ 1 × 0.1 × 1
+ 1 × 0.9 × 0.1
0 × 1 × 1 × 1
+ 1 × 0.1 × 1 × 1
+ 1 × 0.9 × 0.1 × 1
+ 1 × 0.9 × 0.9 × 0.1
t = 0
t = 1
t = 2
t = 3

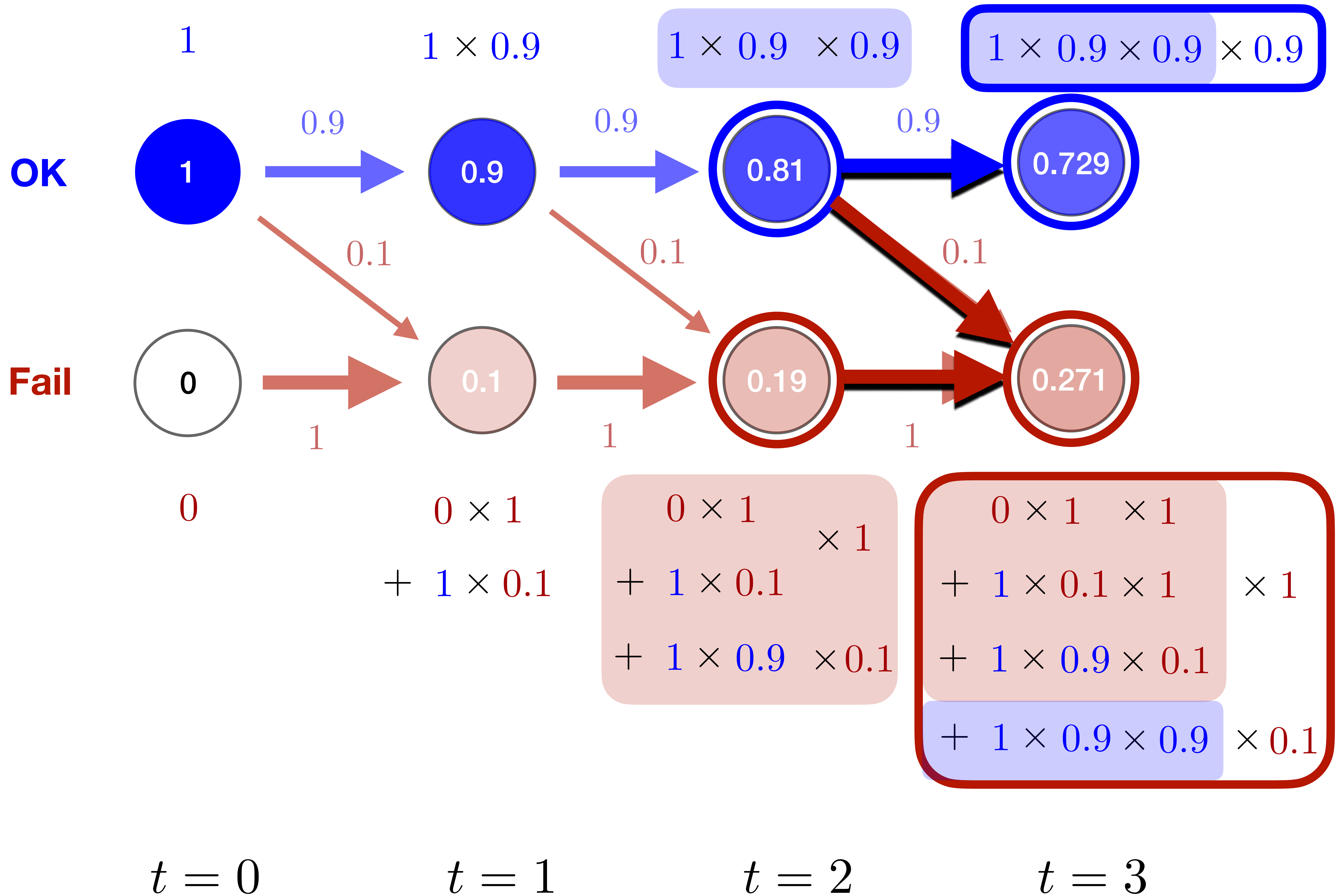


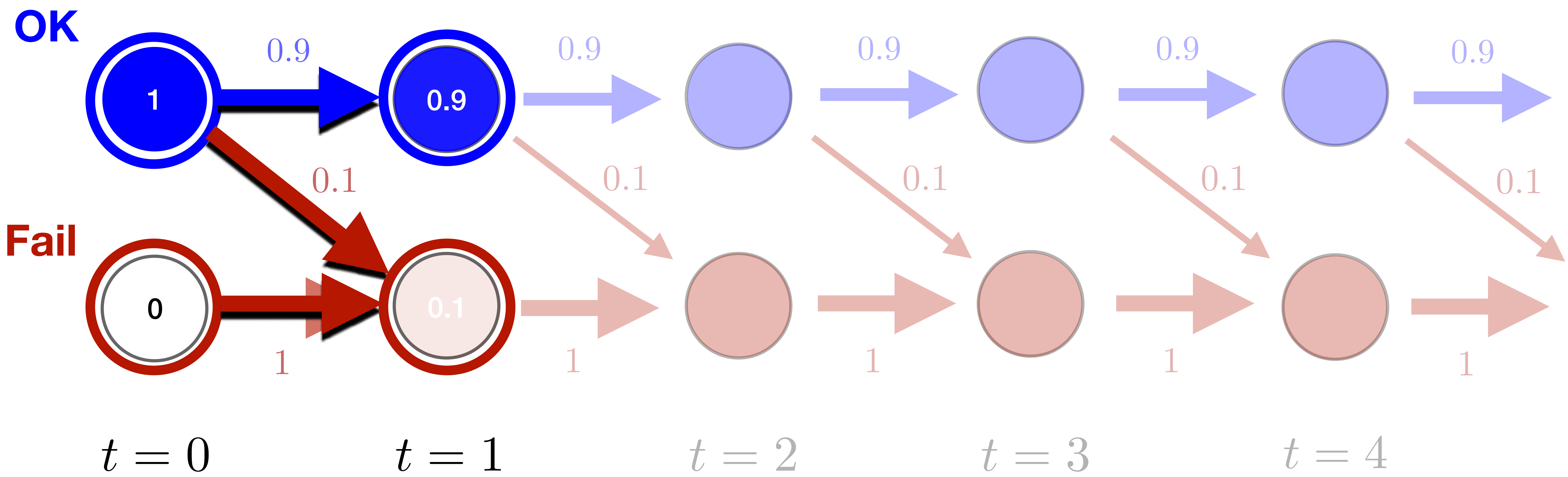
OK
Fail
1
1 × 0.9
1 × 0.9 × 0.9
1 × 0.9 × 0.9 × 0.9
0.9
0.9
0.9
1
0.9
0.81
0.729
0.1
0.1
0.1
0
0.1
0.19
0.271
1
1
1
0
0 × 1
+ 1 × 0.1
0 × 1 × 1
+ 1 × 0.1 × 1
+ 1 × 0.9 × 0.1
0 × 1 × 1 × 1
+ 1 × 0.1 × 1 × 1
+ 1 × 0.9 × 0.1 × 1
+ 1 × 0.9 × 0.9 × 0.1
t = 0
t = 1
t = 2
t = 3



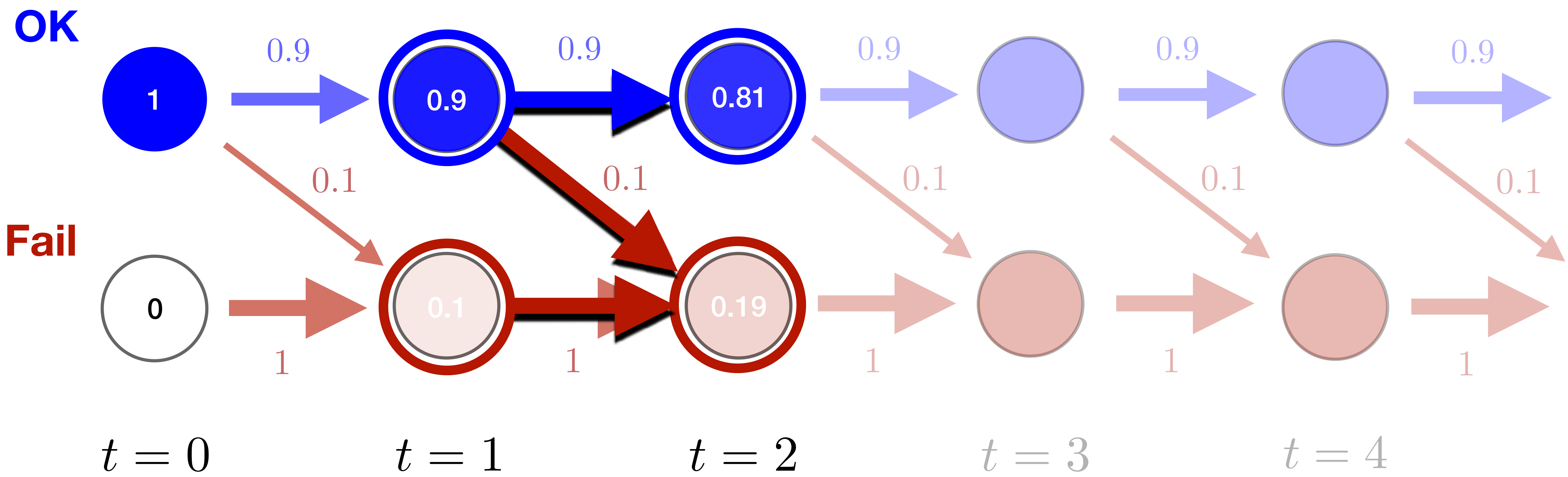
1
1 × 0.9
1 × 0.9 × 0.9
1 × 0.9 × 0.9 × 0.9
OK
Fail
1
0.9
0.81
0.729
0
0.1
0.19
0.271
0.9
0.9
0.9
0.1
0.1
0.1
1
1
1
0
0 × 1
+ 1 × 0.1
0 × 1 × 1
+ 1 × 0.1 × 1
+ 1 × 0.9 × 0.1
0 × 1 × 1 × 1
+ 1 × 0.1 × 1 × 1
+ 1 × 0.9 × 0.1 × 1
+ 1 × 0.9 × 0.9 × 0.1
t = 0
t = 1
t = 2
t = 3

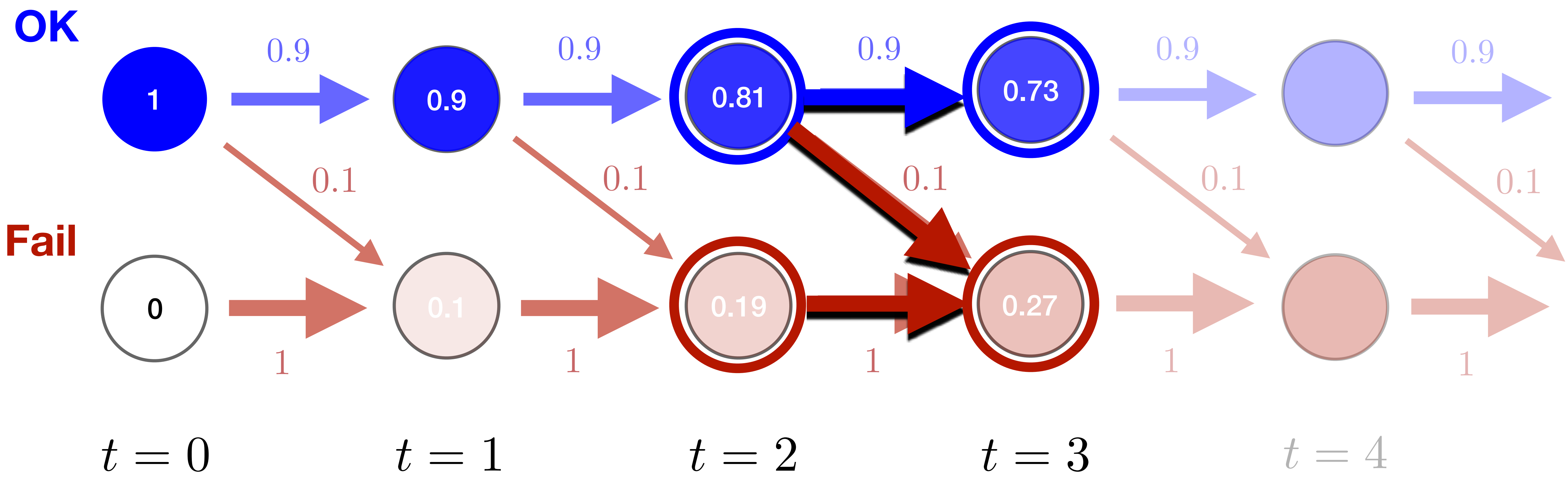

OK
Fail
1
0
1 × 0.9
0 × 1
+ 1 × 0.1
1 × 0.9 × 0.9
0 × 1 × 1
+ 1 × 0.1
+ 1 × 0.9 × 0.1
1 × 0.9 × 0.9 × 0.9
0 × 1 × 1 × 1
+ 1 × 0.1 × 1 × 1
+ 1 × 0.9 × 0.1 × 1
+ 1 × 0.9 × 0.9 × 0.1
0.9
0.1
1
0.81
0.19
0.729
0.271
t = 0
t = 1
t = 2
t = 3

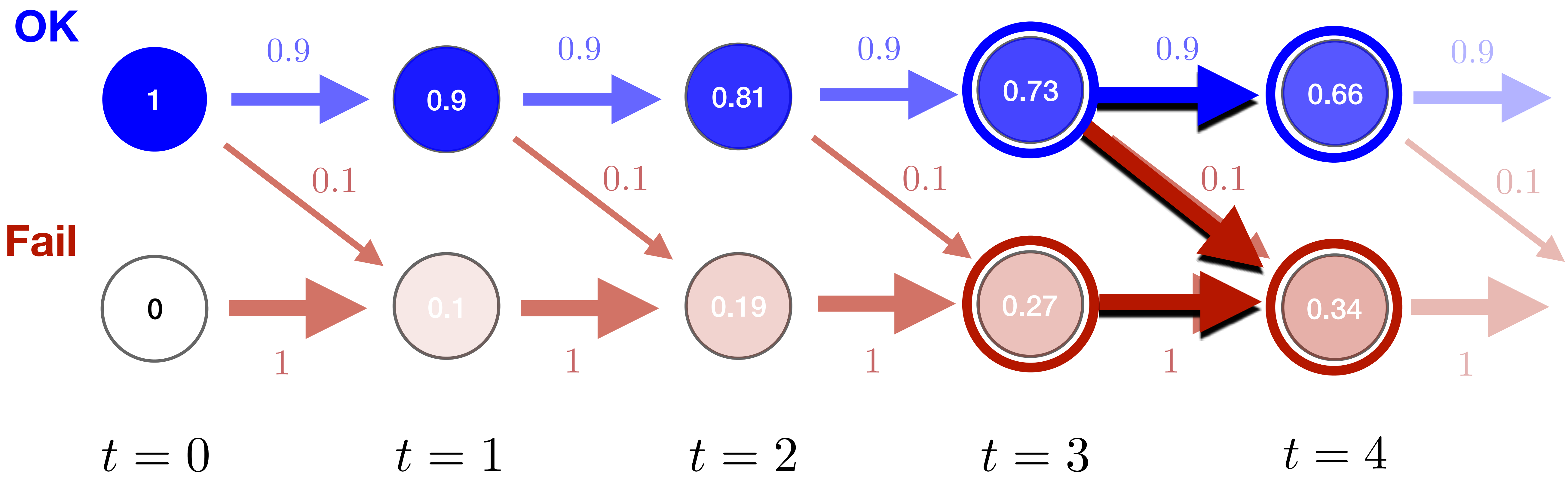

OK
Fail
1
1 × 0.9
1 × 0.9 × 0.9
1 × 0.9 × 0.9 × 0.9
0.9
0.9
0.9
1
0.9
0.81
0.729
0.1
0.1
0.1
0
0.1
0.19
0.271
1
1
1
0
0 × 1
+ 1 × 0.1
0 × 1
+ 1 × 0.1
+ 1 × 0.9
× 1
× 0.1
0 × 1 × 1
+ 1 × 0.1 × 1
+ 1 × 0.9 × 0.1
× 1
+ 1 × 0.9 × 0.9
× 0.1
t = 0
t = 1
t = 2
t = 3

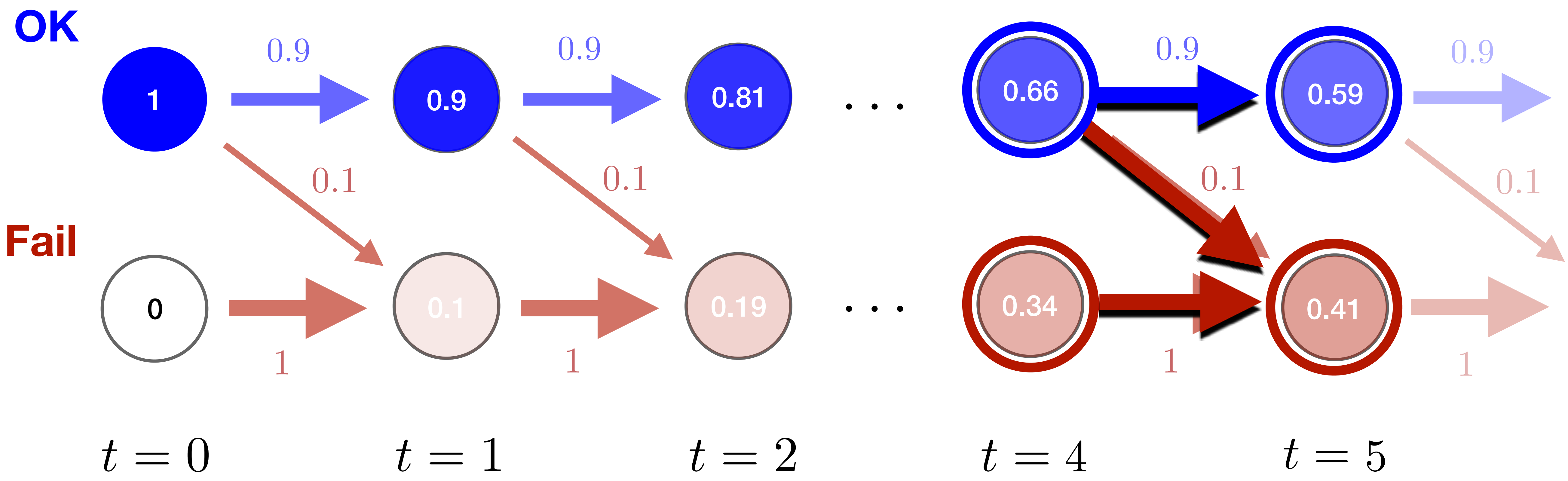


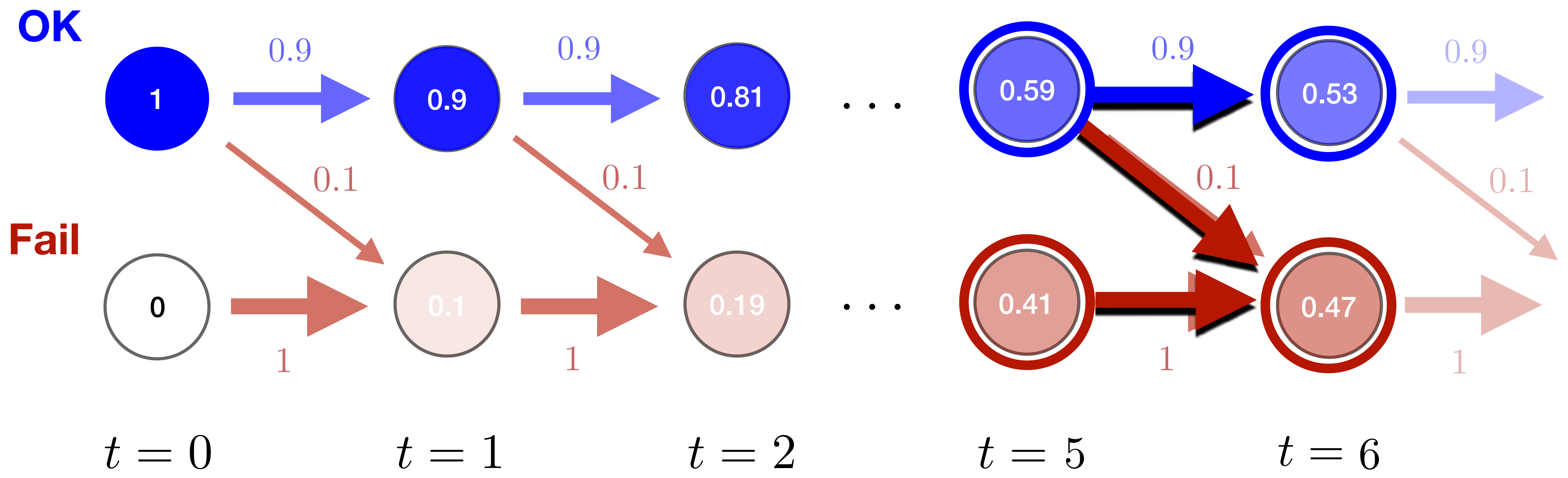
OK
Fail
1
0
0.9
0.1
0.9
0.1
1
0.9
0.1
1
0.9
0.1
1
0.9
0.1
1
0.9
0.1
1
t = 0
t = 1
t = 2
t = 3
t = 4

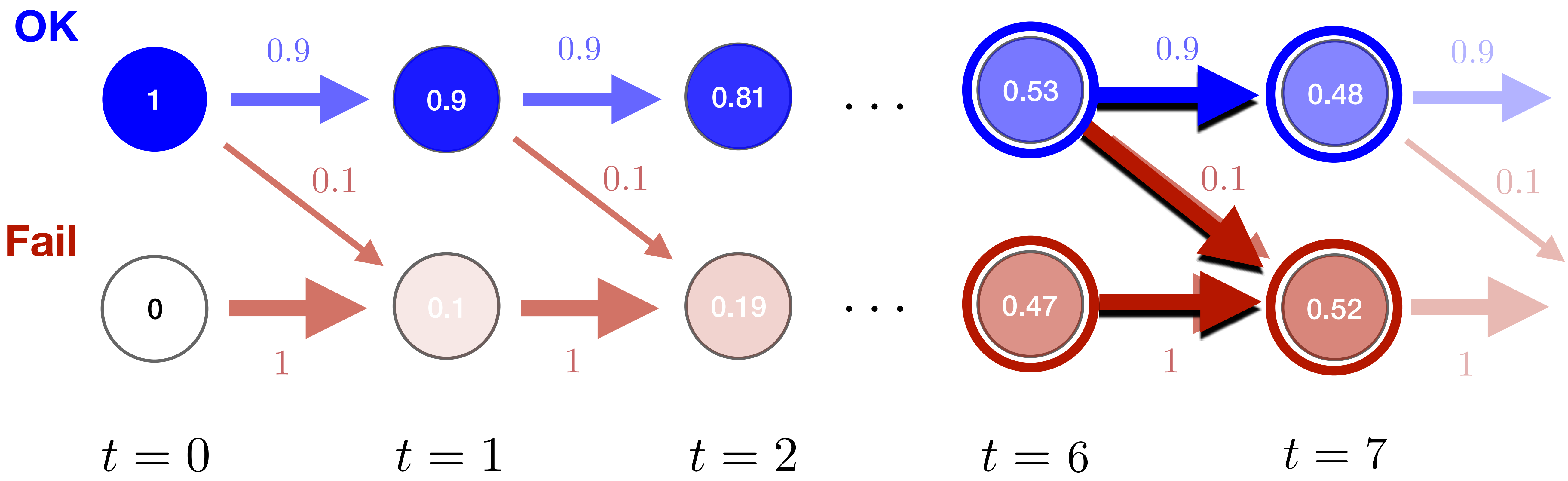

OK
Fail
1
0.9
0.81
0
0.1
0.19
0.9
0.1
1
$t = 0$
$t = 1$
$t = 2$
$t = 3$
$t = 4$


OK
Fail
1
0.9
0.81
0.73
0
0.1
0.19
0.27
0.9
0.1
1
t = 0
t = 1
t = 2
t = 3
t = 4

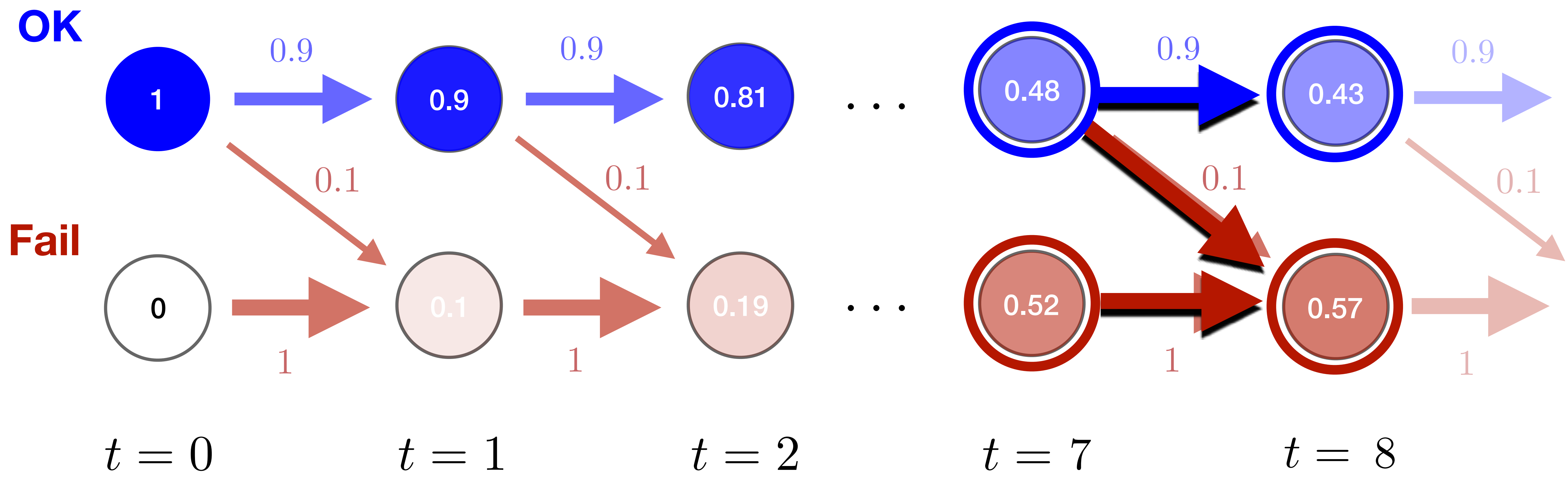

OK
Fail
0.9
0.9
0.9
0.9
0.9
0.1
0.1
0.1
0.1
0.1
1
0.9
0.81
0.73
0.66
0
0.1
0.19
0.27
0.34
1
1
1
1
1
$t = 0$
$t = 1$
$t = 2$
$t = 3$
$t = 4$

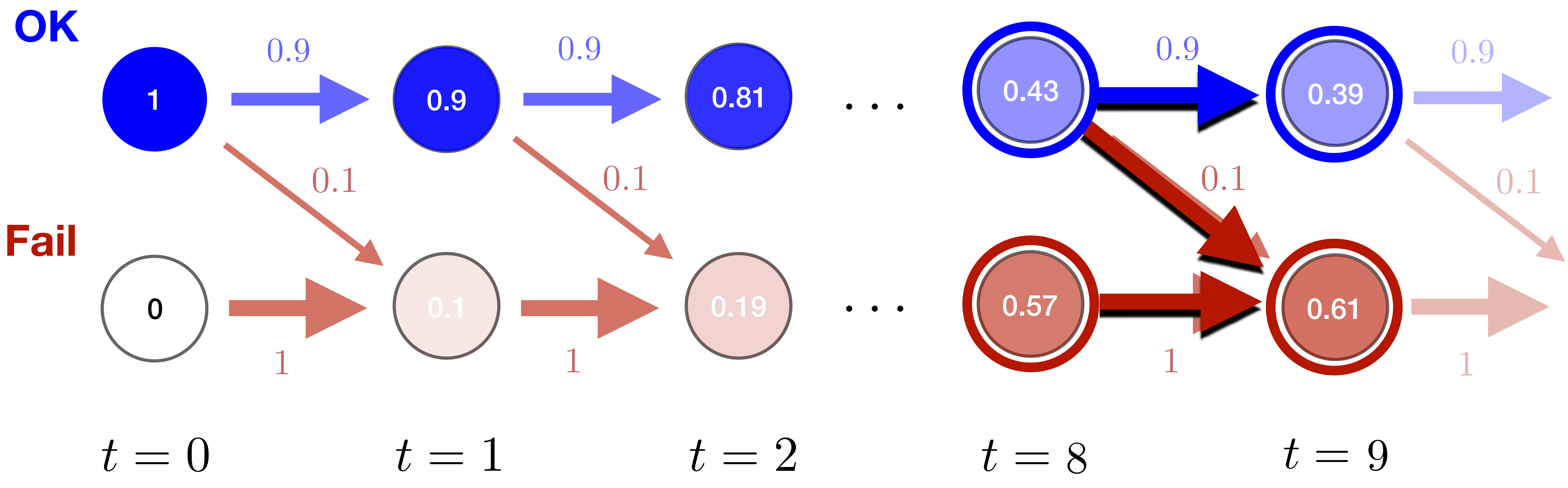

OK
Fail
1
0.9
0.81
0.66
0.59
0
0.1
0.19
0.34
0.41
0.9
0.9
0.9
0.9
0.1
0.1
0.1
0.1
1
1
1
1
$t = 0$
$t = 1$
$t = 2$
$t = 4$
$t = 5$

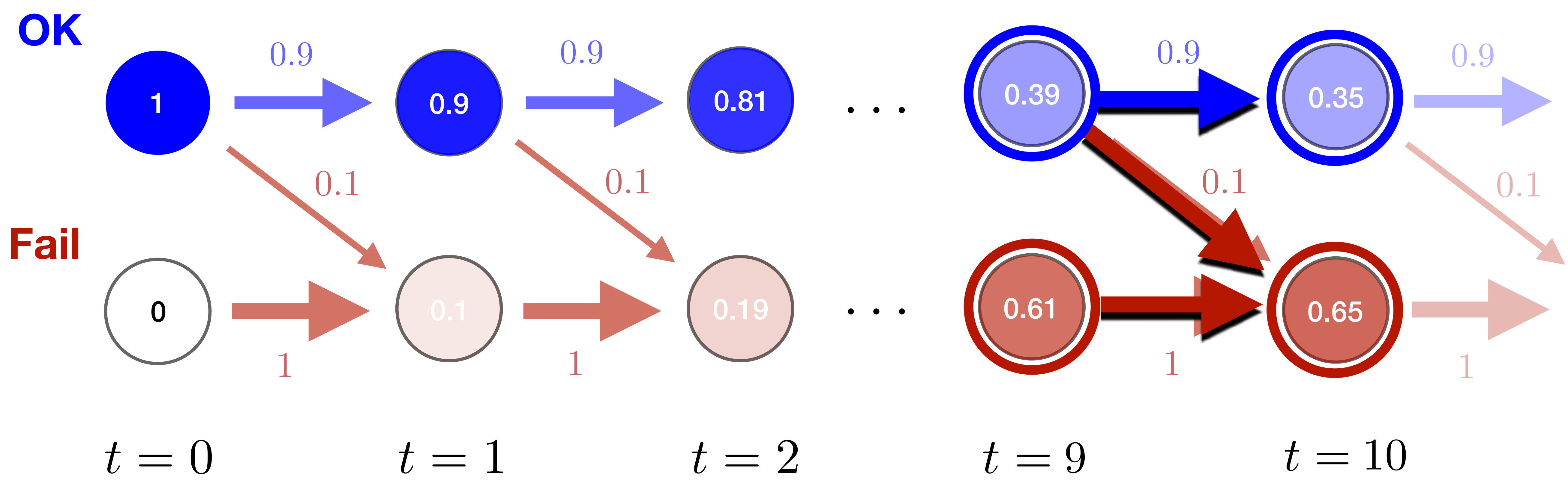

OK
Fail
1
0.9
0.81
0.59
0.53
0
0.1
0.19
0.41
0.47
0.9
0.1
1
t = 0
t = 1
t = 2
t = 5
t = 6

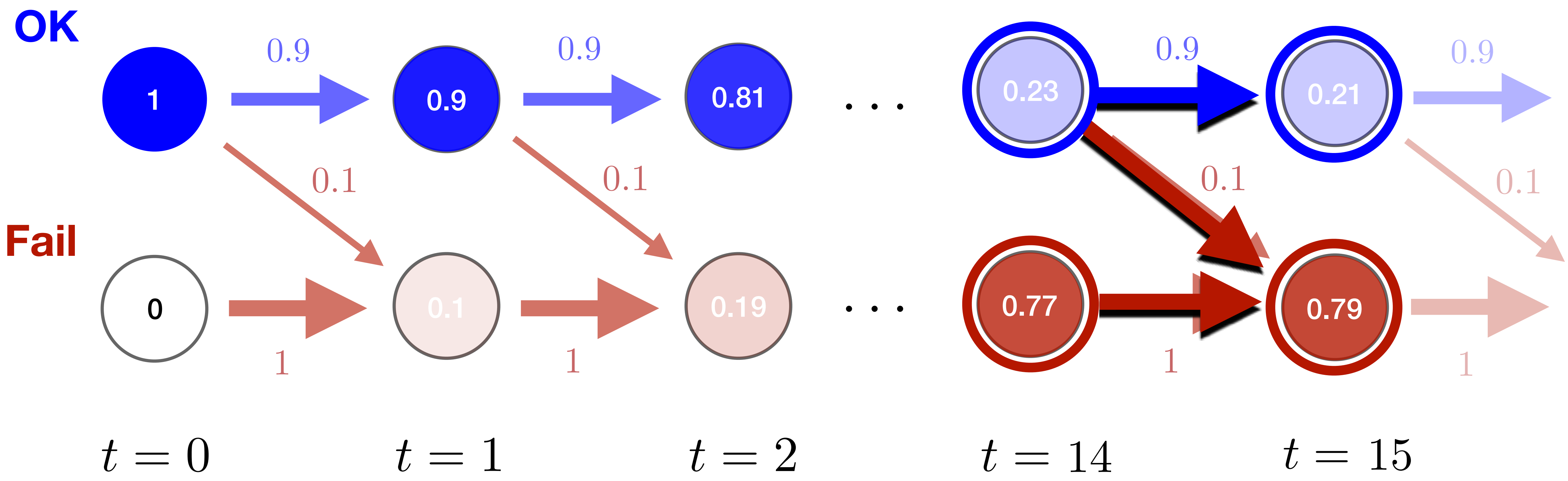


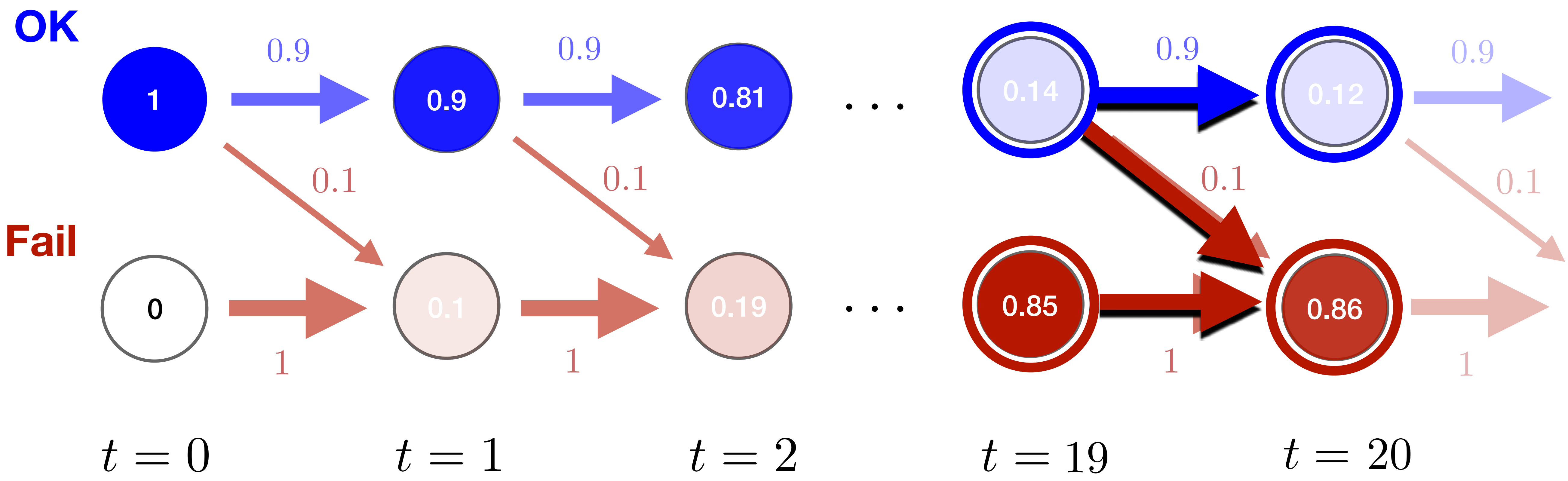
OK
Fail
1
0.9
0.81
0.53
0.48
0
0.1
0.19
0.47
0.52
0.9
0.9
0.9
0.9
0.1
0.1
0.1
0.1
1
1
1
1
. . .
. . .
$t = 0$
$t = 1$
$t = 2$
$t = 6$
$t = 7$

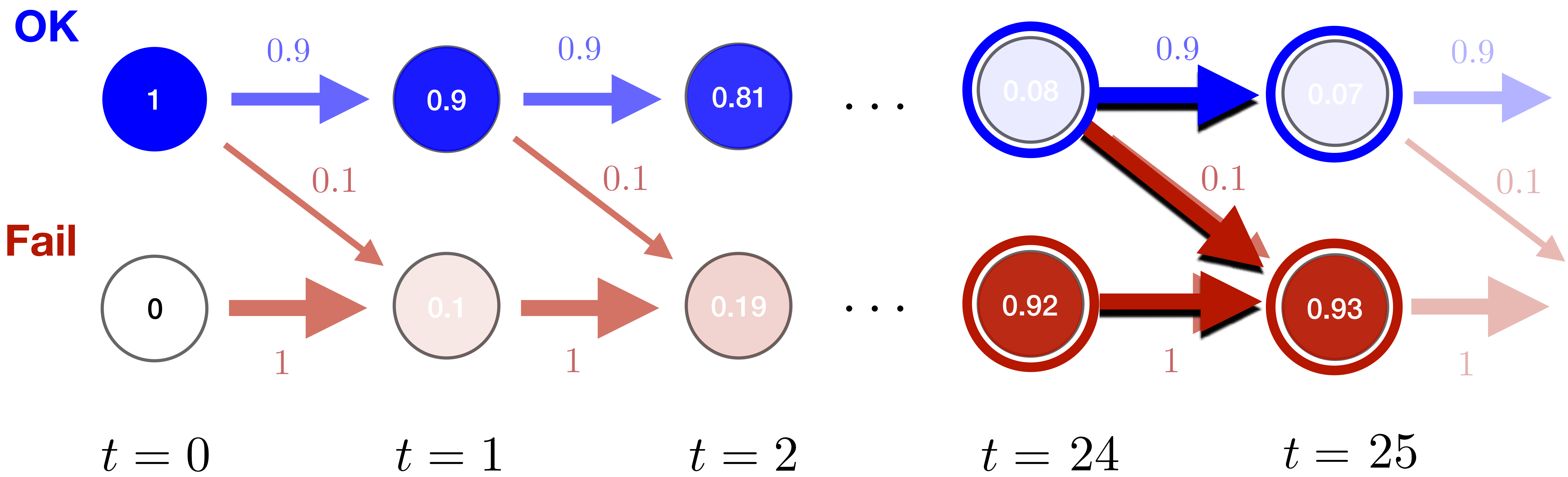

OK
Fail
1
0.9
0.81
0.48
0.43
0
0.1
0.19
0.52
0.57
0.9
0.1
1
. . .
$t = 0$
$t = 1$
$t = 2$
$t = 7$
$t = 8$


OK
Fail
1
0.9
0.81
0.43
0.39
0
0.1
0.19
0.57
0.61
0.9
0.9
0.9
0.9
0.1
0.1
0.1
0.1
1
1
1
1
. . .
. . .
$t = 0$
$t = 1$
$t = 2$
$t = 8$
$t = 9$

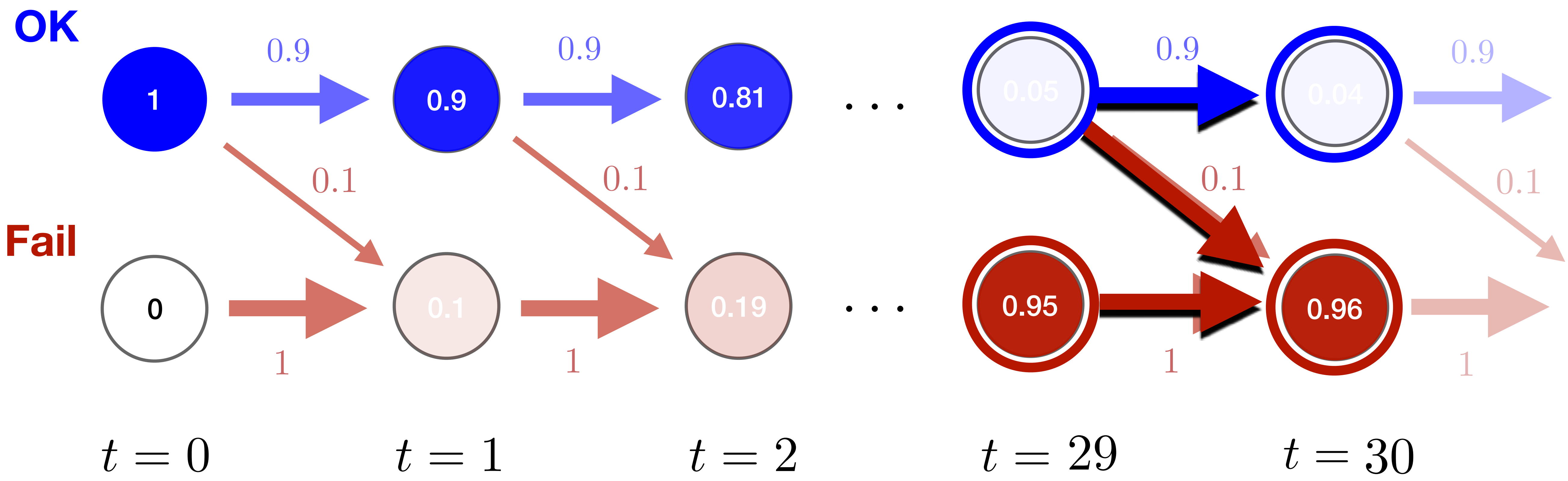

OK
Fail
0.9
0.9
0.9
0.9
0.1
0.1
0.1
0.1
1
1
1
1
1
0.9
0.81
0.39
0.35
0
0.1
0.19
0.61
0.65
t = 0
t = 1
t = 2
t = 9
t = 10


OK
Fail
1
0.9
0.81
0.23
0.21
0
0.1
0.19
0.77
0.79
0.9
0.1
1
t = 0
t = 1
t = 2
t = 14
t = 15
. . .

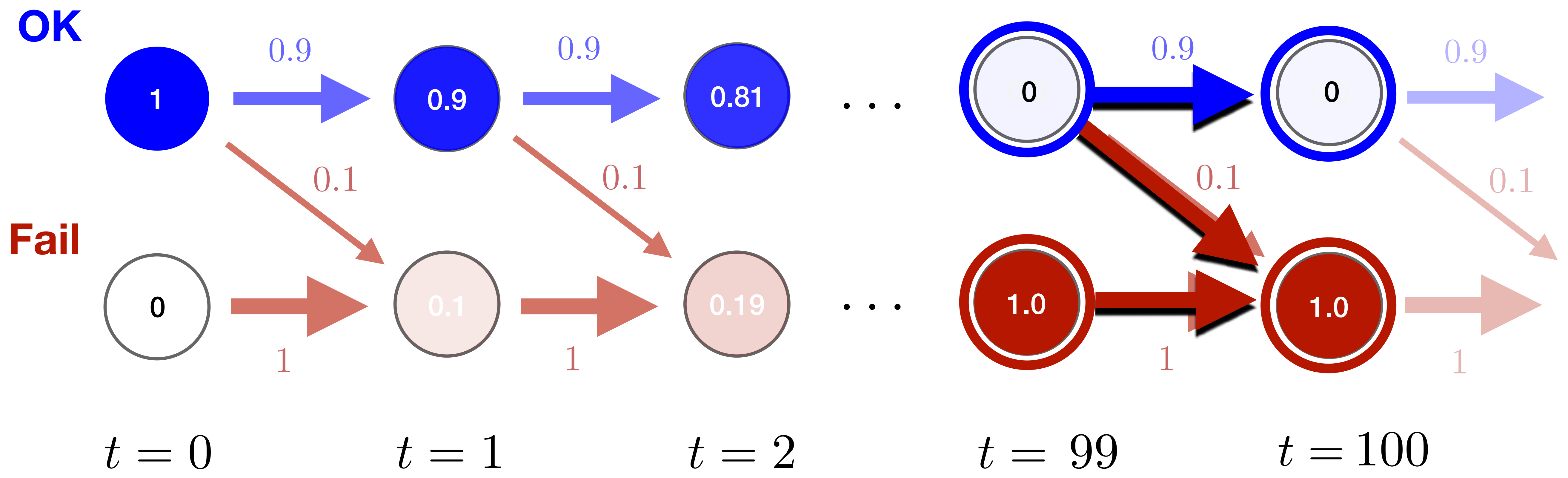

OK
Fail
1
0
0.9
0.1
0.9
0.1
1
0.81
0.19
1
0.14
0.85
0.9
0.1
1
0.12
0.86
0.9
0.1
1
t = 0
t = 1
t = 2
t = 19
t = 20



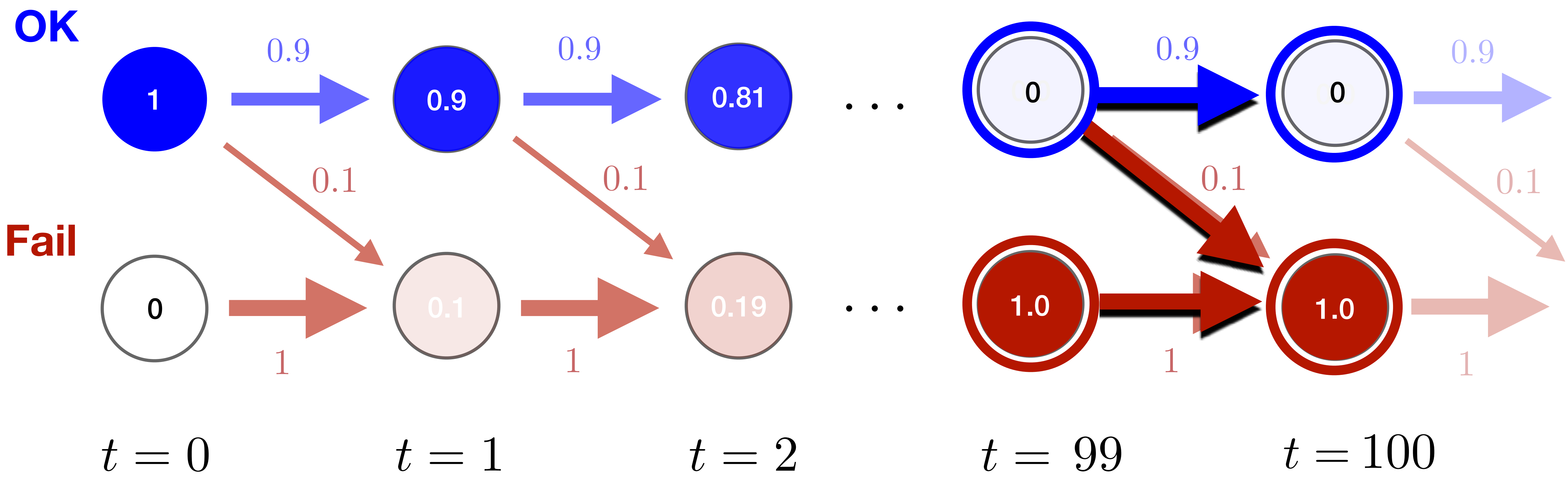
OK
Fail
1
0.9
0.81
0.08
0.07
0
0.1
0.19
0.92
0.93
0.9
0.9
0.9
0.9
0.1
0.1
0.1
0.1
1
1
1
1
. . .
. . .
$t = 0$
$t = 1$
$t = 2$
$t = 24$
$t = 25$

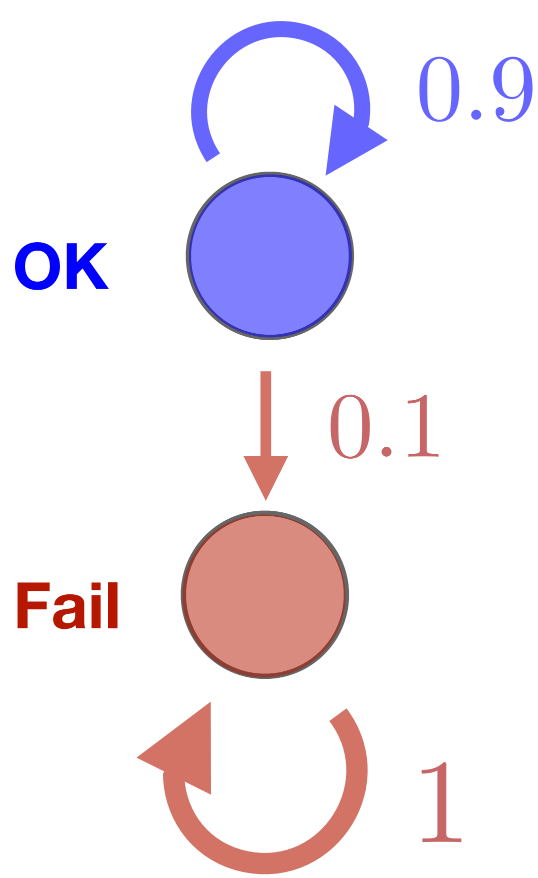


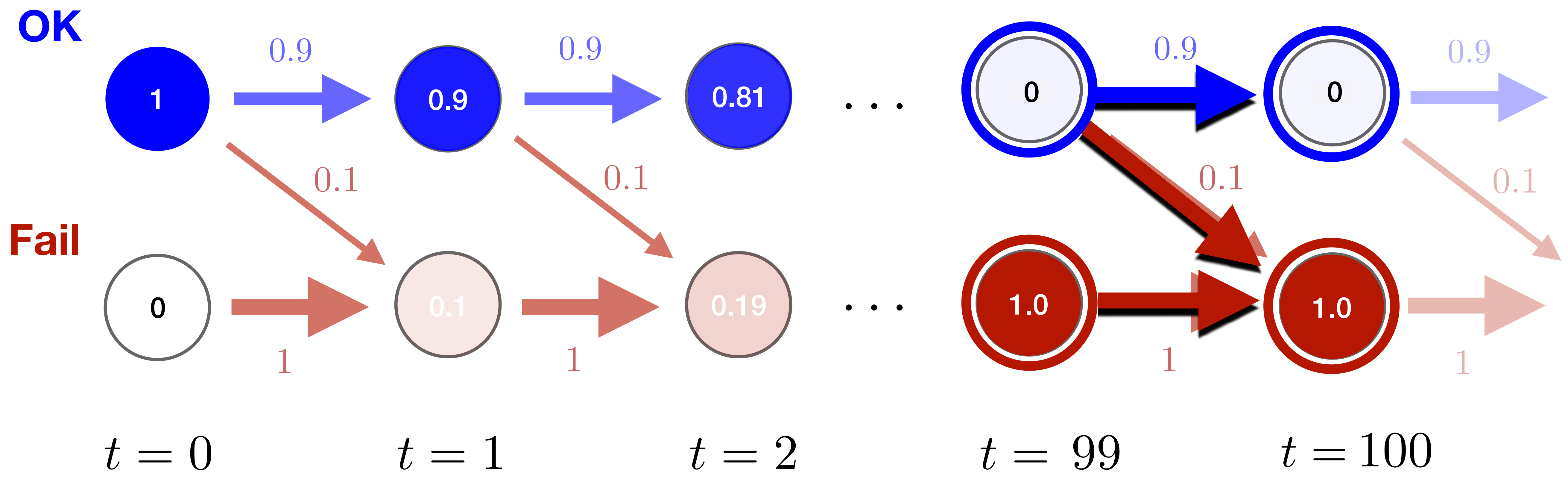
OK
Fail
1
0.9
0.81
0.05
0.04
0
0.1
0.19
0.95
0.96
0.9
0.1
1
. . .
t = 0
t = 1
t = 2
t = 29
t = 30

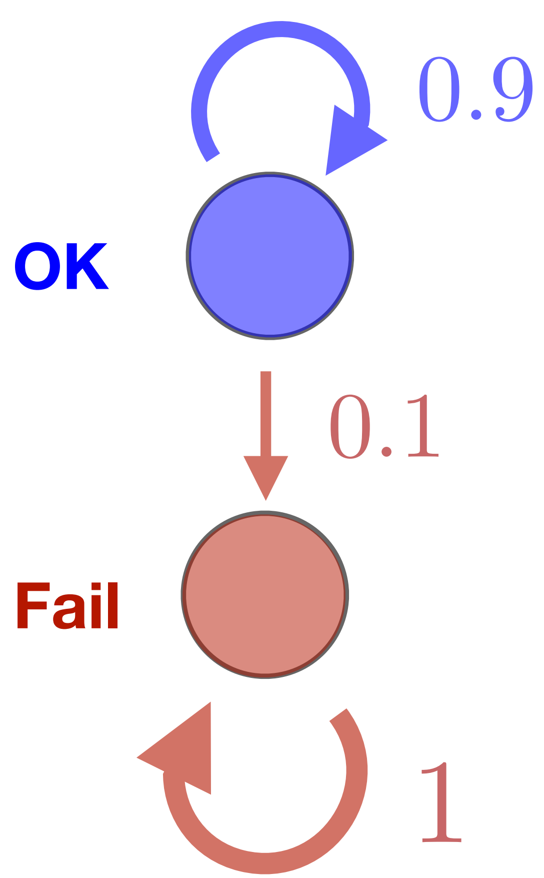


OK
Fail
1
0.9
0.81
0
0
0
0.1
0.19
1.0
1.0
0.9
0.9
0.9
0.9
0.1
0.1
0.1
0.1
1
1
1
1
. . .
. . .
$t = 0$
$t = 1$
$t = 2$
$t = 99$
$t = 100$

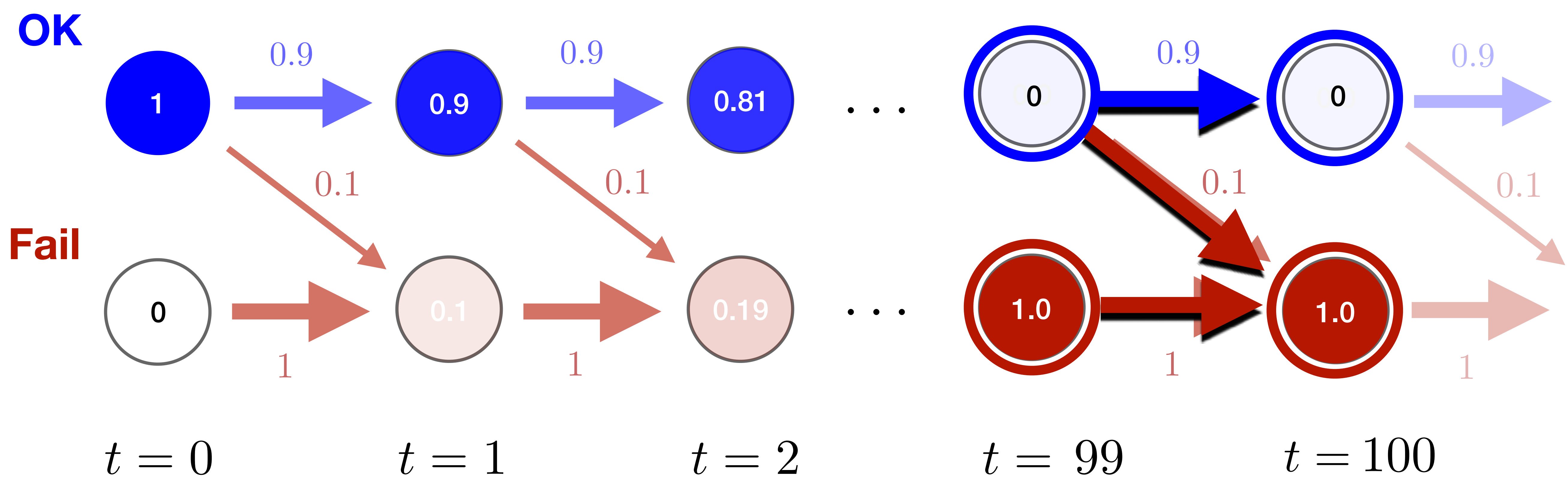

OK
Fail
1
0.9
0.81
0
0
0
0.1
0.19
1.0
1.0
0.9
0.9
0.9
0.9
0.1
0.1
0.1
0.1
1
1
1
1
. . .
. . .
$t = 0$
$t = 1$
$t = 2$
$t = 99$
$t = 100$

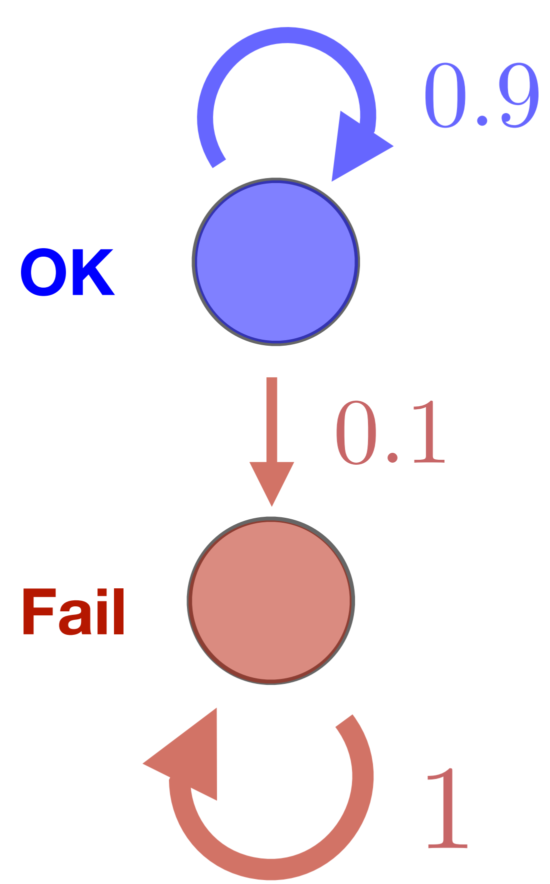

0.9
OK
0.1
Fail
1





$$\begin{bmatrix} p_1(1) \\ p_2(1) \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.1 & 1 \end{bmatrix} \begin{bmatrix} p_1(0) \\ p_2(0) \end{bmatrix}$$

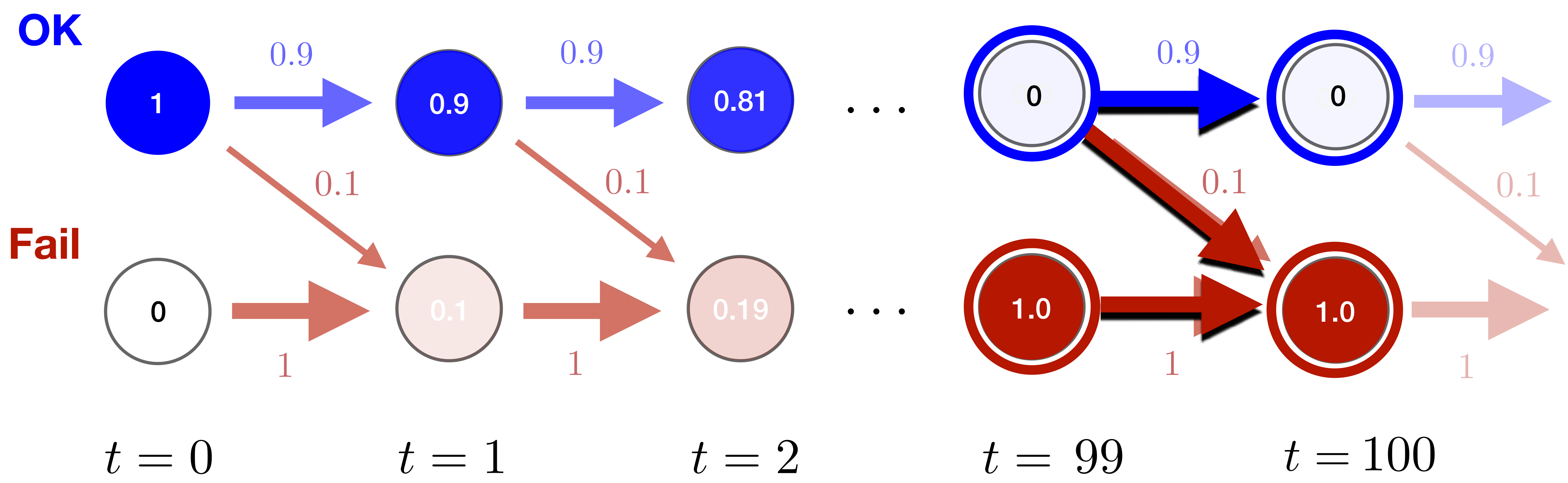


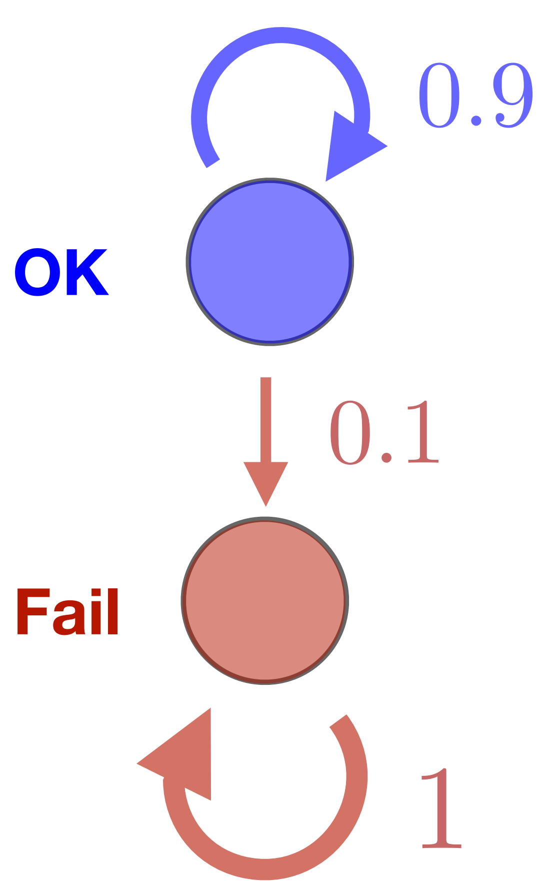


$$\begin{bmatrix} p_1(1) \\ p_2(1) \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.1 & 1 \end{bmatrix} \begin{bmatrix} p_1(0) \\ p_2(0) \end{bmatrix}$$

$$p_1(1) = 0.9 \times p_1(0) + 0 \times p_2(0)$$

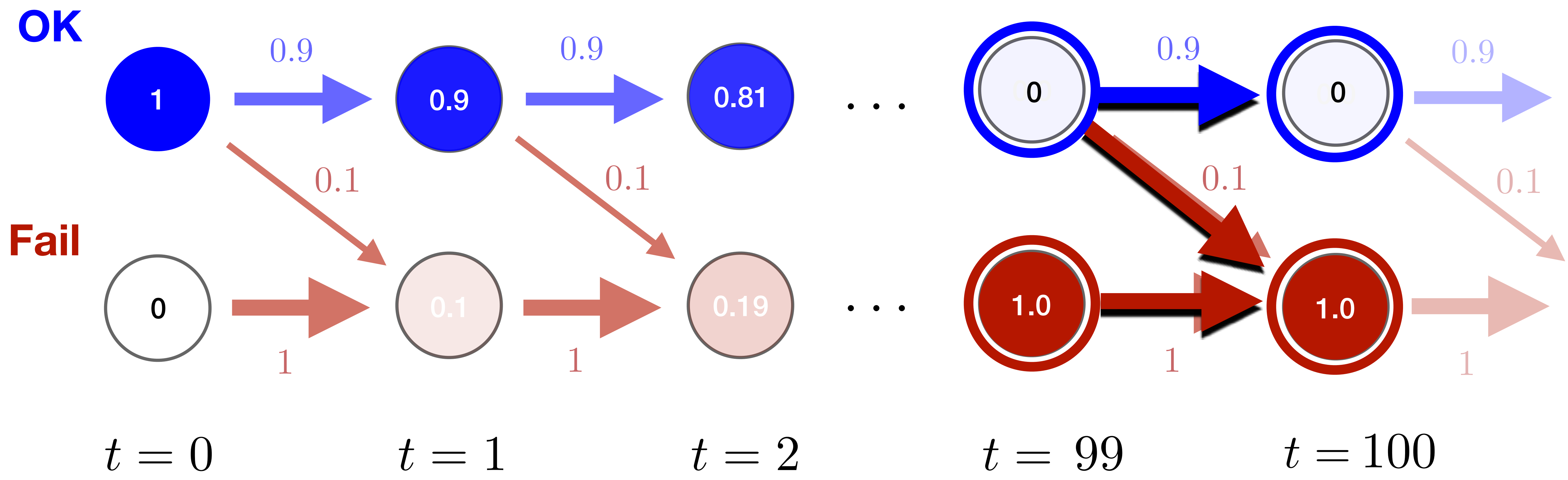
$$p_2(1) = 0.1 \times p_1(0) + 1 \times p_2(0)$$

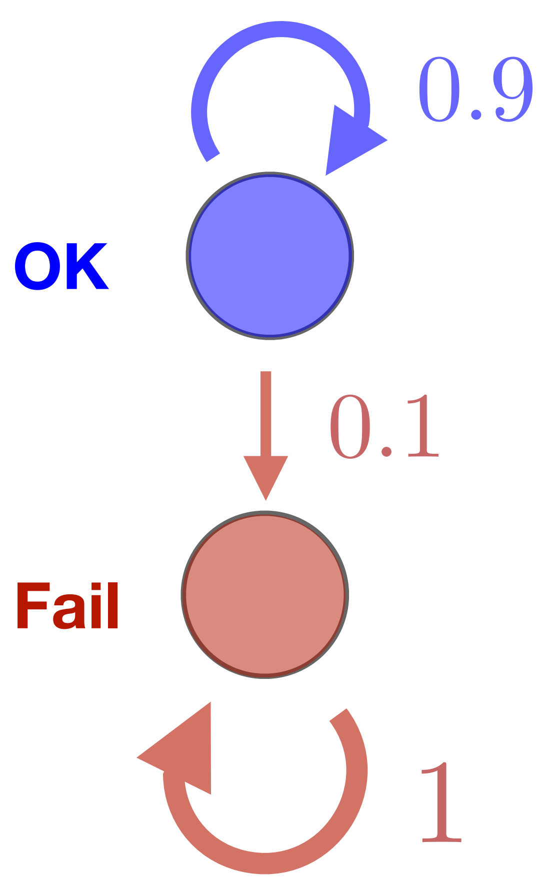




$$\begin{bmatrix} p_1(1) \\ p_2(1) \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.1 & 1 \end{bmatrix} \begin{bmatrix} p_1(0) \\ p_2(0) \end{bmatrix}$$

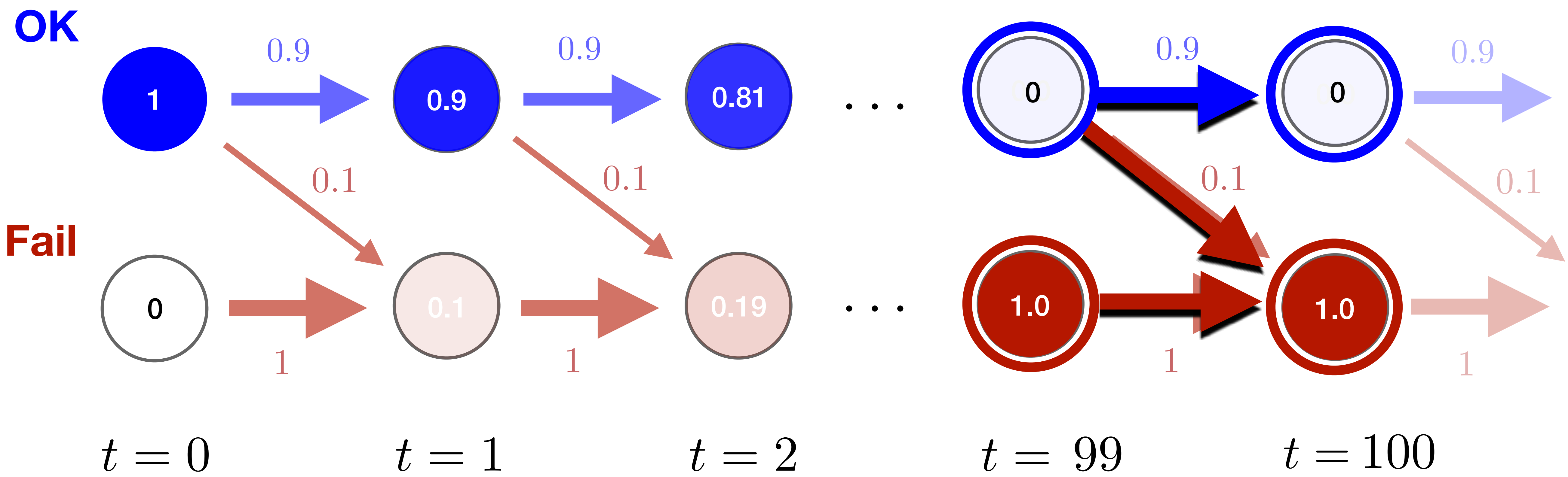
$$\begin{bmatrix} p_1(2) \\ p_2(2) \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.1 & 1 \end{bmatrix} \begin{bmatrix} p_1(1) \\ p_2(1) \end{bmatrix}$$

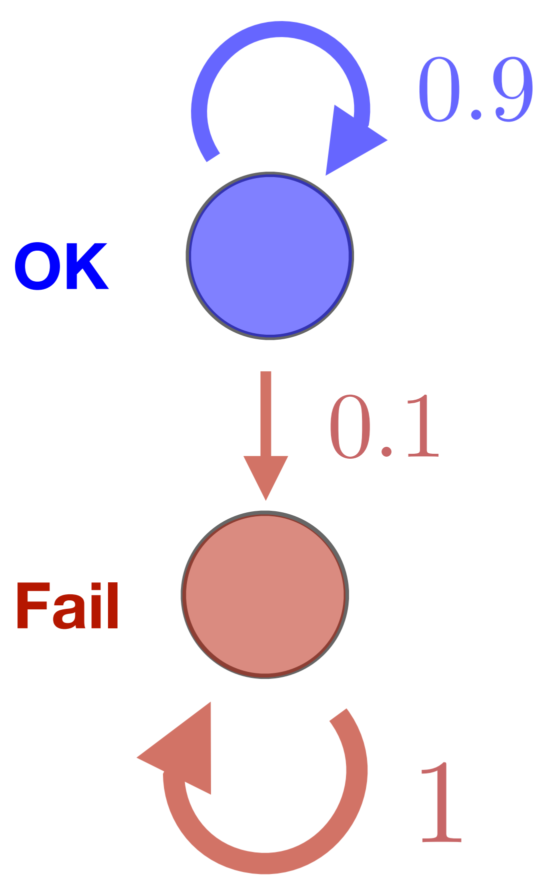




$$\begin{bmatrix} p_1(1) \\ p_2(1) \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.1 & 1 \end{bmatrix} \begin{bmatrix} p_1(0) \\ p_2(0) \end{bmatrix}$$

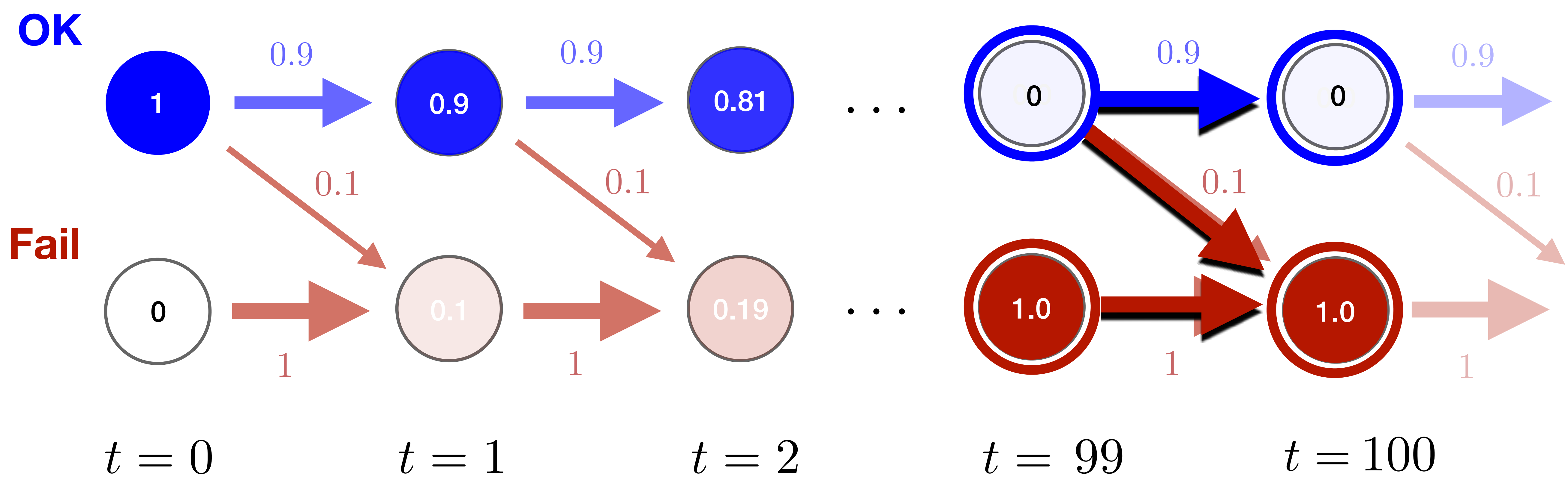
$$\begin{bmatrix} p_1(2) \\ p_2(2) \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.1 & 1 \end{bmatrix} \begin{bmatrix} 0.9 & 0 \\ 0.1 & 1 \end{bmatrix} \begin{bmatrix} p_1(0) \\ p_2(0) \end{bmatrix}$$

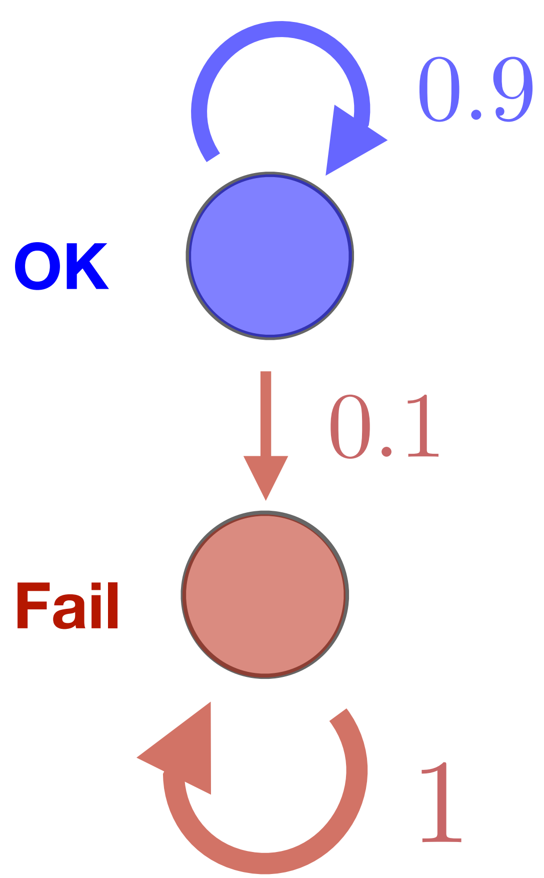




$$\begin{bmatrix} p_1(1) \\ p_2(1) \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.1 & 1 \end{bmatrix} \begin{bmatrix} p_1(0) \\ p_2(0) \end{bmatrix}$$

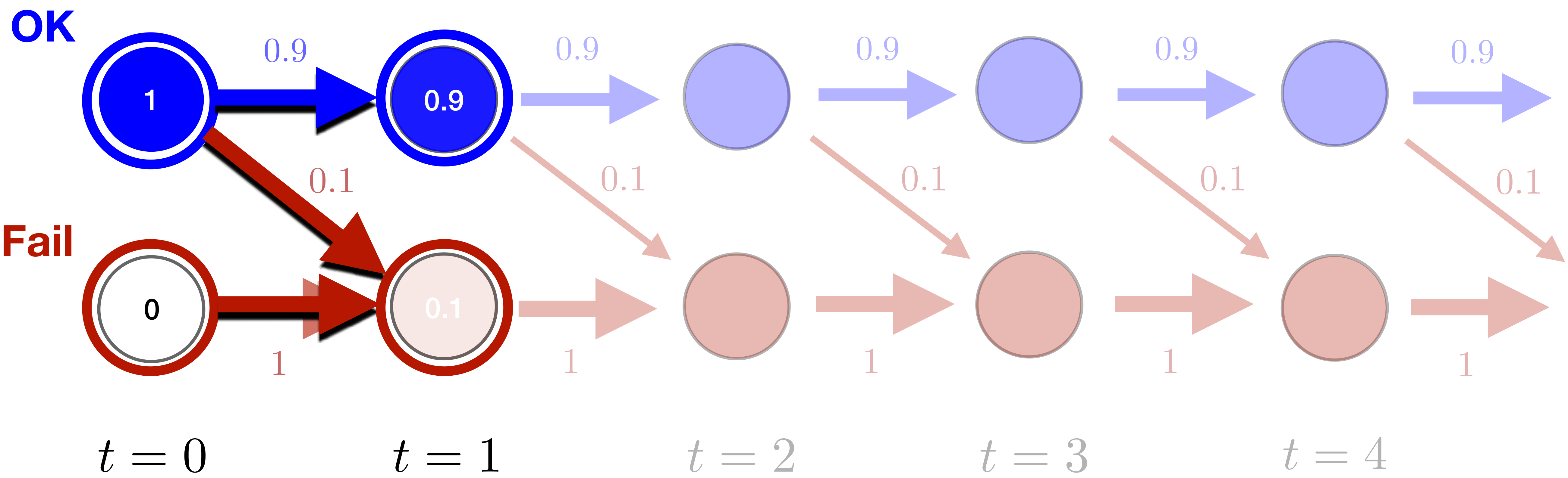
$$\begin{bmatrix} p_1(2) \\ p_2(2) \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.1 & 1 \end{bmatrix}^2 \begin{bmatrix} p_1(0) \\ p_2(0) \end{bmatrix}$$

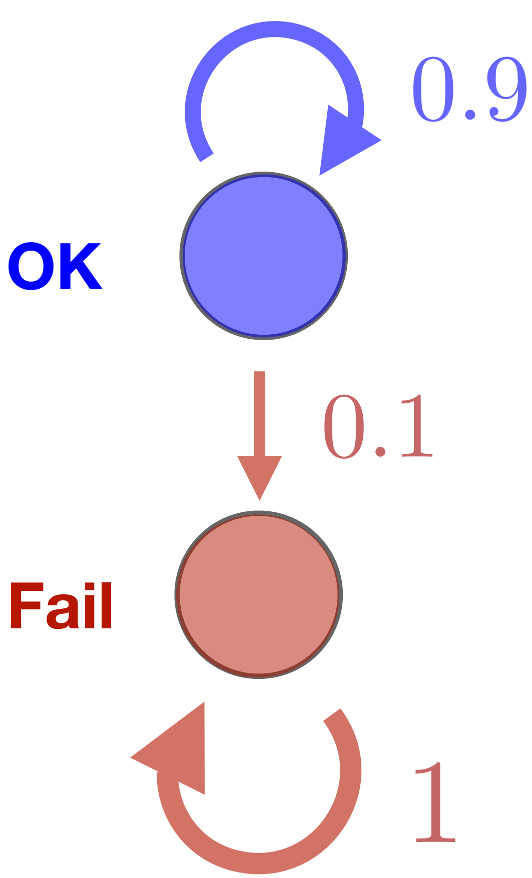




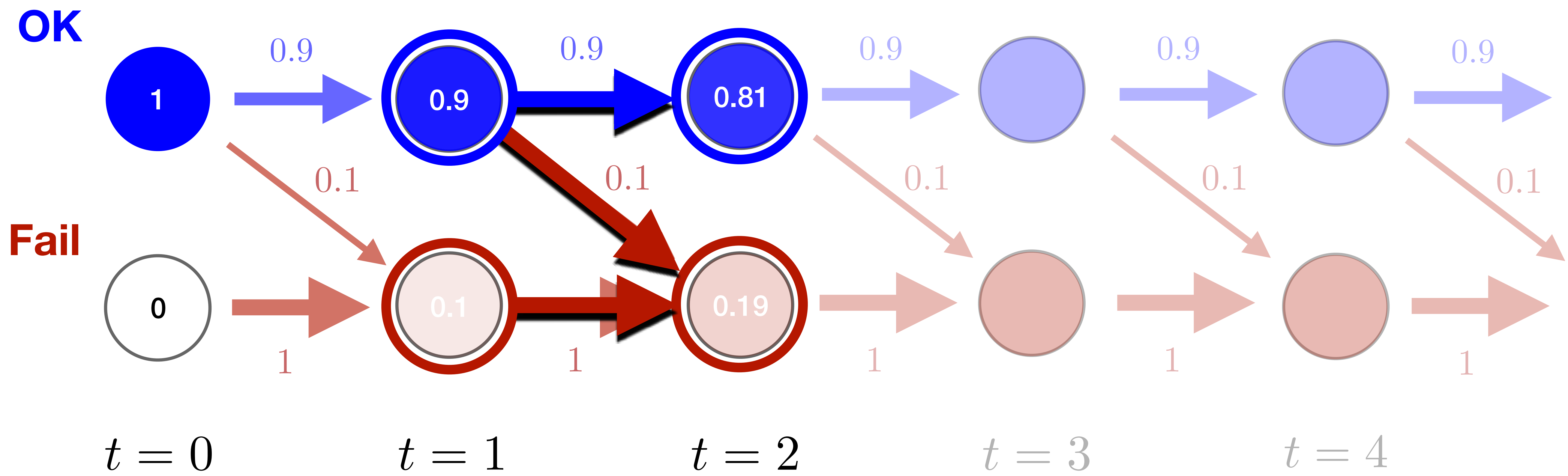
$$\begin{bmatrix} p_1(1) \\ p_2(1) \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.1 & 1 \end{bmatrix} \begin{bmatrix} p_1(0) \\ p_2(0) \end{bmatrix}$$

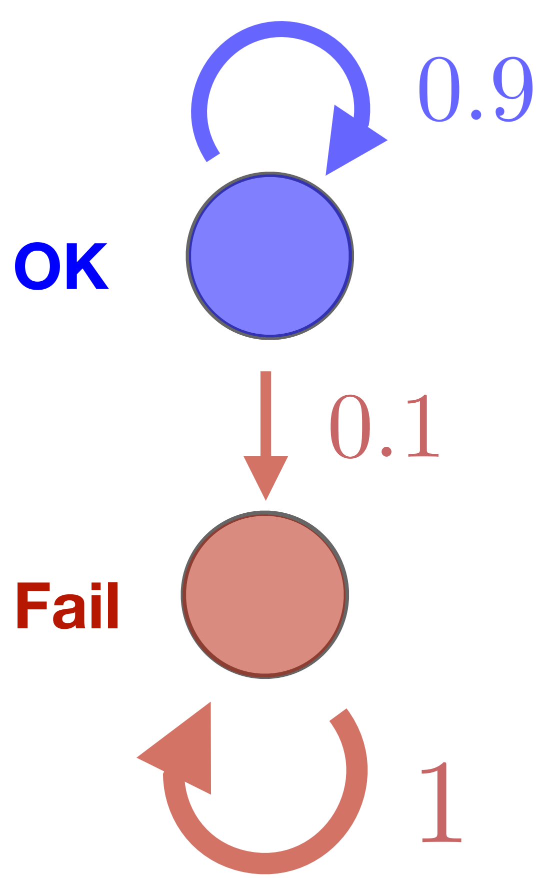
$$\begin{bmatrix} p_1(k) \\ p_2(k) \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.1 & 1 \end{bmatrix}^k \begin{bmatrix} p_1(0) \\ p_2(0) \end{bmatrix}$$



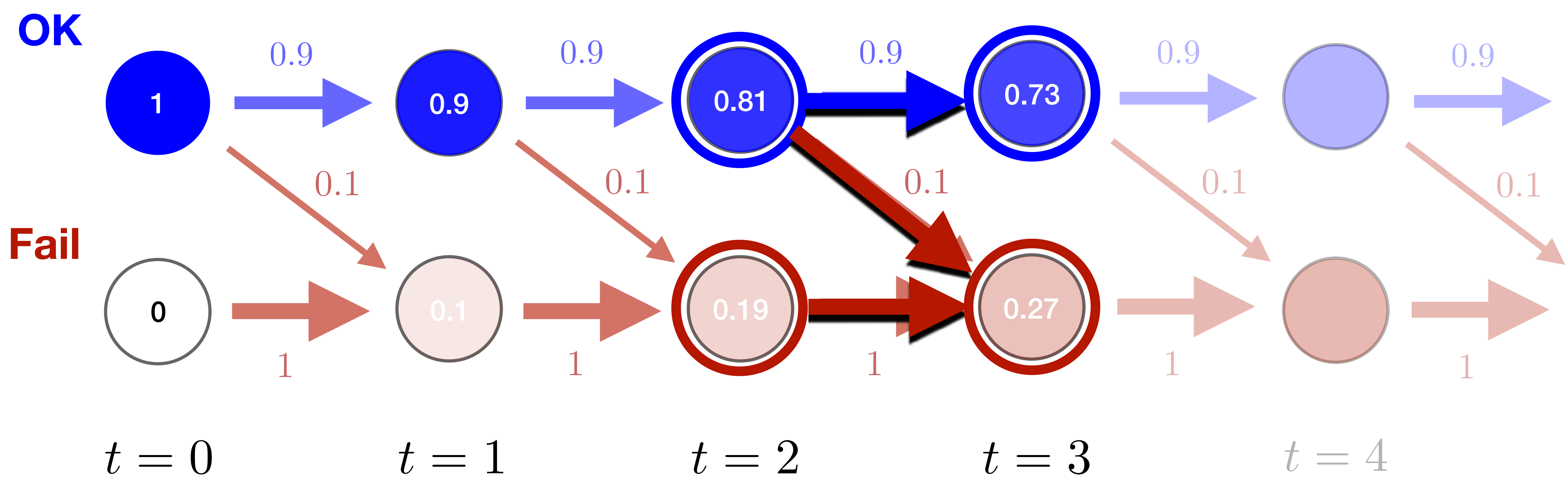


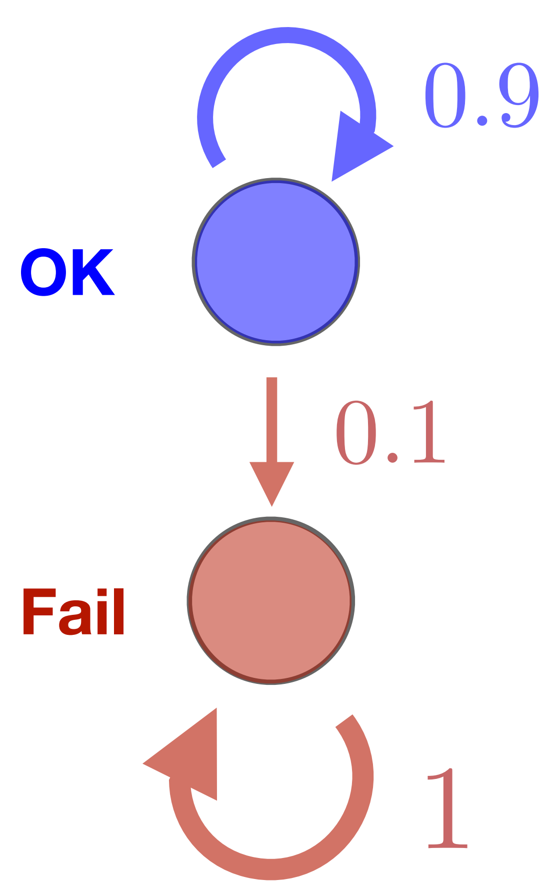
$$\begin{bmatrix} 0.9 \\ 0.1 \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.1 & 1 \end{bmatrix}^{1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$



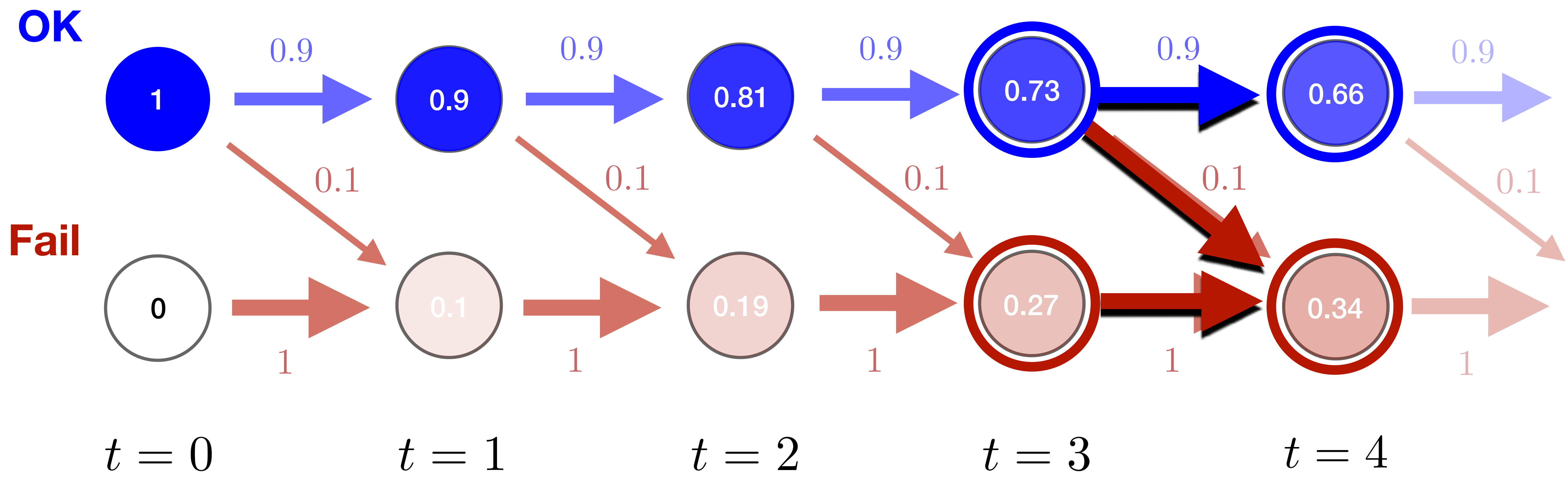


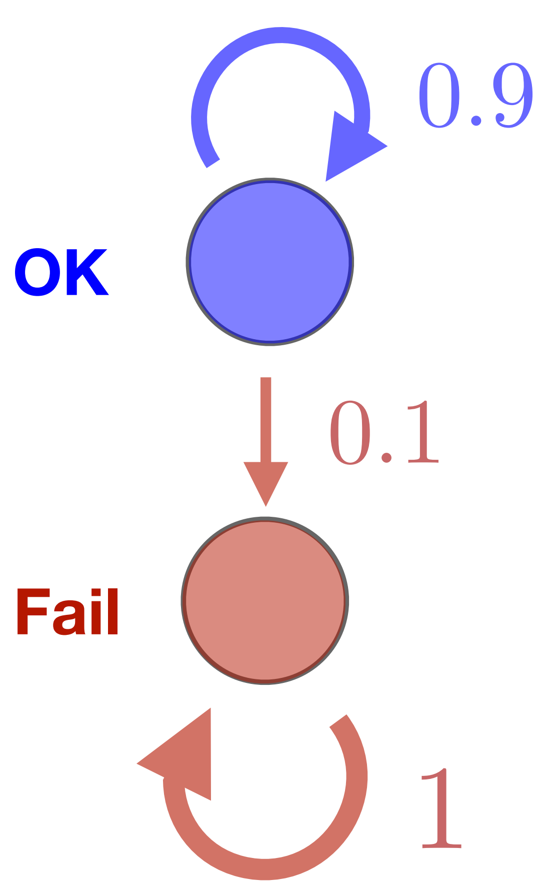
$$\begin{bmatrix} 0.81 \\ 0.19 \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.1 & 1 \end{bmatrix}^2 \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$



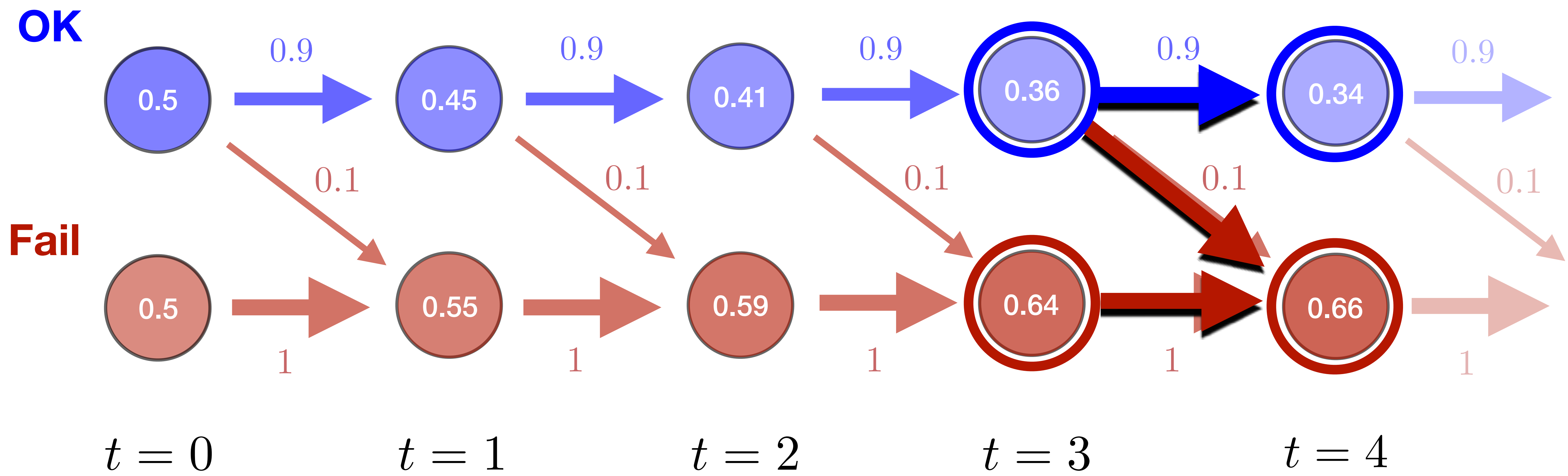


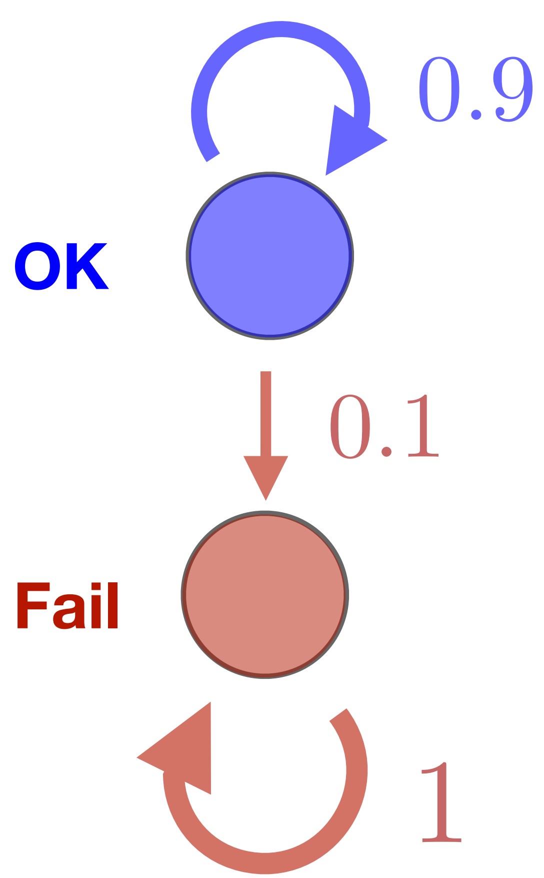
$$\begin{bmatrix}0.73\\0.27\end{bmatrix} = \begin{bmatrix}0.9 & 0\\0.1 & 1\end{bmatrix}^3 \begin{bmatrix}1\\0\end{bmatrix}$$



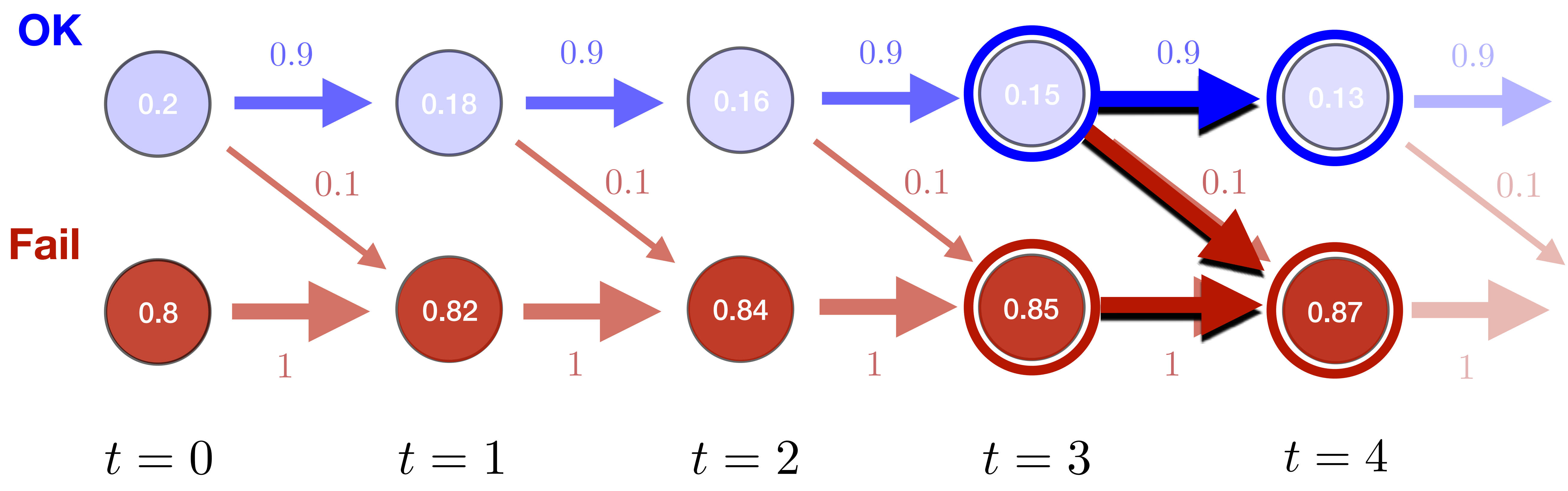


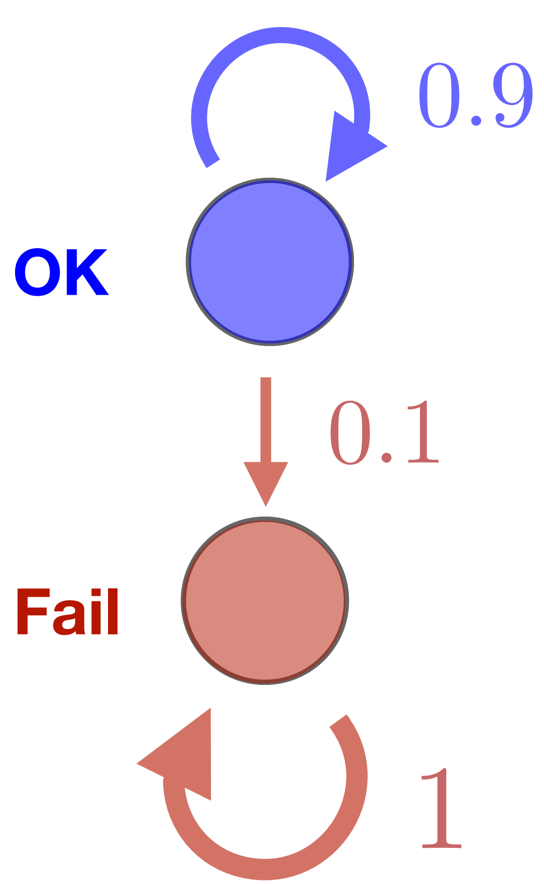
$$\begin{bmatrix} 0.66 \\ 0.34 \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.1 & 1 \end{bmatrix}^4 \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$



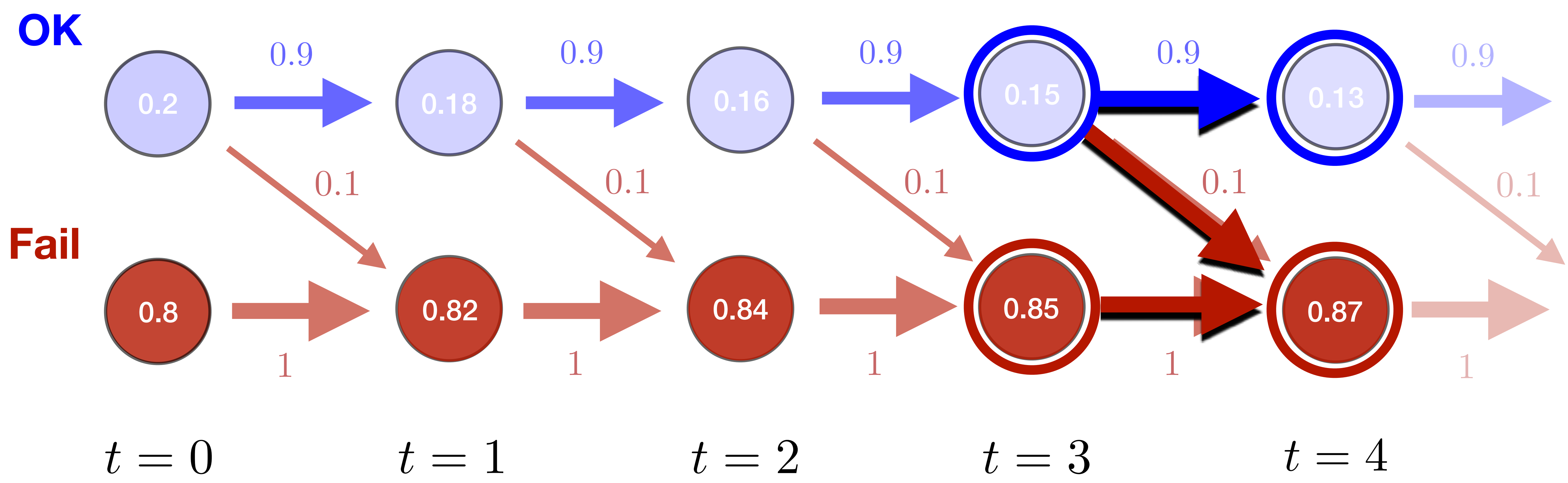


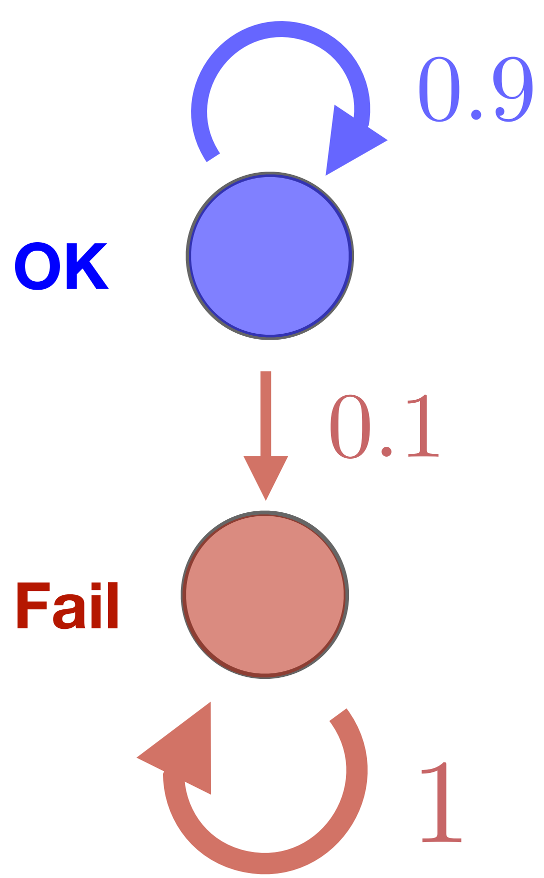
$$\begin{bmatrix} 0.66 \\ 0.34 \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.1 & 1 \end{bmatrix}^4 \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$





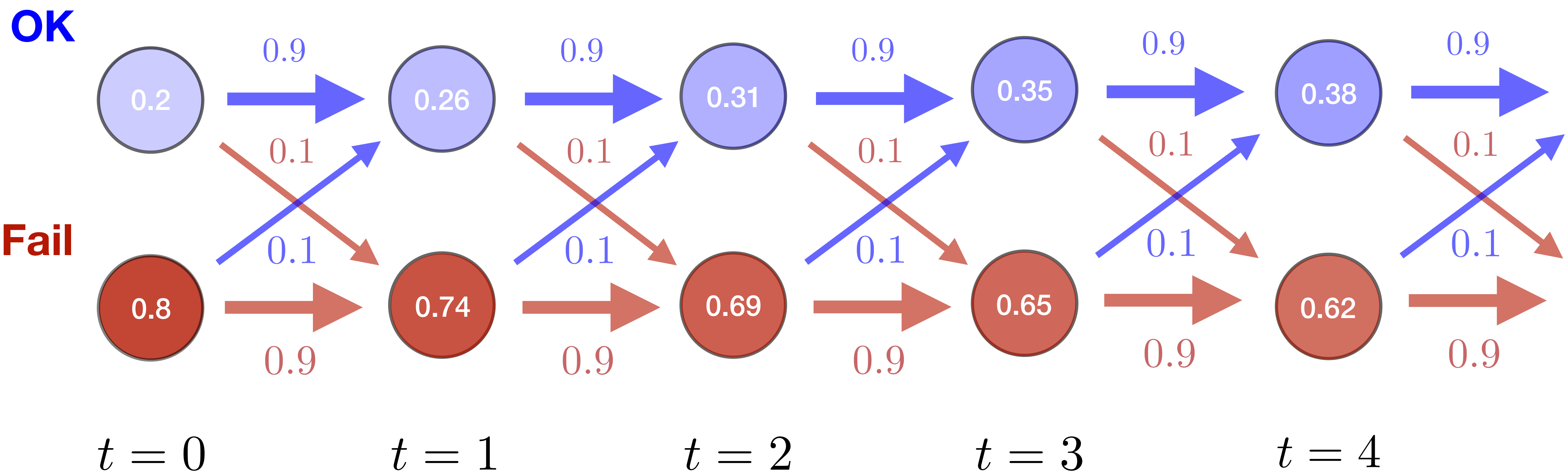
$$\begin{bmatrix} 0.13 \\ 0.87 \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.1 & 1 \end{bmatrix}^4 \begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix}$$

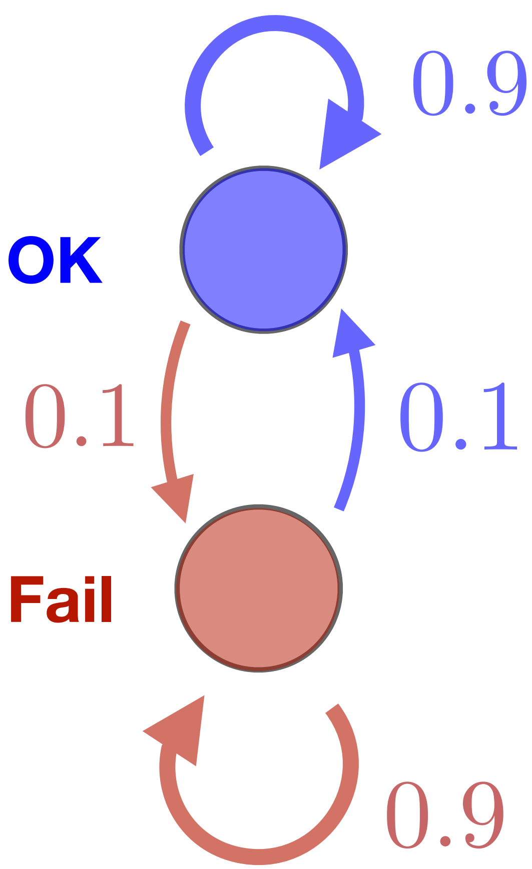




$$\begin{bmatrix} p_1(k) \\ p_2(k) \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.1 & 1 \end{bmatrix}^k \begin{bmatrix} p_1(0) \\ p_2(0) \end{bmatrix}$$

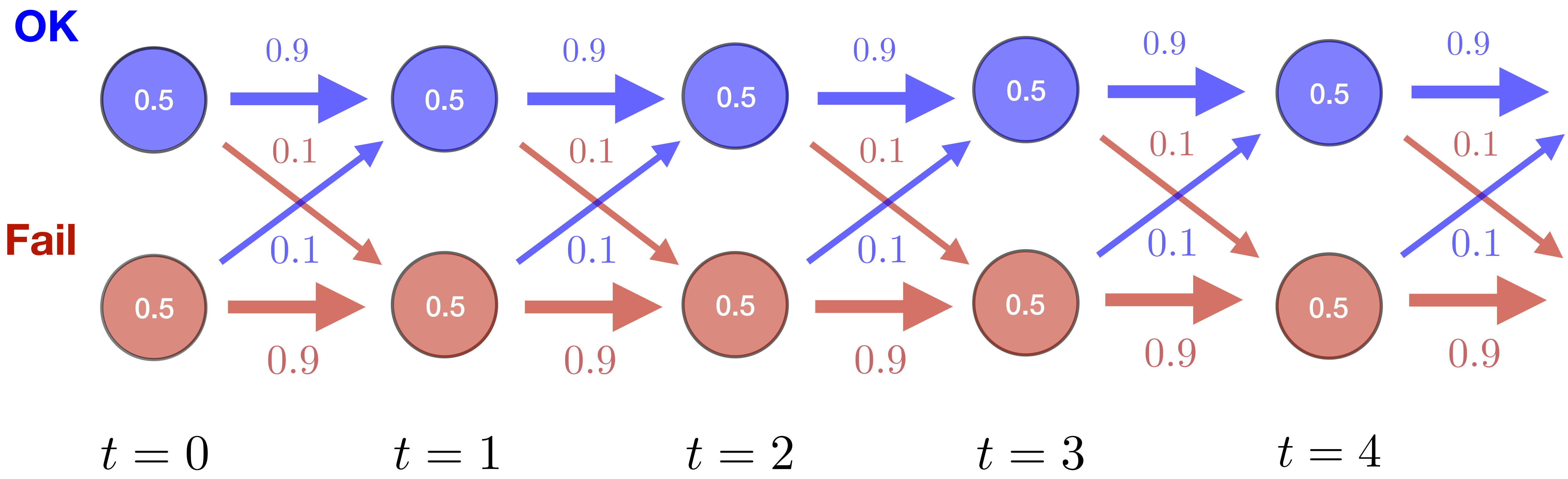
$$\lim_{k\to\infty} \begin{bmatrix} p_1(k) \\ p_2(k) \end{bmatrix} = \begin{bmatrix} p_1(\infty) \\ p_2(\infty) \end{bmatrix} = \begin{bmatrix} 0.0 \\ 1.0 \end{bmatrix}$$

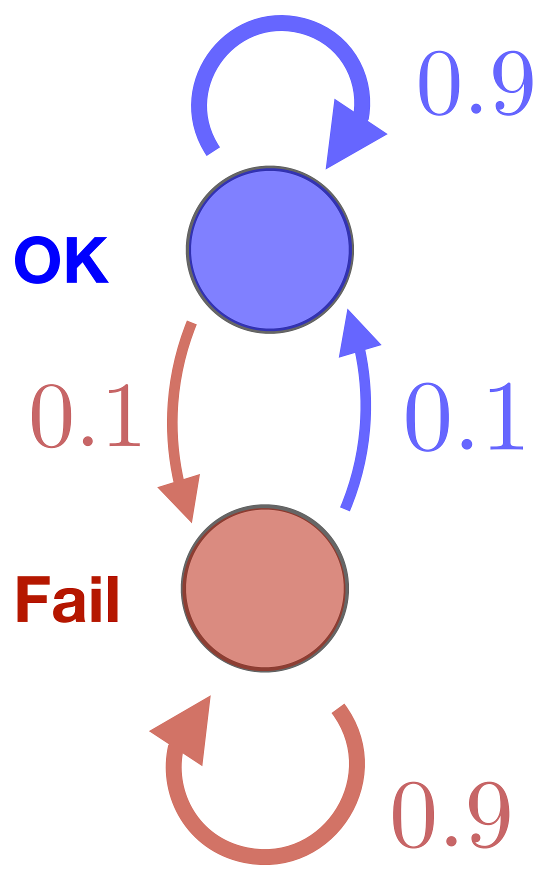




$$\begin{bmatrix} p_1(k) \\ p_2(k) \end{bmatrix} = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}^k \begin{bmatrix} p_1(0) \\ p_2(0) \end{bmatrix}$$

$$\lim_{k\to\infty} \begin{bmatrix} p_1(k) \\ p_2(k) \end{bmatrix} = \begin{bmatrix} p_1(\infty) \\ p_2(\infty) \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

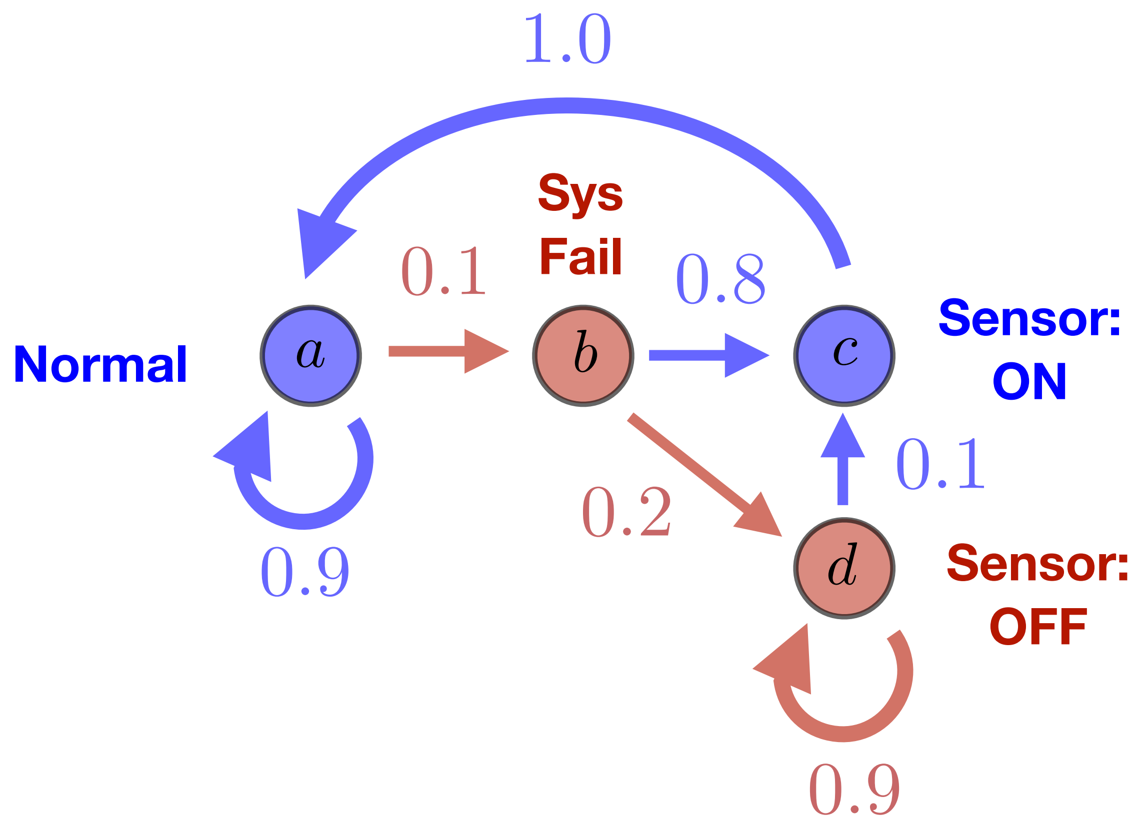


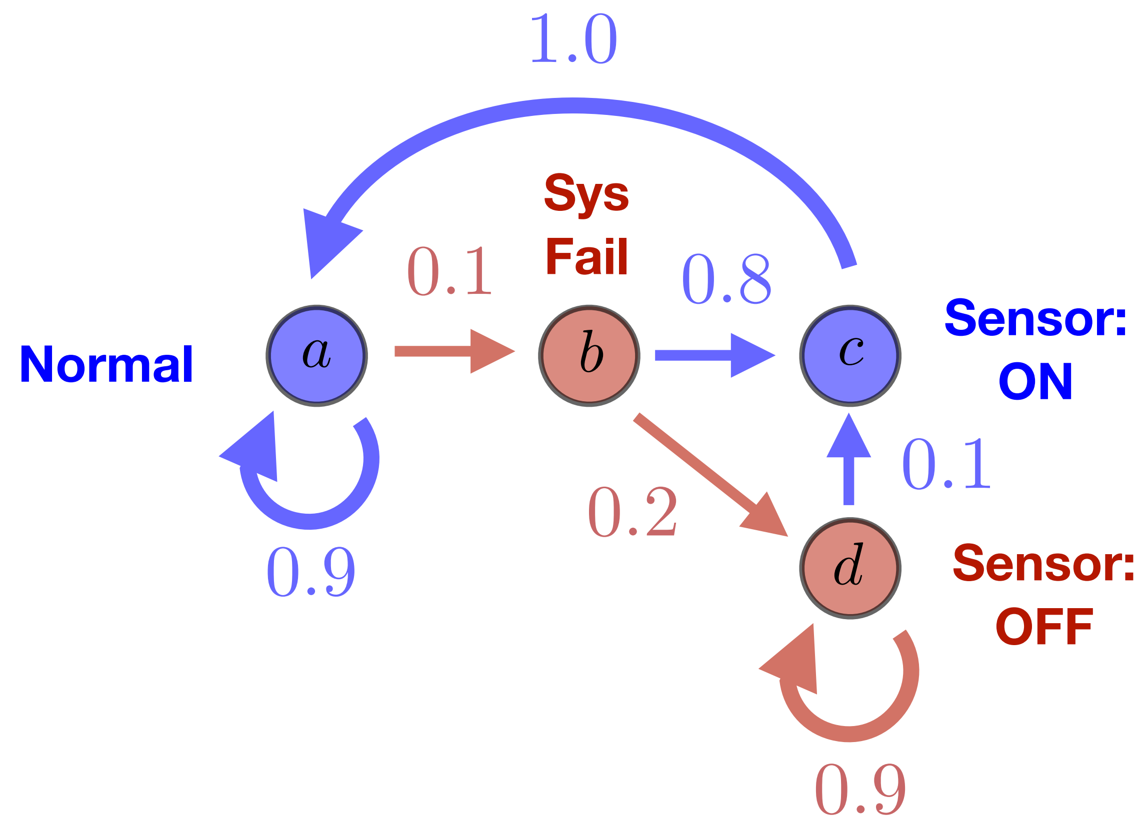


$$\begin{bmatrix} p_1(k) \\ p_2(k) \end{bmatrix} = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}^k \begin{bmatrix} p_1(0) \\ p_2(0) \end{bmatrix}$$

$$\lim_{k\to\infty} \begin{bmatrix} p_1(k) \\ p_2(k) \end{bmatrix} = \begin{bmatrix} p_1(\infty) \\ p_2(\infty) \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

examples


1.0
Sys
Fail
0.1
0.8
Normal
a
b
c
Sensor:
ON
0.1
0.2
0.9
d
Sensor:
OFF
0.9

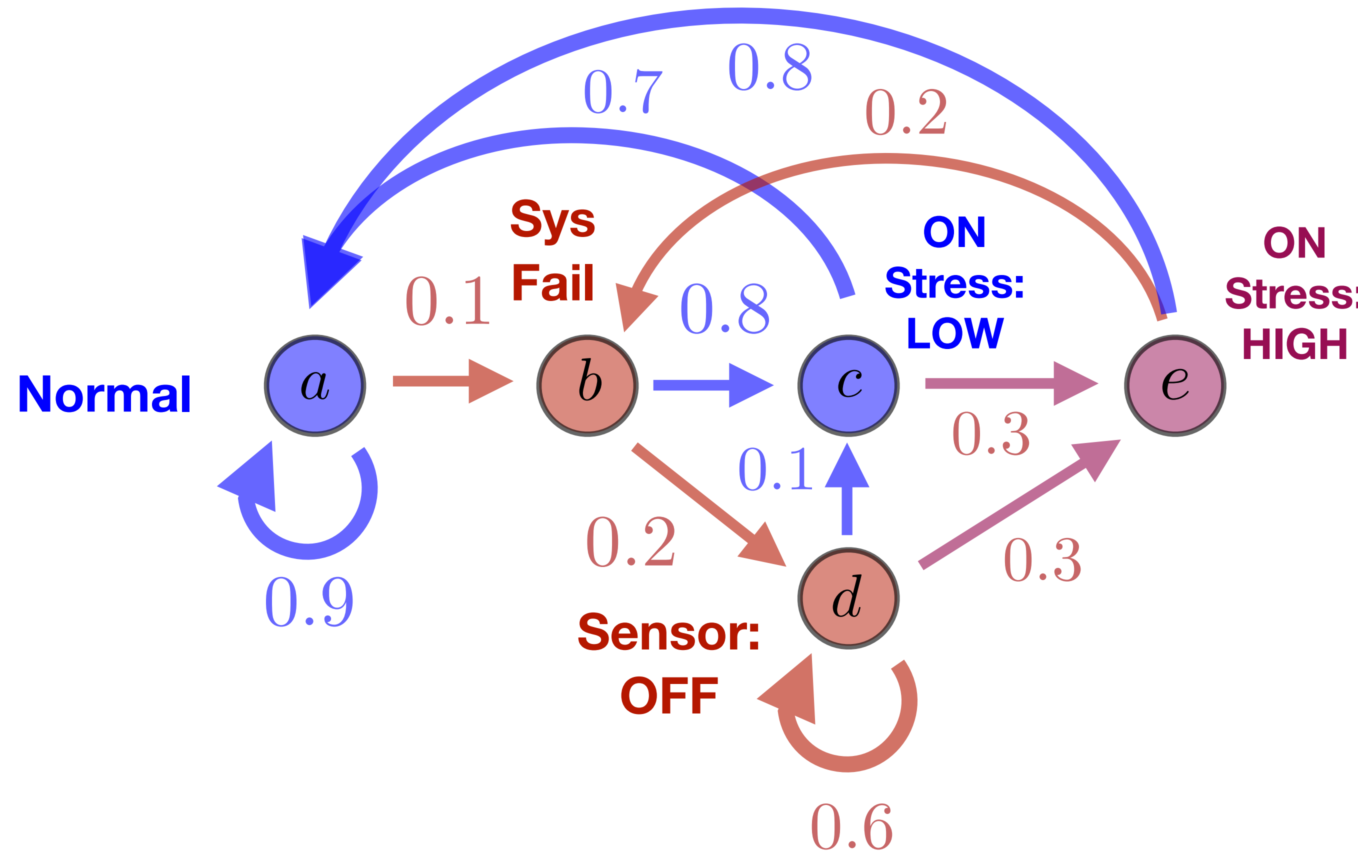


Update Equation:

$$\begin{bmatrix} p_a(k) \\ p_b(k) \\ p_c(k) \\ p_d(k) \end{bmatrix} = \begin{bmatrix} 0.9 & 0 & 1.0 & 0 \\ 0.1 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0.1 \\ 0 & 0.2 & 0 & 0.9 \end{bmatrix}^k \begin{bmatrix} p_a(0) \\ p_b(0) \\ p_c(0) \\ p_d(0) \end{bmatrix}$$

Steady State Dist.

$$\begin{bmatrix} p_a(\infty) \\ p_b(\infty) \\ p_c(\infty) \\ p_d(\infty) \end{bmatrix} = \begin{bmatrix} 0.71 \\ 0.07 \\ 0.07 \\ 0.14 \end{bmatrix}$$



Update Equation:

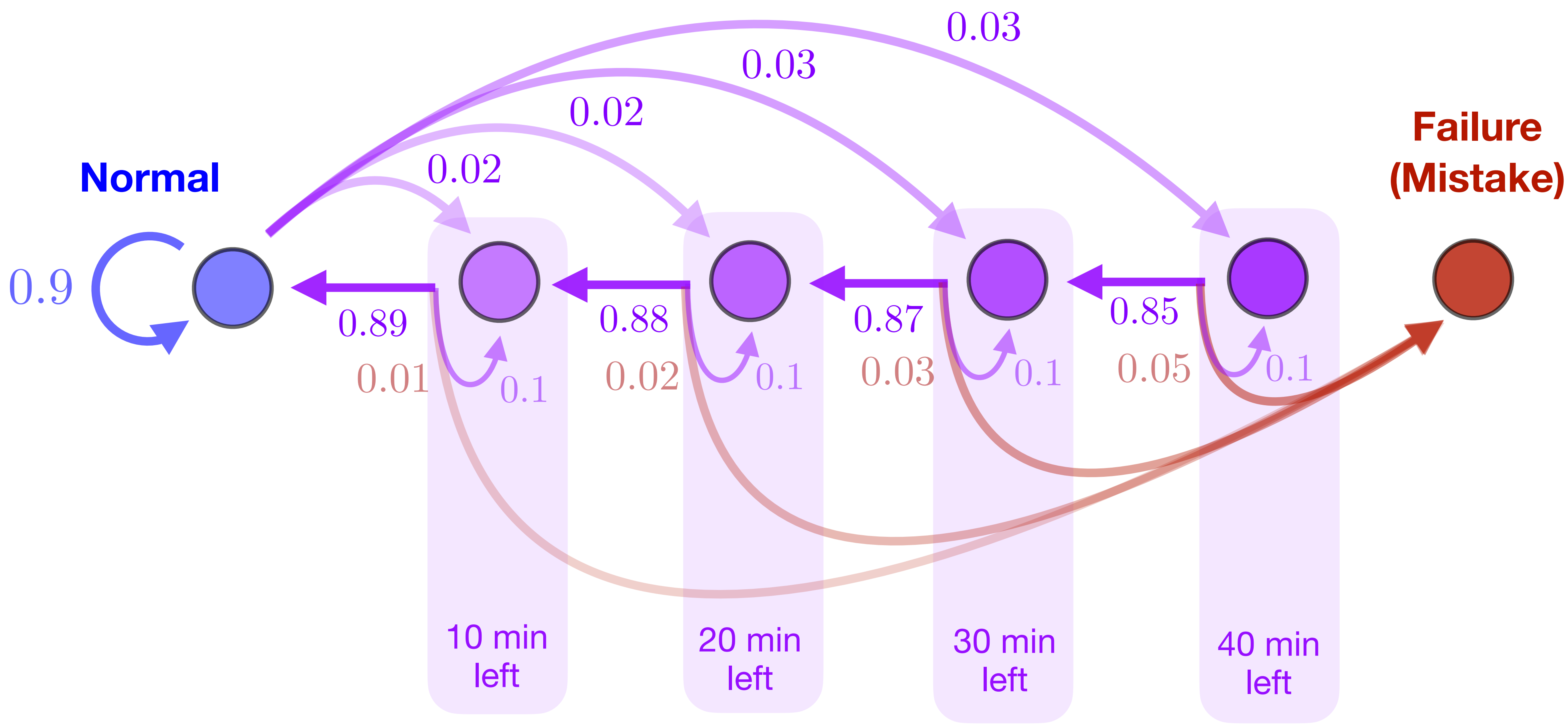
$$\begin{bmatrix} p_a(k) \\ p_b(k) \\ p_c(k) \\ p_d(k) \\ p_e(k) \end{bmatrix} = \begin{bmatrix} 0.9 & 0 & 0.7 & 0 & 0.8 \\ 0.1 & 0 & 0 & 0 & 0.2 \\ 0 & 0.8 & 0 & 0.1 & 0 \\ 0 & 0.2 & 0 & 0.6 & 0 \\ 0 & 0 & 0.3 & 0.3 & 0 \end{bmatrix}^k \begin{bmatrix} p_a(0) \\ p_b(0) \\ p_c(0) \\ p_d(0) \\ p_e(0) \end{bmatrix}$$

Steady State Dist.

$$\begin{bmatrix} p_a(\infty) \\ p_b(\infty) \\ p_c(\infty) \\ p_d(\infty) \\ p_e(\infty) \end{bmatrix} = \begin{bmatrix} 0.77 \\ 0.08 \\ 0.07 \\ 0.04 \\ 0.03 \end{bmatrix}$$

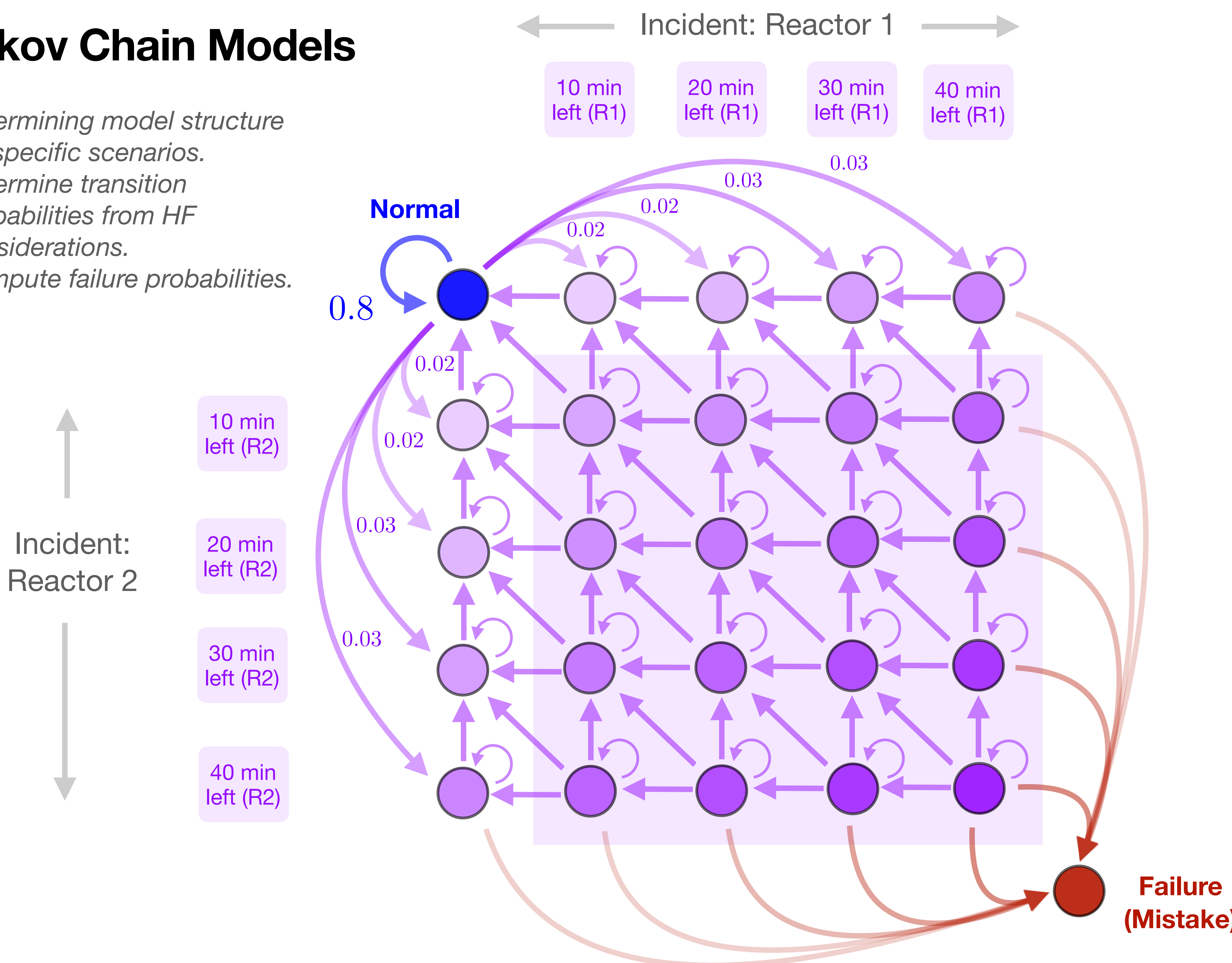
Markov Chain Models

1. *Determining model structure for specific scenarios.*
2. *Determine transition probabilities from HF considerations.*
3. *Compute failure probabilities.*



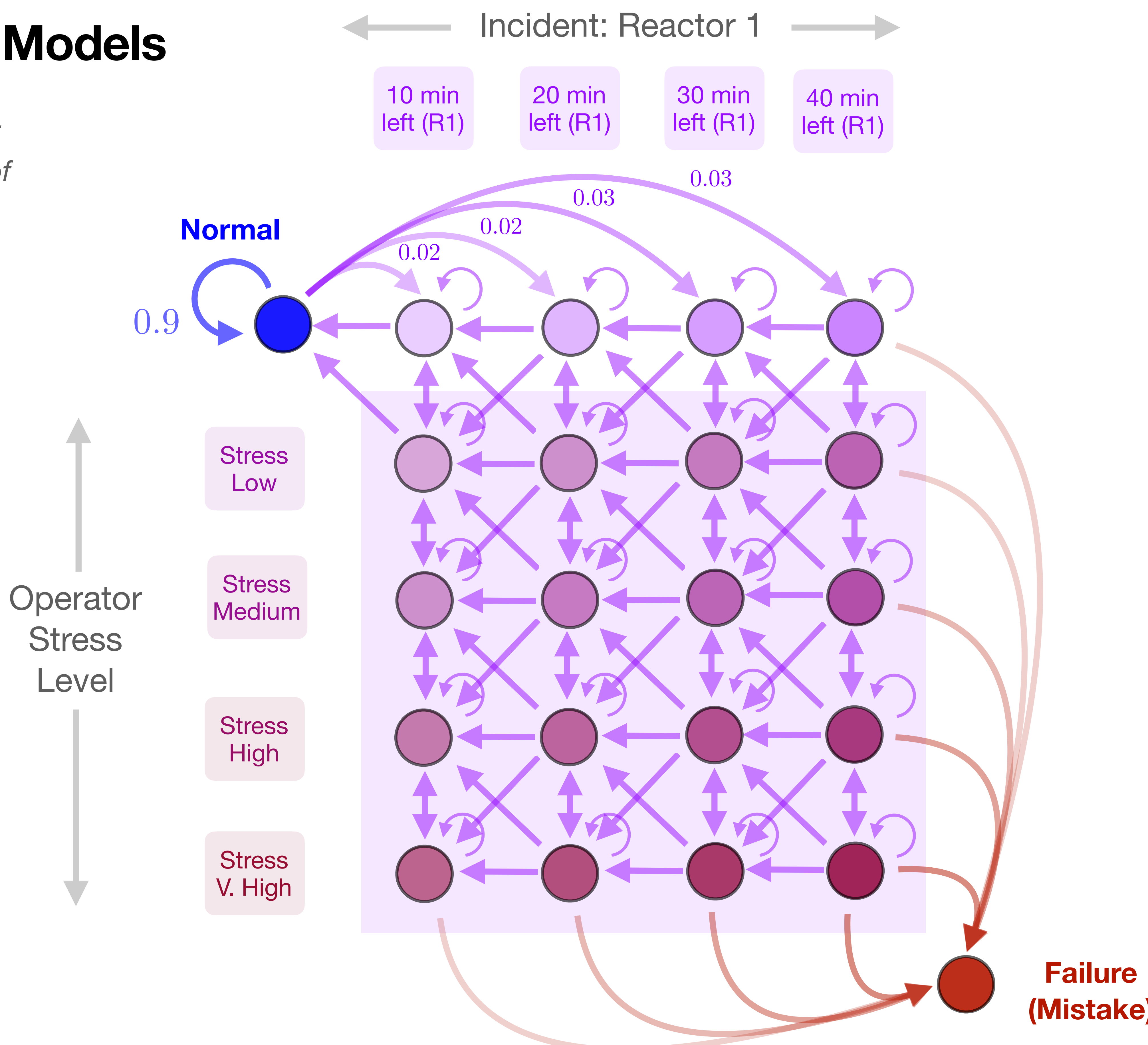
Markov Chain Models

1. *Determining model structure for specific scenarios.*
2. *Determine transition probabilities from HF considerations.*
3. *Compute failure probabilities.*



Markov Chain Models

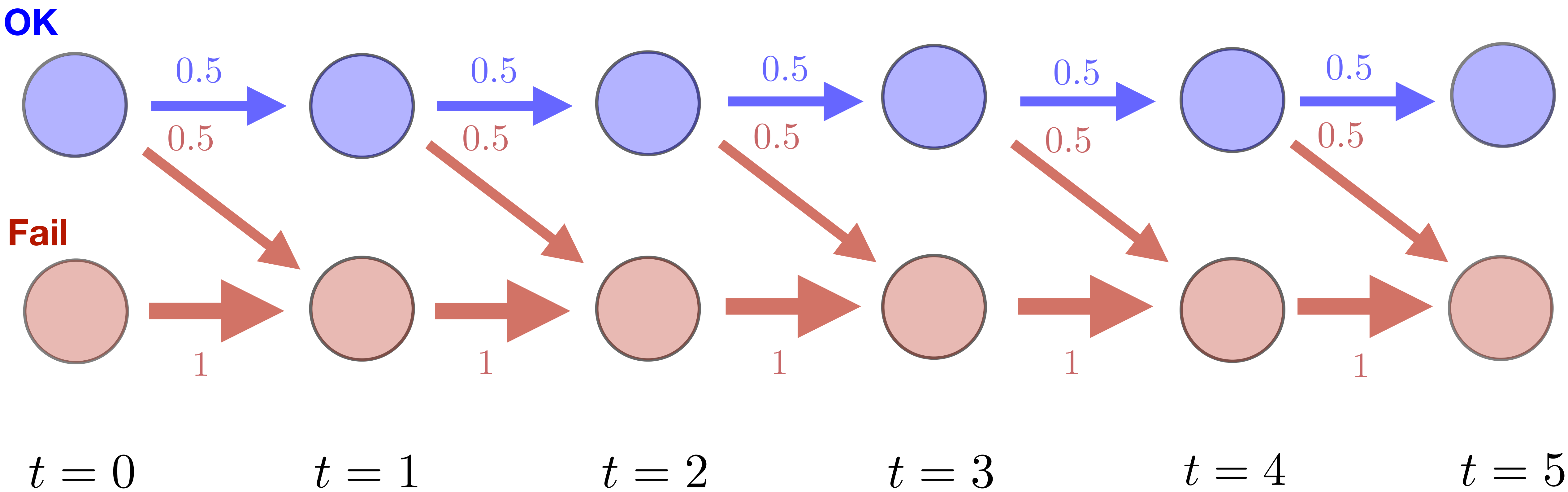
4. *Modeling different HF components as part of the state space.*



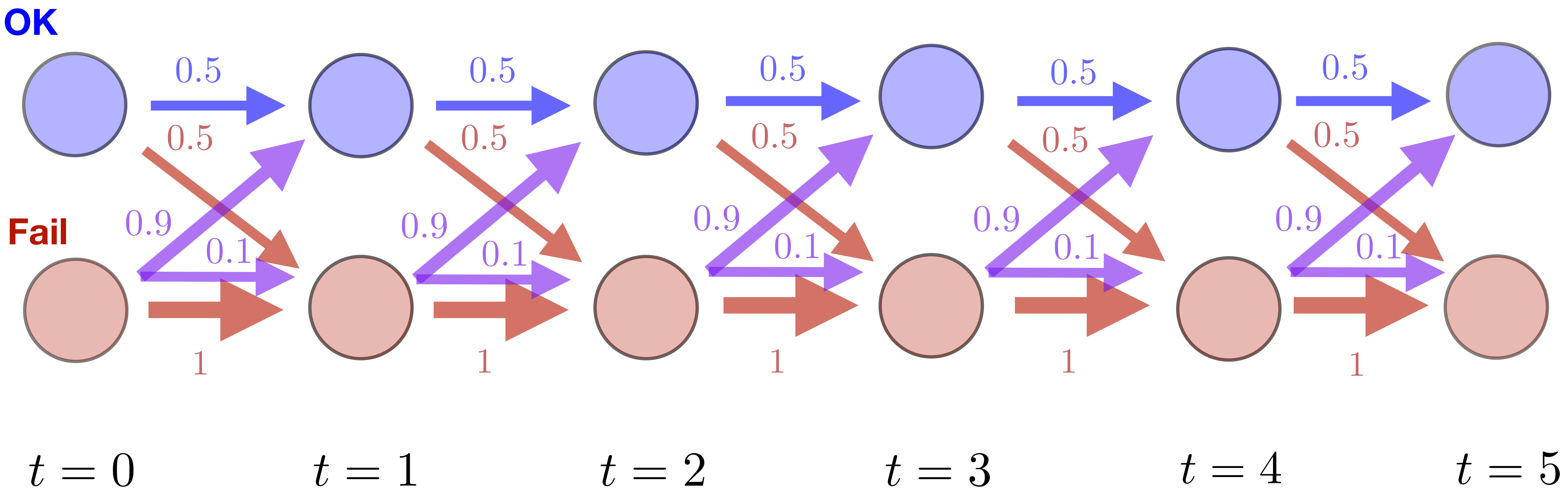
Markov Decision Process (MDP): Examples

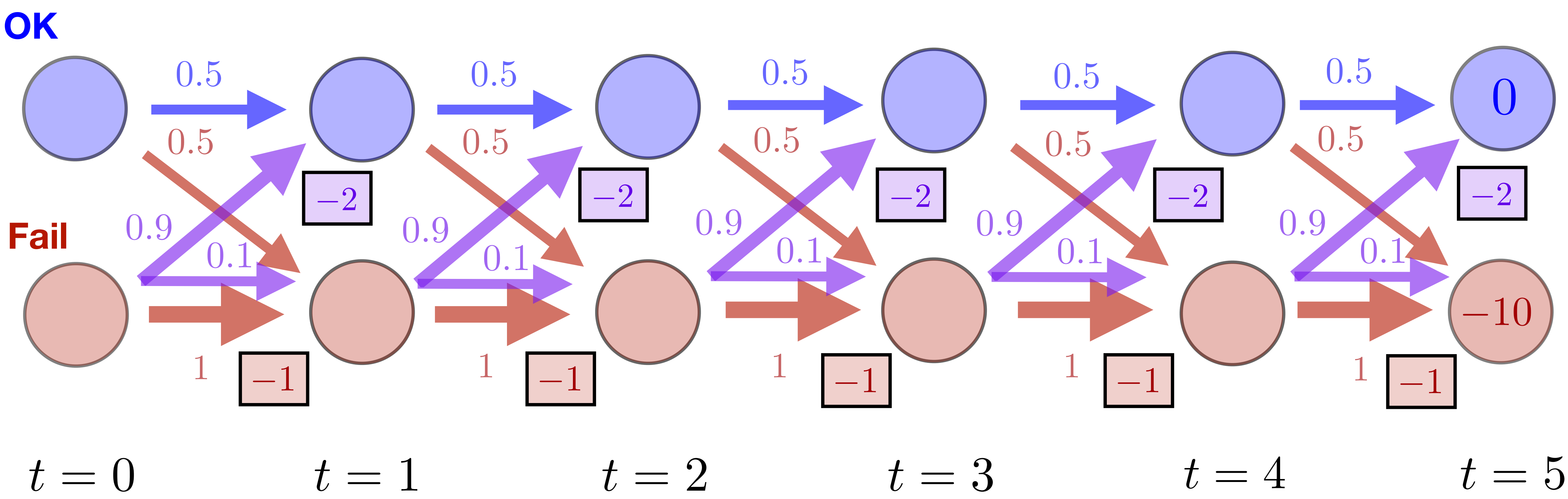
Stochastic Processes

Major sources:

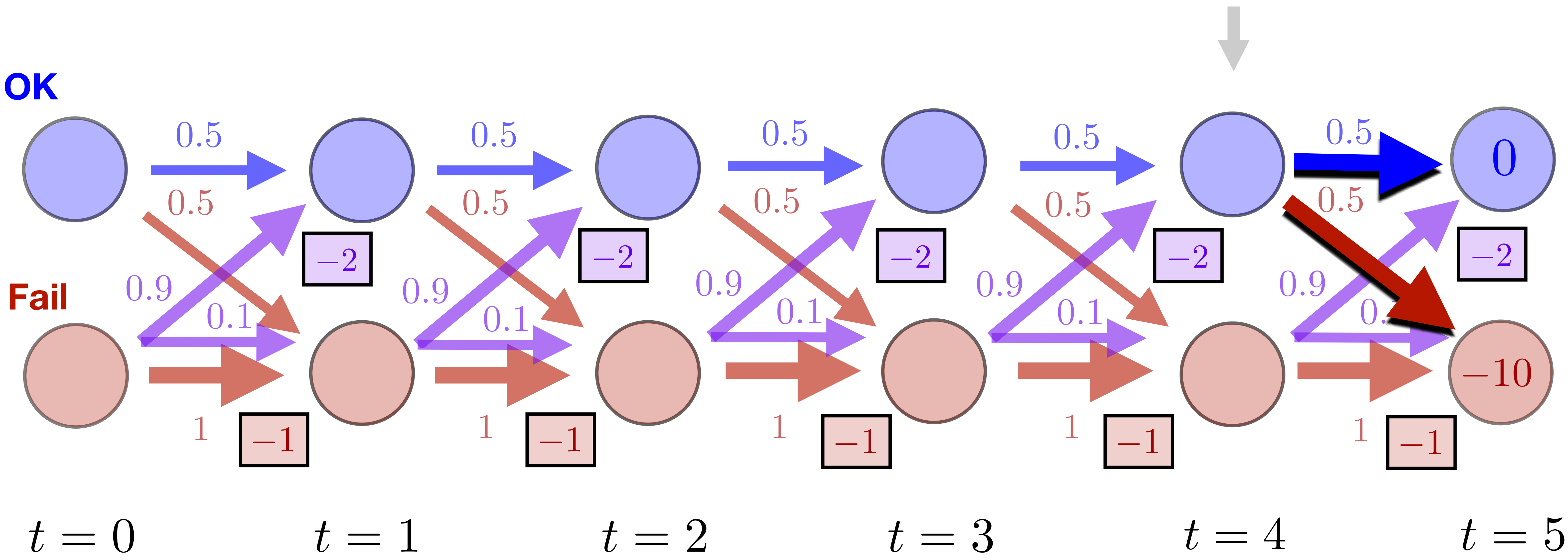


OK
Fail
0.5
0.5
1
$t = 0$
$t = 1$
$t = 2$
$t = 3$
$t = 4$
$t = 5$


OK
Fail
0.5
0.5
0.9
0.1
1
$t = 0$
$t = 1$
$t = 2$
$t = 3$
$t = 4$
$t = 5$



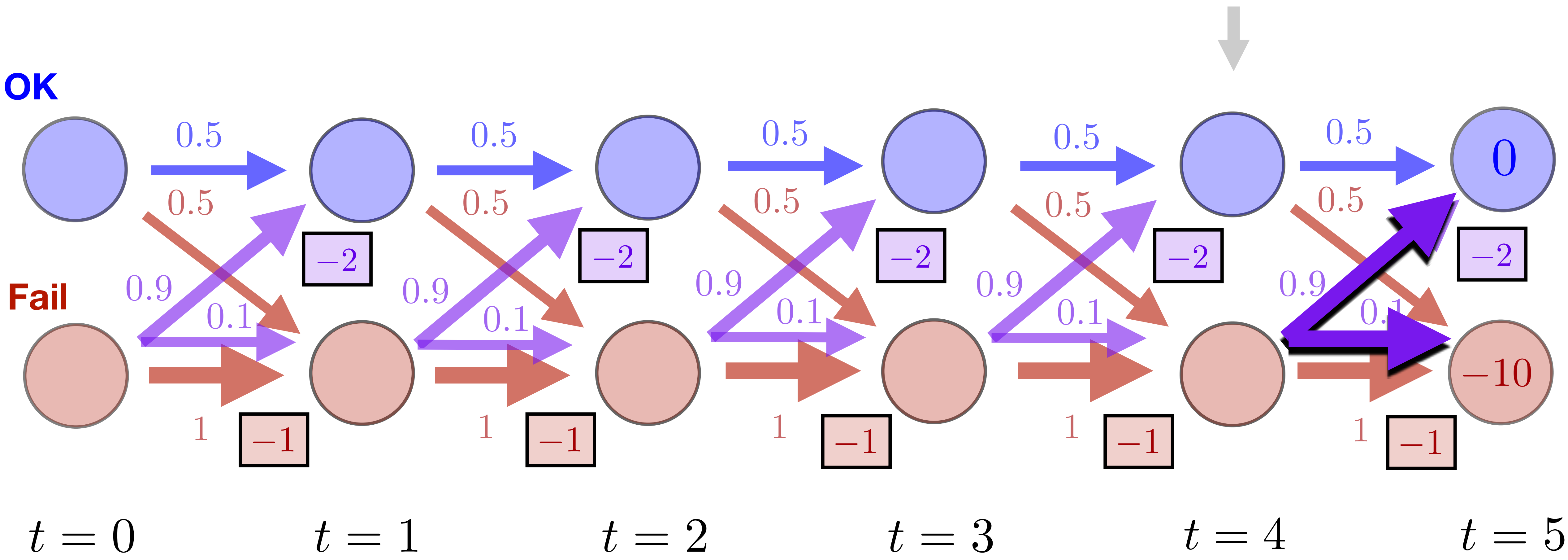
OK
Fail
0.5
0.5
0.9
0.1
1
−2
−1
0
−10
t = 0
t = 1
t = 2
t = 3
t = 4
t = 5



From State 1

action 1

$$-5 = 0 \times 0.5 + -10 \times 0.5$$

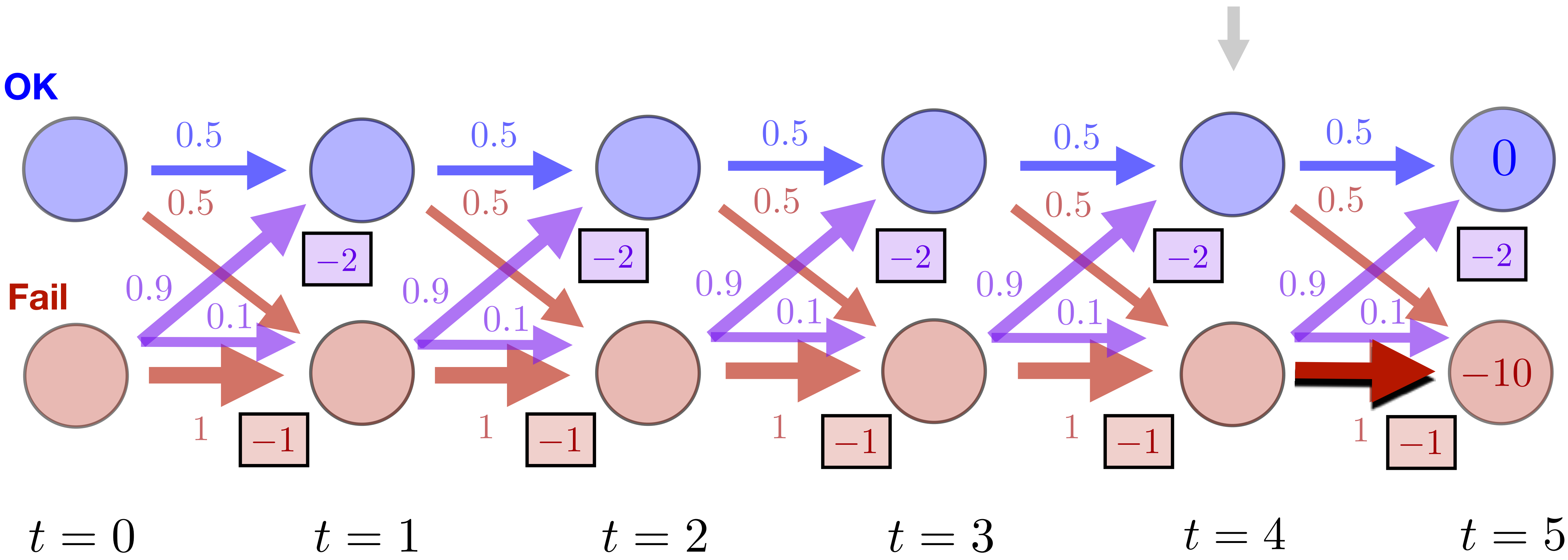


From State 1

action 1 $\quad -5 = 0 \times 0.5 + (-10) \times 0.5$

From State 2

action 1 $\quad -3 = 0 \times 0.9 + (-10) \times 0.1 + (-2)$



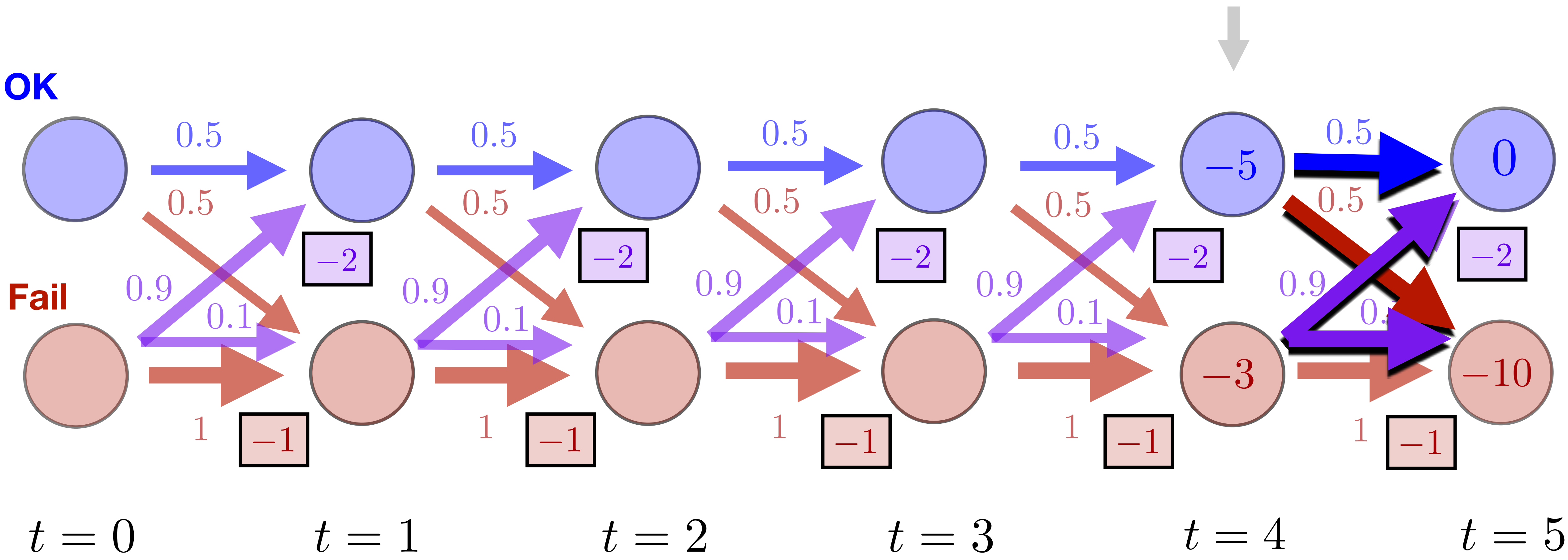
From State 1

action 1 $-5 = 0 \times 0.5 + -10 \times 0.5$

From State 2

action 1 $-3 = 0 \times 0.9 + -10 \times 0.1 + -2$

action 2 $-11 = -10 \times 1 + -1$



From State 1

action 1 (max)

$$-5 = 0 \times 0.5 + (-10) \times 0.5$$

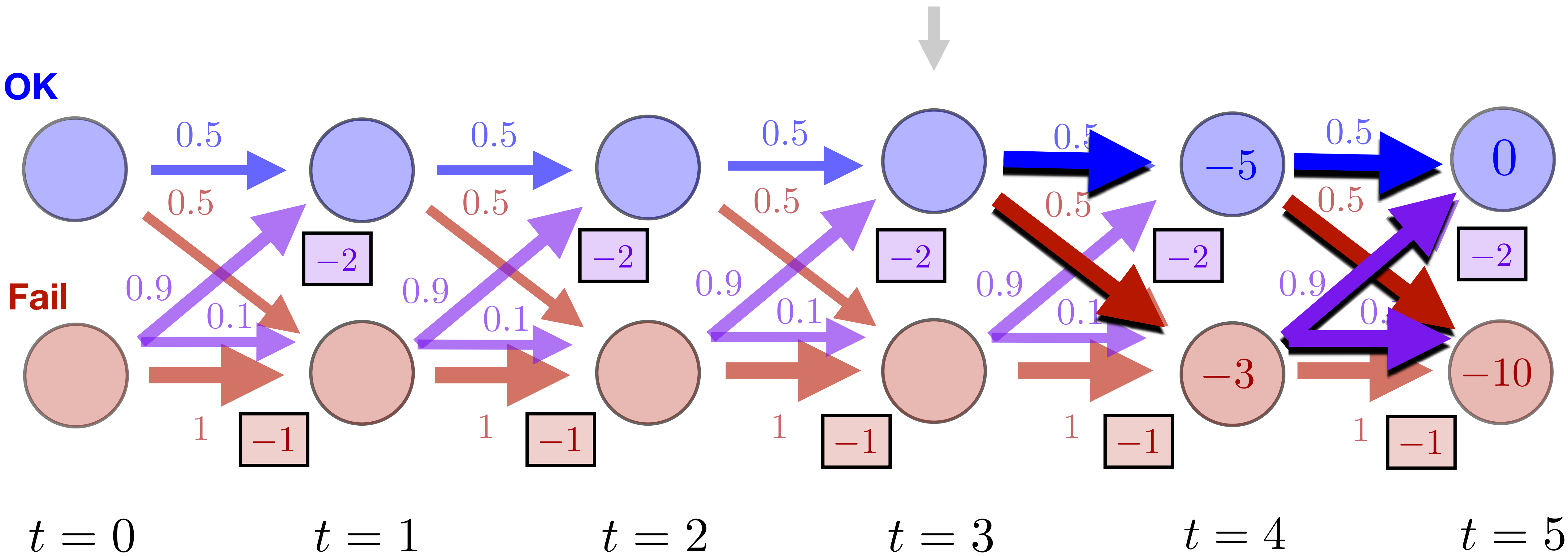
From State 2

action 1 (max)

$$-3 = 0 \times 0.9 + (-10) \times 0.1 + (-2)$$

action 2

$$-11 = (-10) \times 1 + (-1)$$



From State 1

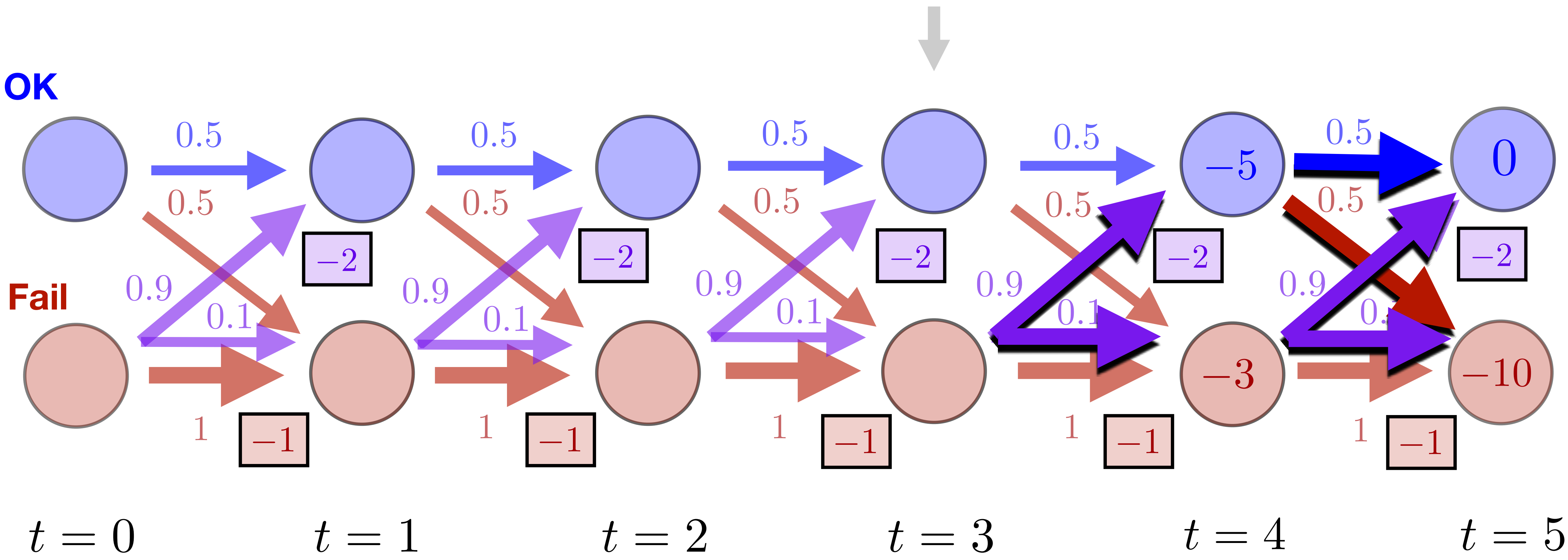
action 1

$$-4 = -5 \times 0.5 + -3 \times 0.5$$

From State 2

action 1

action 2



From State 1

action 1

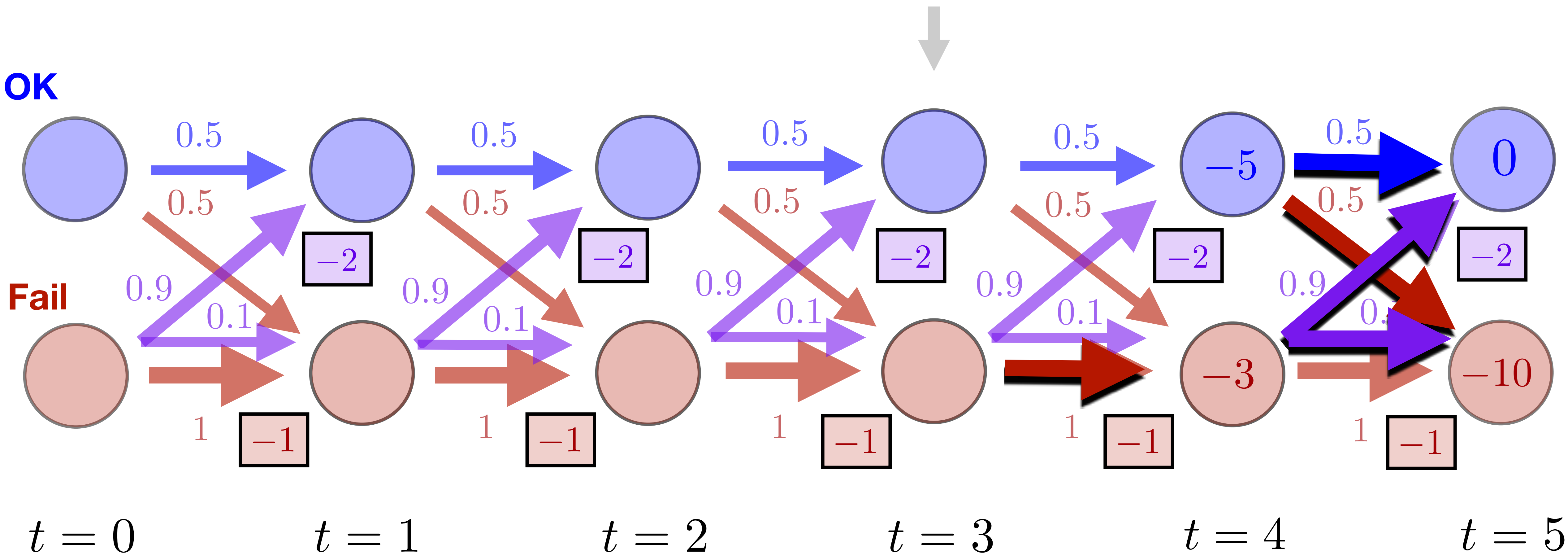
$$-4 = -5 \times 0.5 + -3 \times 0.5$$

From State 2

action 1

$$-6.8 = -5 \times 0.9 + -3 \times 0.1 + -2$$

action 2



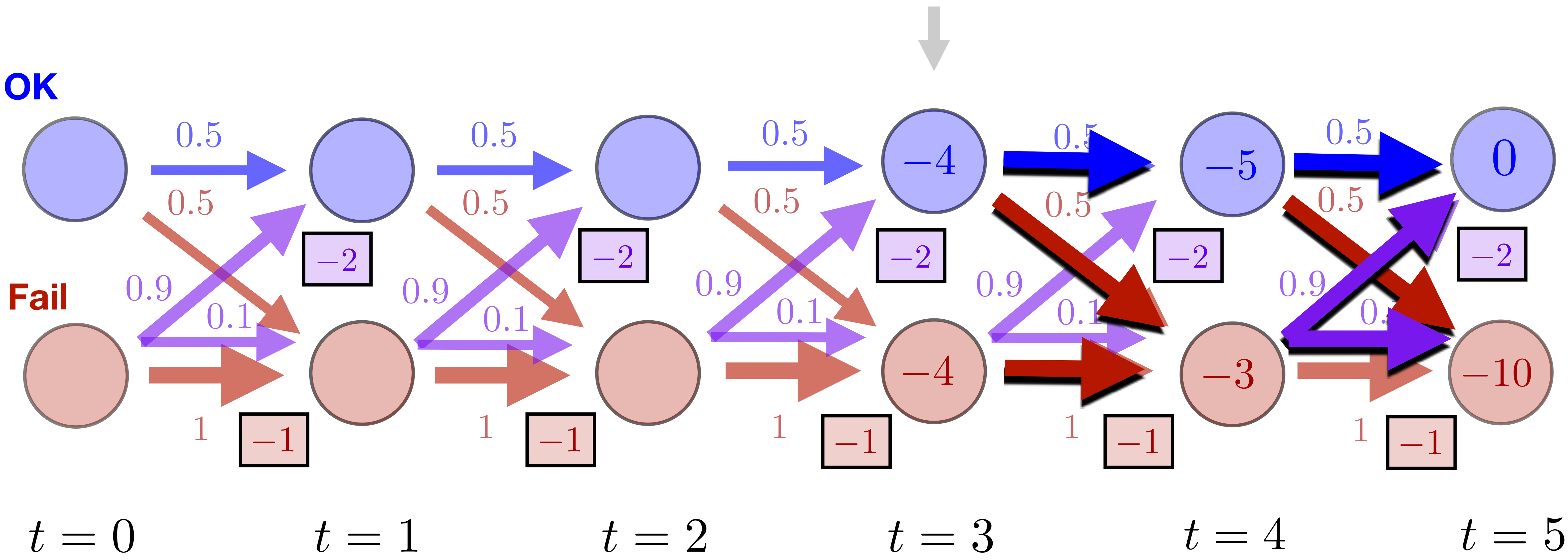
From State 1

action 1 $\quad -4 = -5 \times 0.5 + -3 \times 0.5$

From State 2

action 1 $\quad -6.8 = -5 \times 0.9 + -3 \times 0.1 + -2$

action 2 $\quad -4 = -3 \times 1 + -1$



max

From State 1

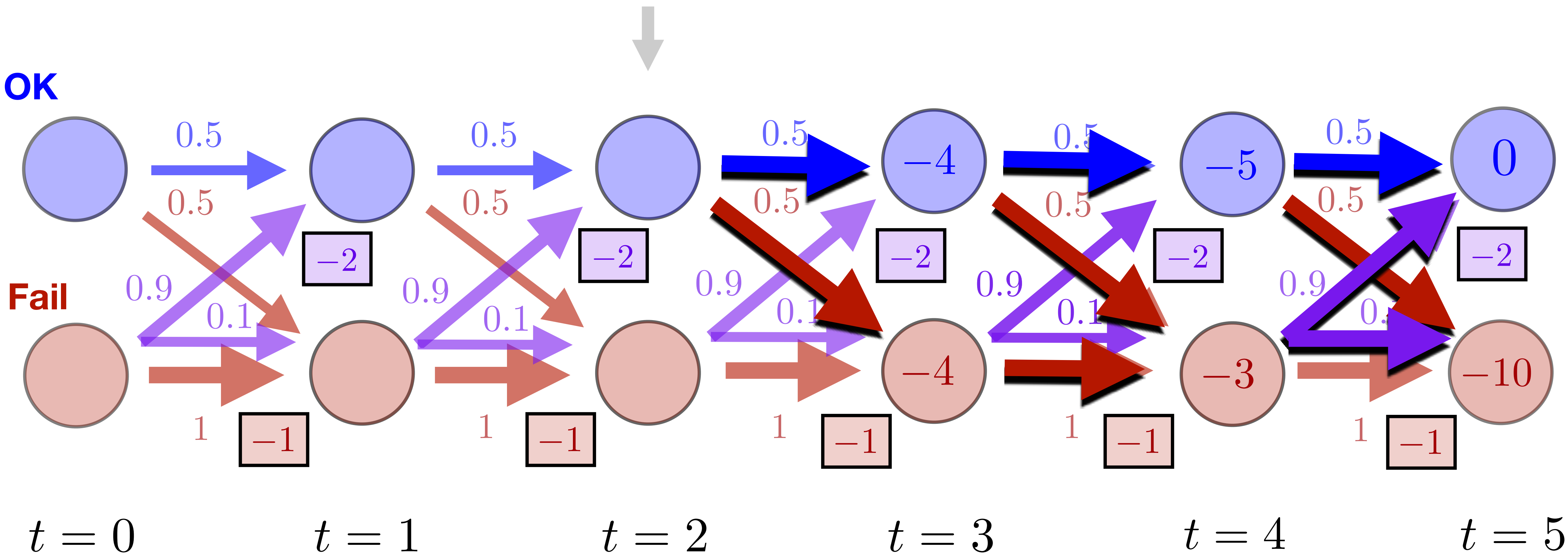
action 1 $-4 = -5 \times 0.5 + -3 \times 0.5$

From State 2

action 1 $-6.8 = -5 \times 0.9 + -3 \times 0.1 + -2$

max

action 2 $-4 = -3 \times 1 + -1$



From State 1

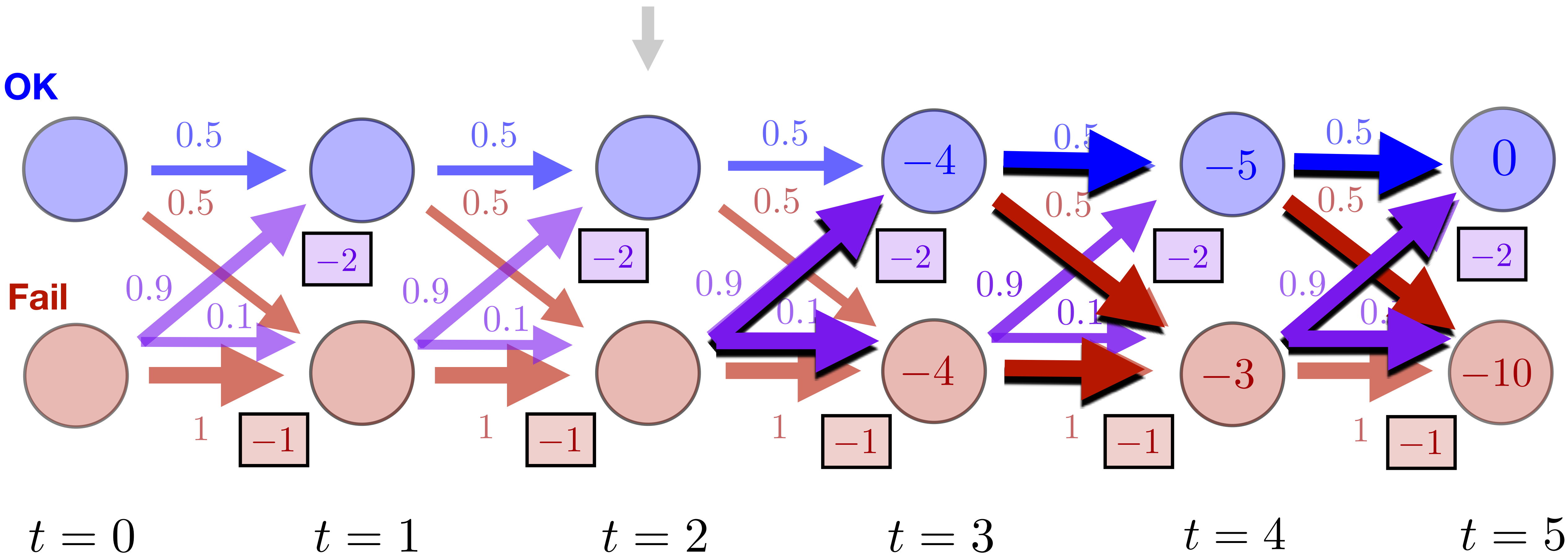
action 1

$$-4 = -4 \times 0.5 + -4 \times 0.5$$

From State 2

action 1

action 2



From State 1

action 1

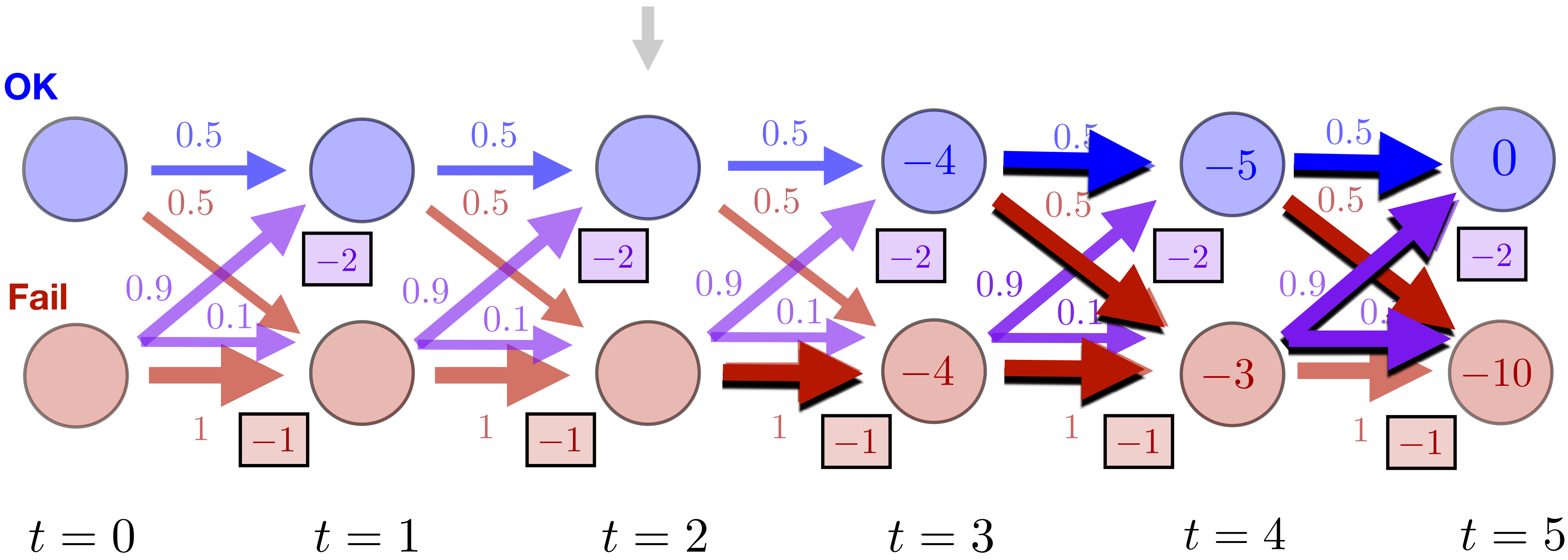
$$-4 = (-4) \times 0.5 + (-4) \times 0.5$$

From State 2

action 1

$$-6 = (-4) \times 0.9 + (-4) \times 0.1 + (-2)$$

action 2



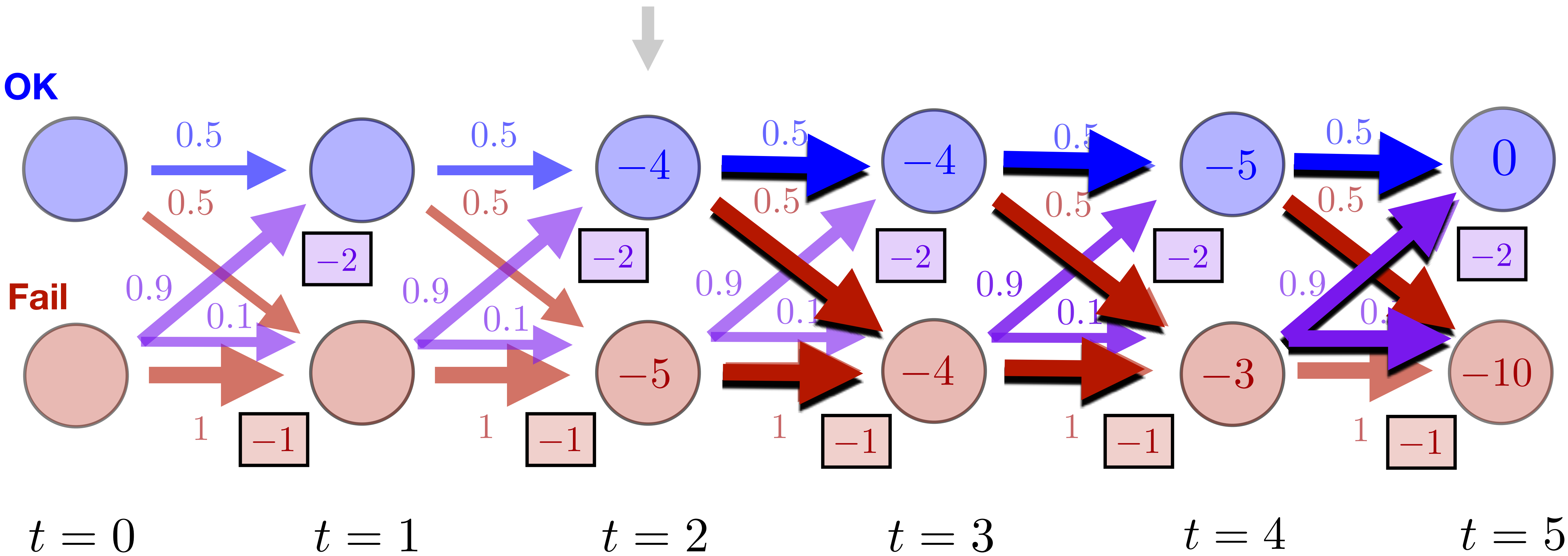
From State 1

action 1 $\quad -4 = -4 \times 0.5 + -4 \times 0.5$

From State 2

action 1 $\quad -6 = -4 \times 0.9 + -4 \times 0.1 + -2$

action 2 $\quad -5 = -4 \times 1 + -1$



From State 1

max

action 1

$$-4 = -4 \times 0.5 + -4 \times 0.5$$

From State 2

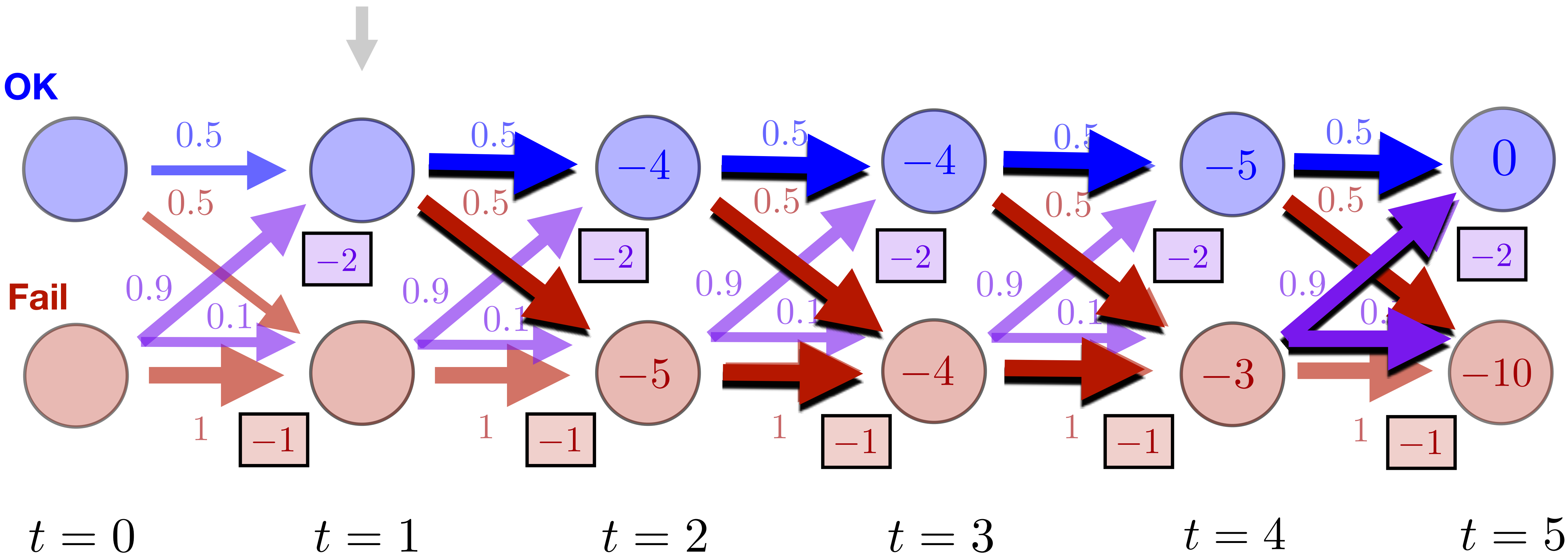
action 1

$$-6 = -4 \times 0.9 + -4 \times 0.1 + -2$$

max

action 2

$$-5 = -4 \times 1 + -1$$



From State 1

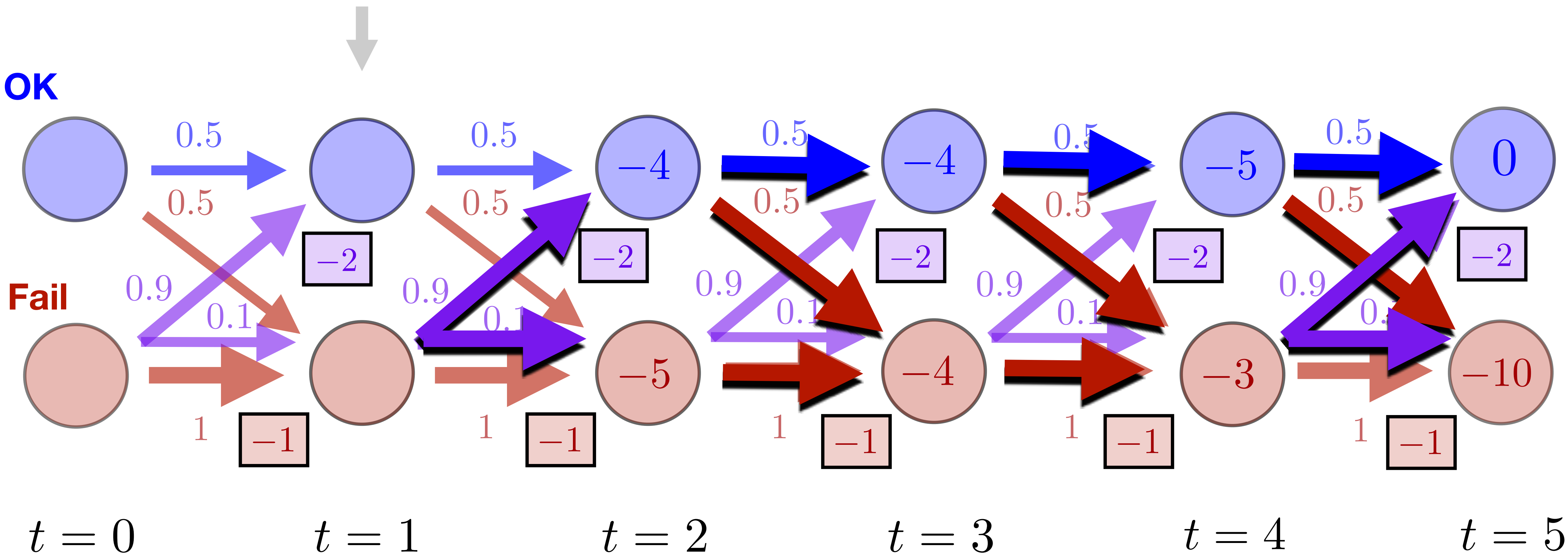
action 1

$$-4.5 = -4 \times 0.5 + -5 \times 0.5$$

From State 2

action 1

action 2



From State 1

action 1

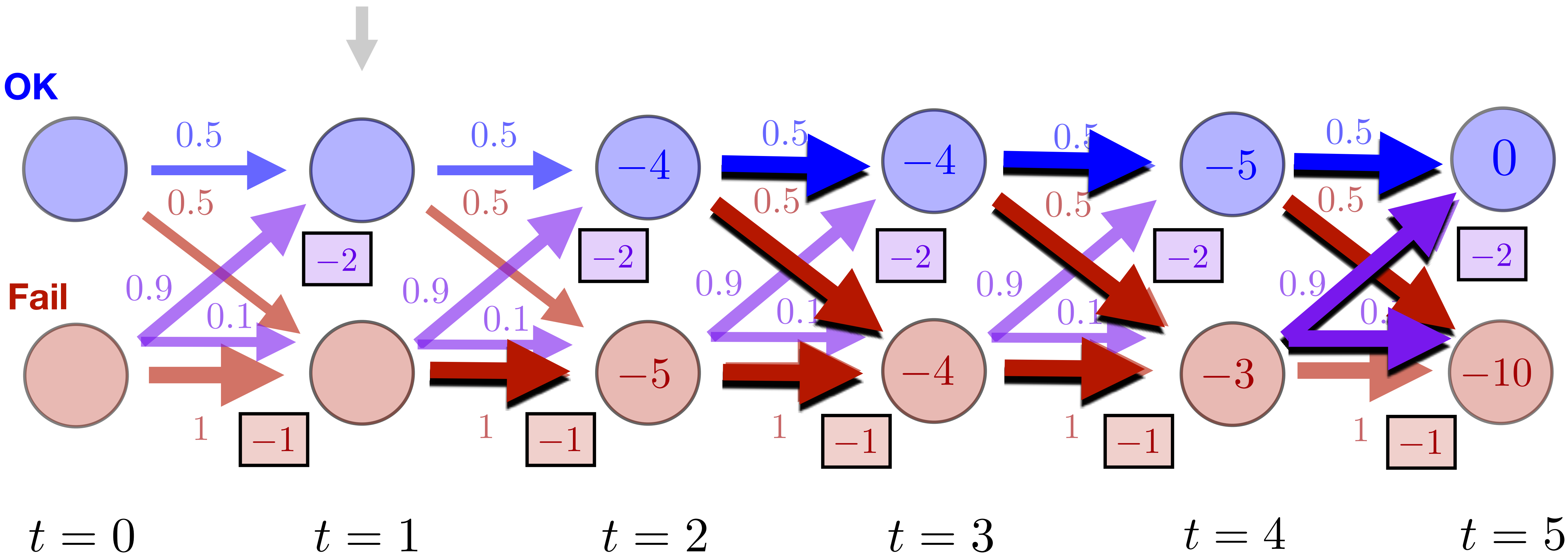
$$-4.5 = -4 \times 0.5 + -5 \times 0.5$$

From State 2

action 1

$$-6.1 = -4 \times 0.9 + -5 \times 0.1 + -2$$

action 2



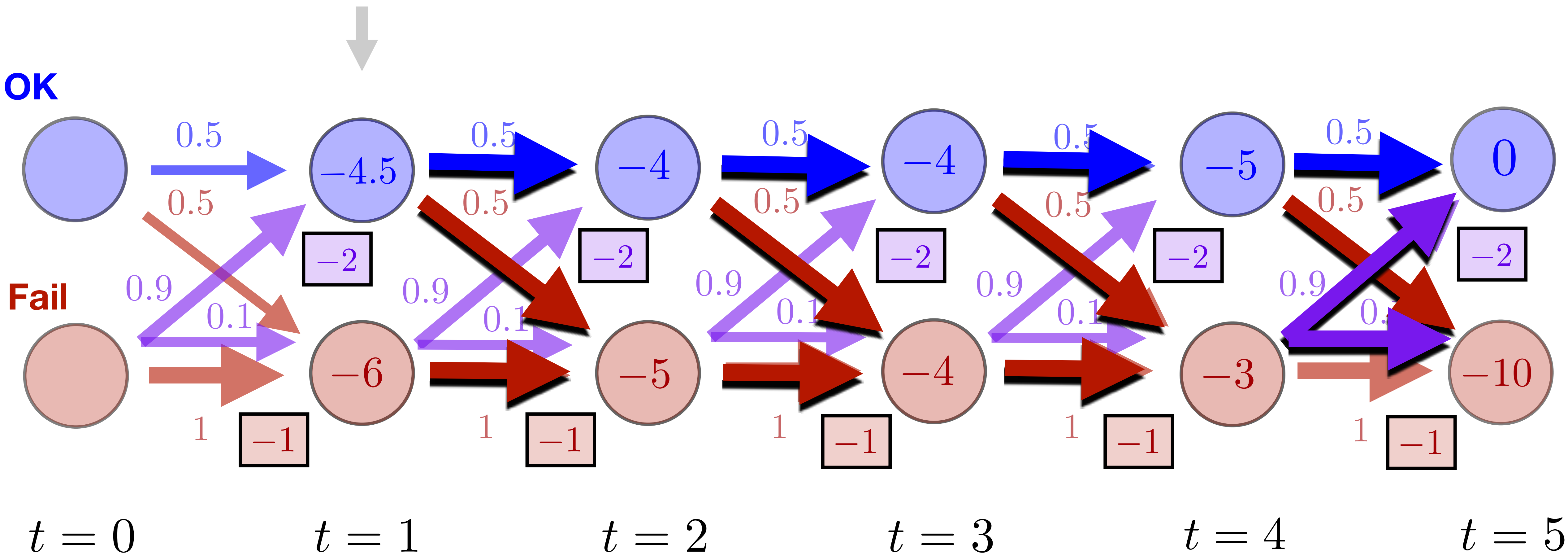
From State 1

action 1 $\quad -4.5 = -4 \times 0.5 + -5 \times 0.5$

From State 2

action 1 $\quad -6.1 = -4 \times 0.9 + -5 \times 0.1 + -2$

action 2 $\quad -6 = -5 \times 1 + -1$



From State 1

action 1

max

$$-4.5 = -4 \times 0.5 + -5 \times 0.5$$

From State 2

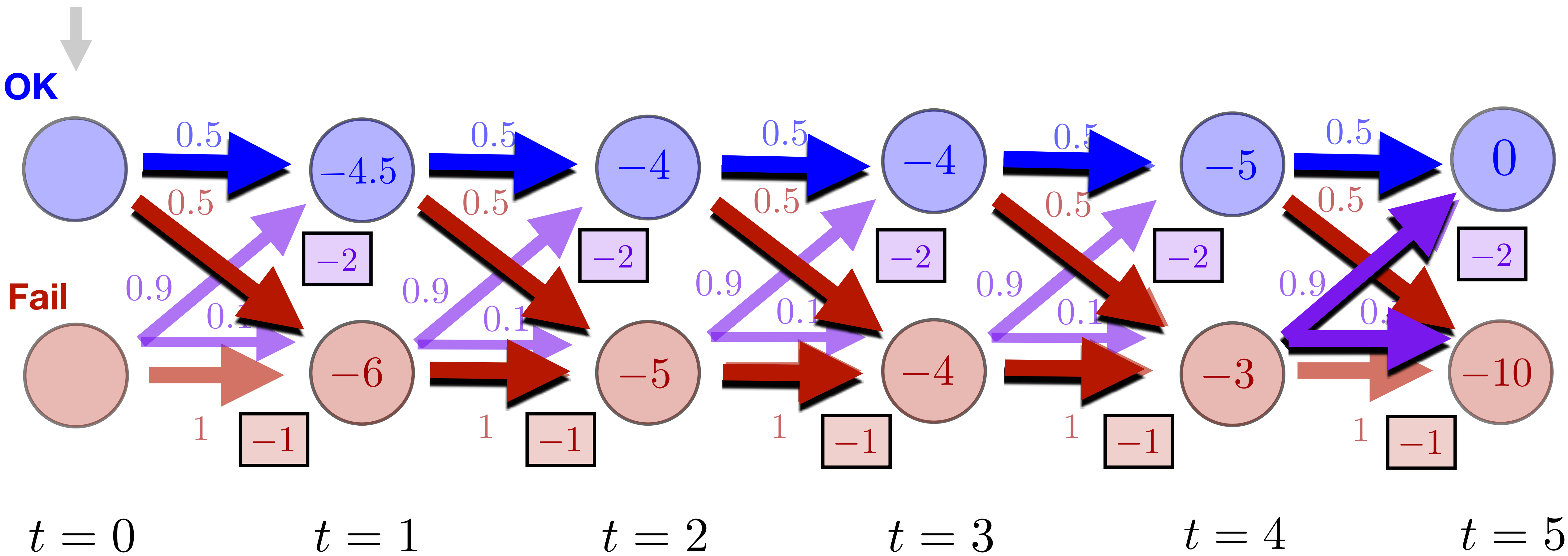
action 1

$$-6.1 = -4 \times 0.9 + -5 \times 0.1 + -2$$

max

action 2

$$-6 = -5 \times 1 + -1$$



From State 1

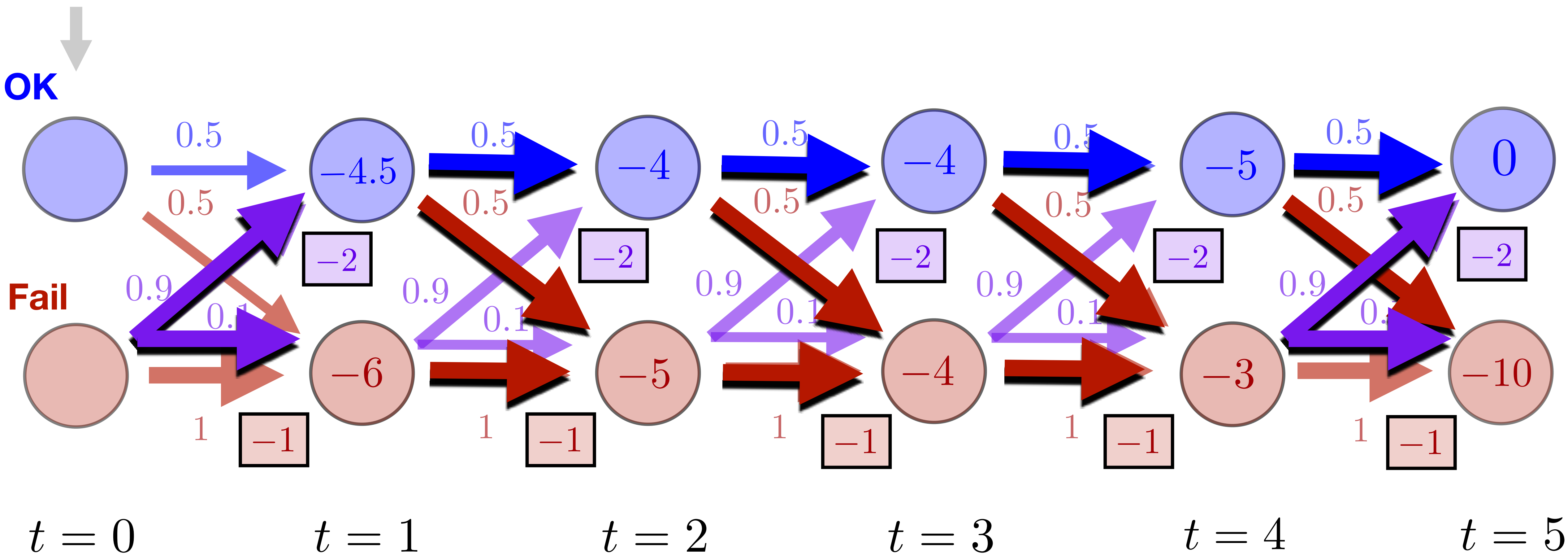
action 1

$$-5.3 = -4.5 \times 0.5 + -6 \times 0.5$$

From State 2

action 1

action 2



From State 1

action 1

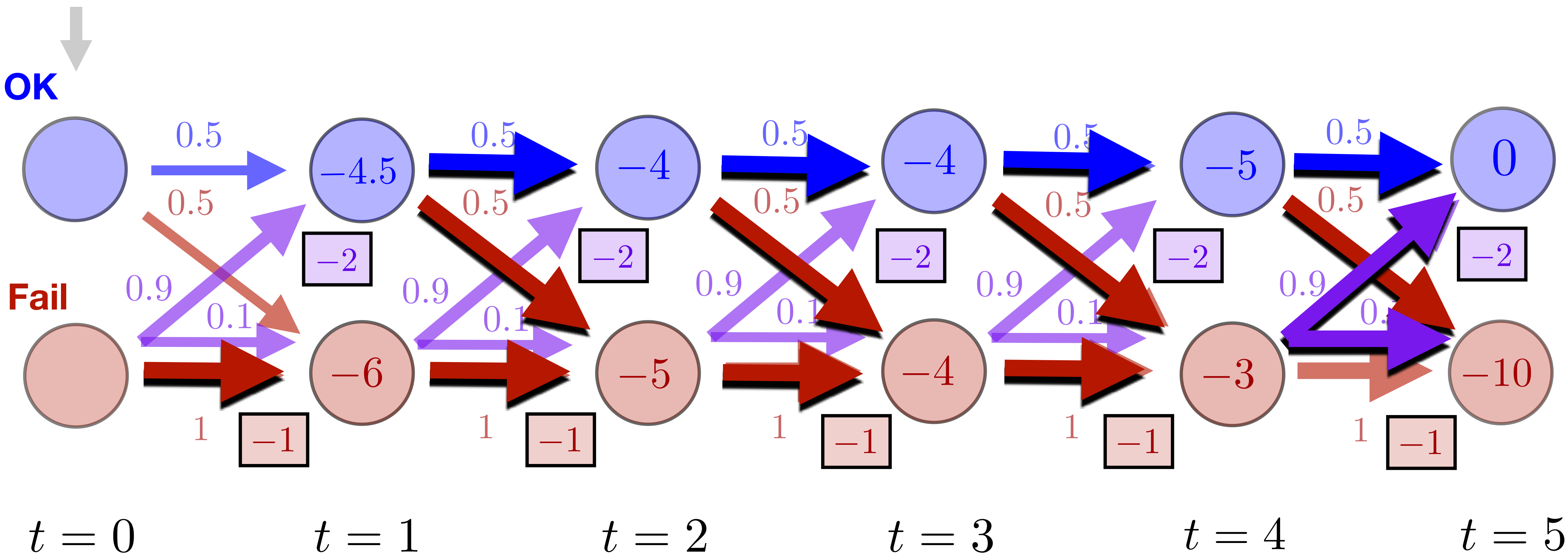
$$-5.3 = -4.5 \times 0.5 + -6 \times 0.5$$

From State 2

action 1

$$-6.7 = -4.5 \times 0.9 + -6 \times 0.1 + -2$$

action 2



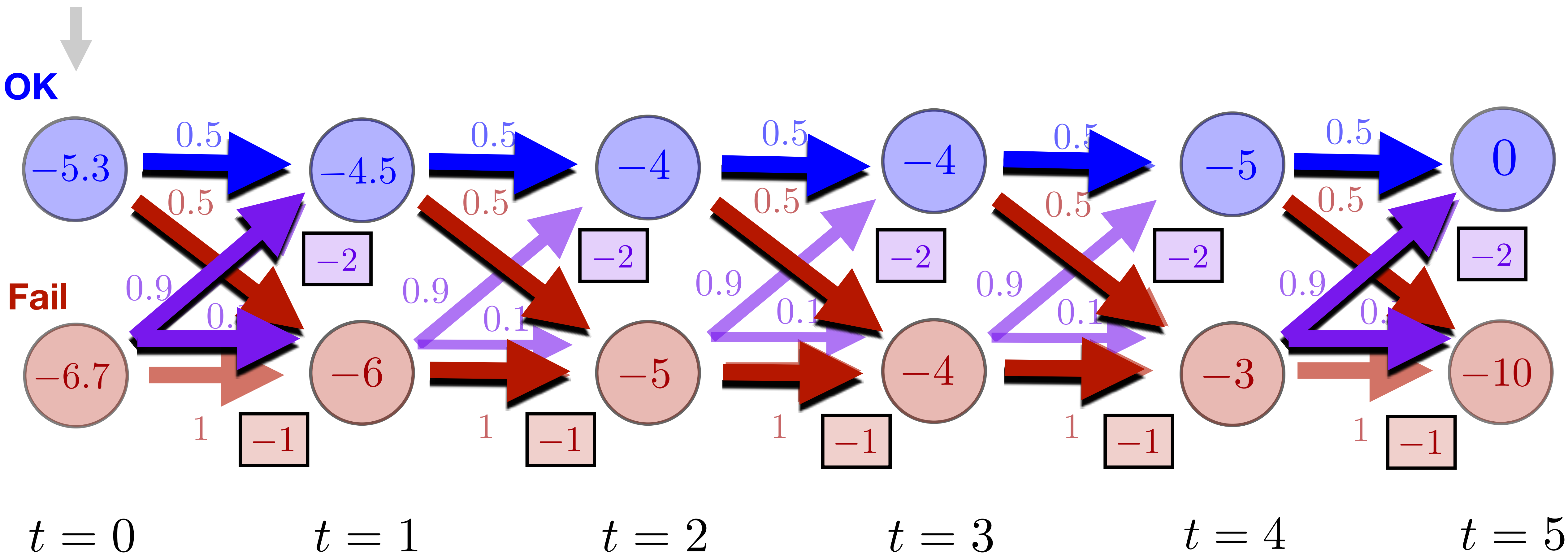
From State 1

action 1 $$-5.3 = -4.5 \times 0.5 + -6 \times 0.5$$

From State 2

action 1 $$-6.7 = -4.5 \times 0.9 + -6 \times 0.1 + -2$$

action 2 $$-7 = -6 \times 1 + -1$$



From State 1

max

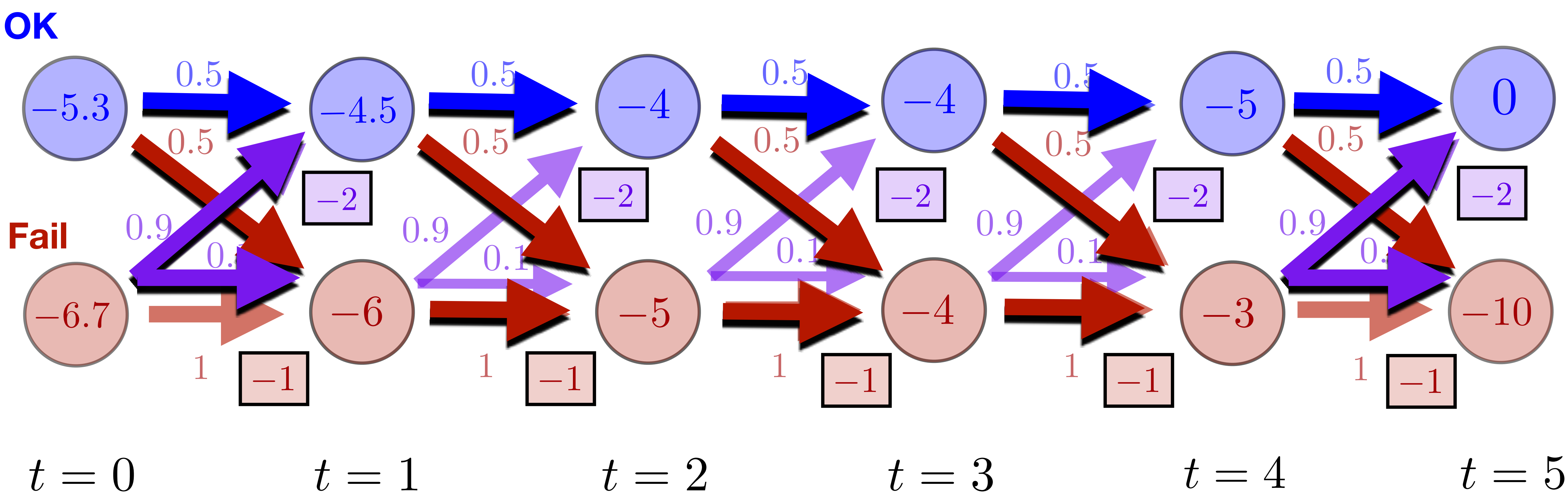
action 1 $$-5.3 = -4.5 \times 0.5 + -6 \times 0.5$$

From State 2

max

action 1 $$-6.7 = -4.5 \times 0.9 + -6 \times 0.1 + -2$$

action 2 $$-7 = -6 \times 1 + -1$$



Action-Value

from state s, action a

$$q_{sa}(k-1) = \sum_{s'} v_{s'}(k)\, P(s'|s,a) + r_{sa}(k)$$

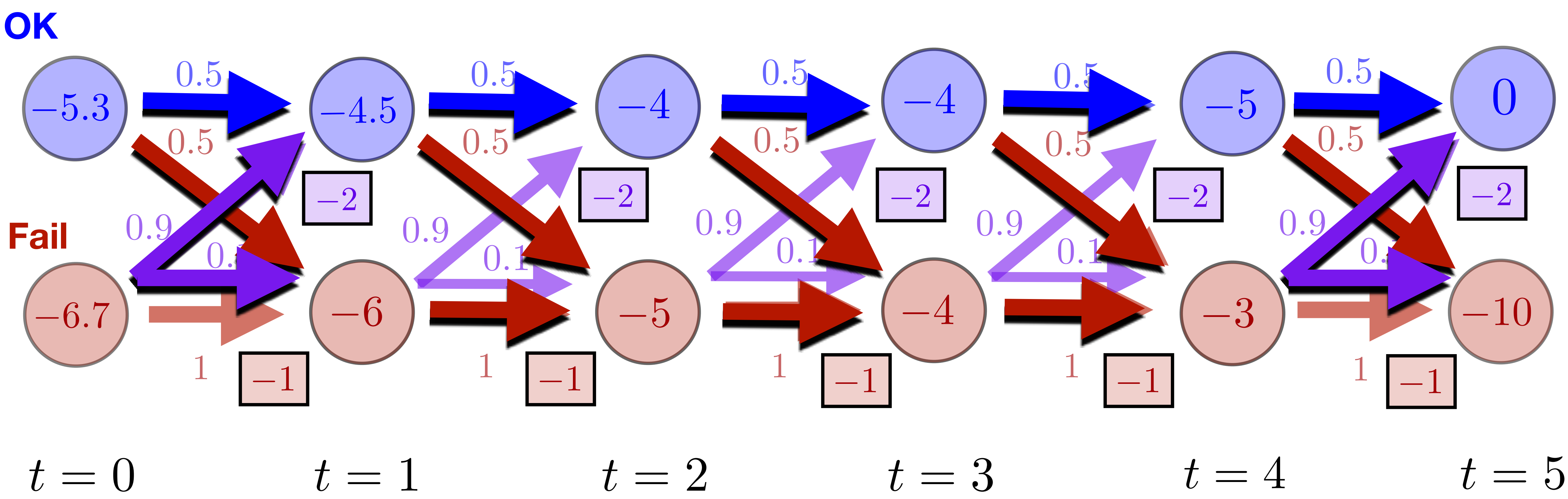
"q-value" general form

From State 2

action 1

$$-6.7 = -4.5 \times 0.9 + -6 \times 0.1 + -2$$

example



Action-Value

from state s, action a

$$q_{sa}(k-1) = \sum_{s'} v_{s'}(k)\, P(s'|s,a) + r_{sa}(k)$$

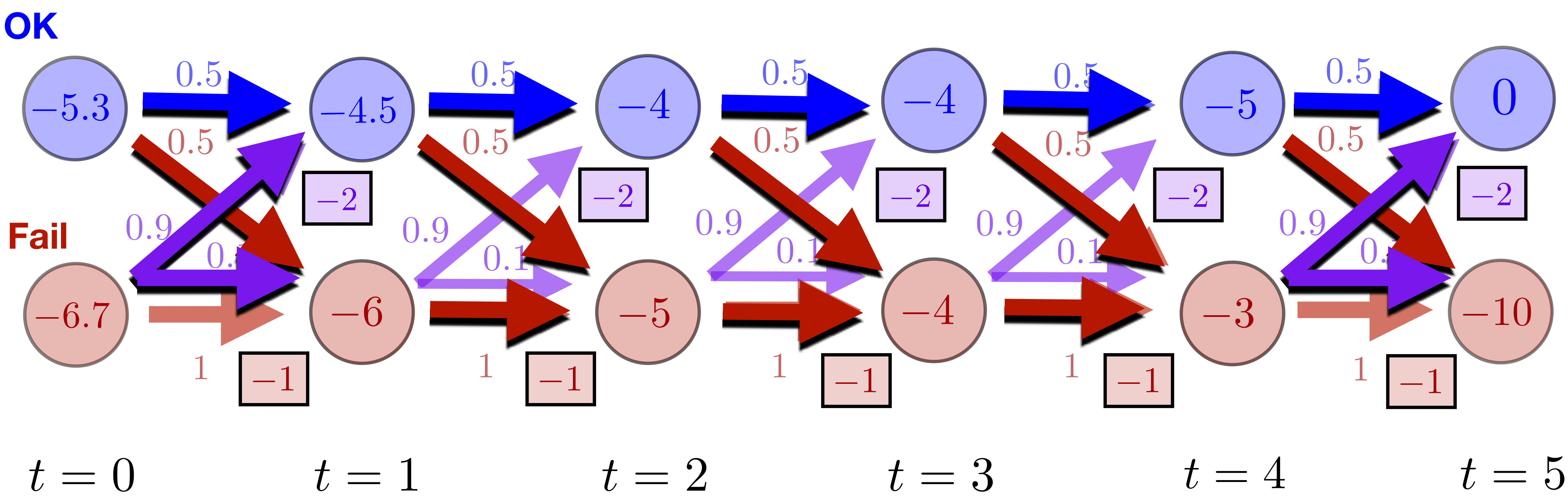
"q-value" general form

State-Value

for each state s

$$v_s(k-1) = \max_{a \in \mathcal{A}_s} q_{sa}(k-1)$$

"Reward-to-go"



Action-Value

from state s, action a

$$q_{sa}(k-1) = \sum_{s'} v_{s'}(k)\, P(s'|s,a) + r_{sa}(k)$$

"q-value" general form

State-Value

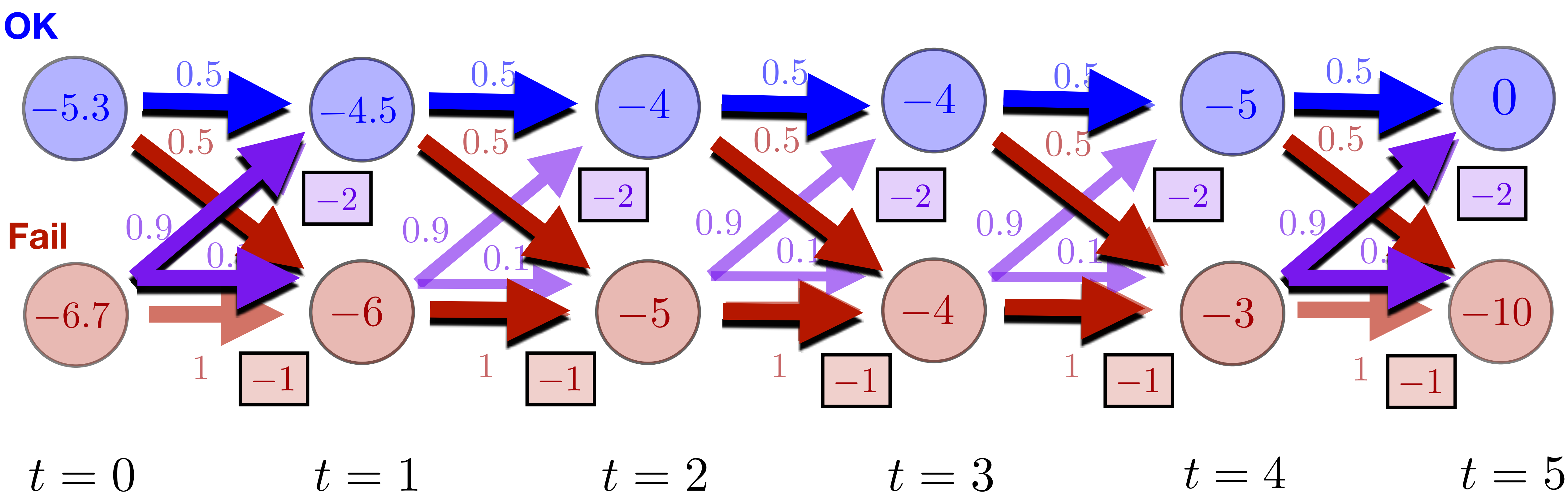
for each state s

$$v_s(k-1) = \max_{a \in \mathcal{A}_s} q_{sa}(k-1)$$

"Reward-to-go"

Bellman Equation
Dynamic Programming

$$= \max_{a \in \mathcal{A}_s} \sum_{s'} v_{s'}(k)\, P(s'|s,a) + r_{sa}(k)$$



State-Value Meaning

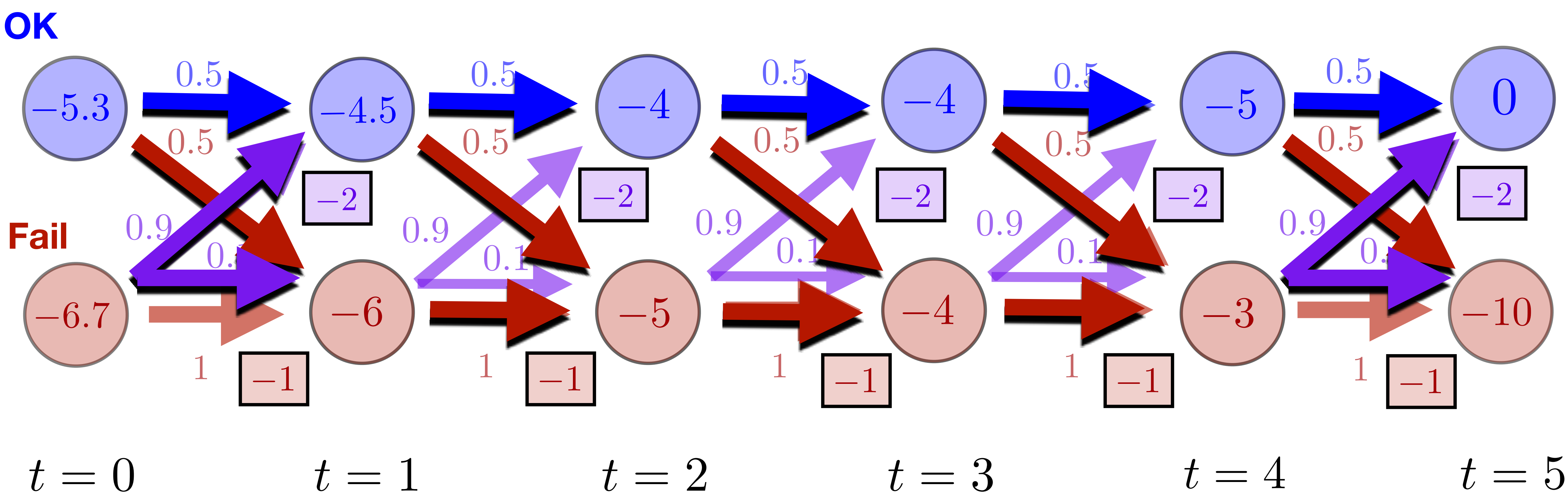
from state s $\quad v_s(k) =$ *expected reward till the end of the time horizon if you start in state s at time k and use optimal policy*

State-Value

for each state s $\quad v_s(k-1) = \max_{a \in \mathcal{A}_s} q_{sa}(k-1)$ *"Reward-to-go"*

Bellman Equation
Dynamic Programming

$$= \max_{a \in \mathcal{A}_s} \sum_{s'} v_{s'}(k) \, P(s'|s,a) + r_{sa}(k)$$



State-Value Meaning

from state s $\quad v_s(0) =$ *expected reward if you start in state s (at time 0) and use the optimal policy*

State-Value

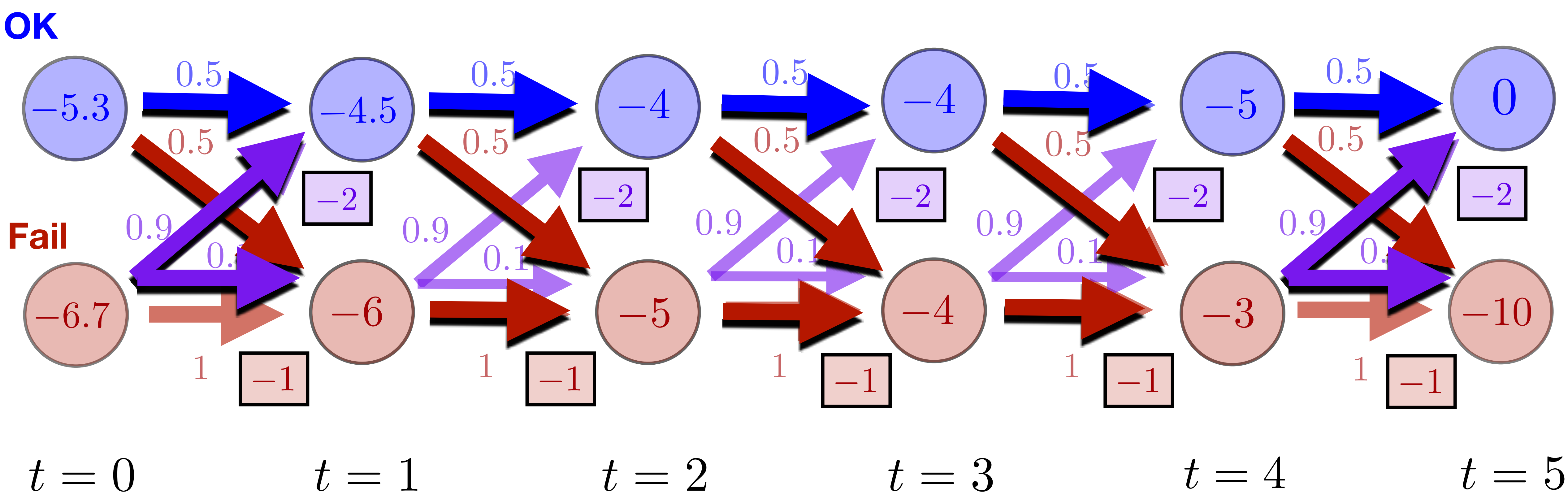
for each state s

$$v_s(k-1) = \max_{a \in \mathcal{A}_s} q_{sa}(k-1)$$

"Reward-to-go"

Bellman Equation
Dynamic Programming

$$= \max_{a \in \mathcal{A}_s} \sum_{s'} v_{s'}(k)\, P(s'|s,a) + r_{sa}(k)$$



Total Reward

(Finite Horizon)

$$R = \sum_s p_s(0)\; v_s(0)$$

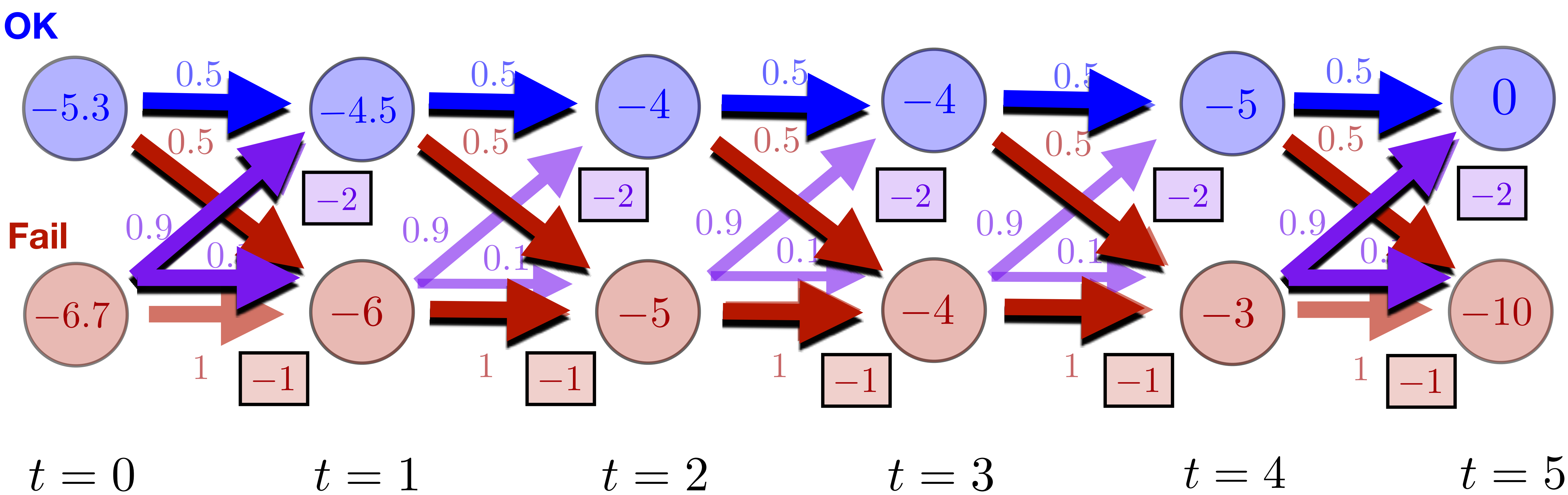
State-Value

for each state s $\quad v_s(k-1) = \max_{a\in\mathcal{A}_s} q_{sa}(k-1)$ *"Reward-to-go"*

Bellman Equation

Dynamic Programming

$$= \max_{a\in\mathcal{A}_s} \sum_{s'} v_{s'}(k)\, P(s'|s,a) + r_{sa}(k)$$



Total Reward

(Finite Horizon)

$$R = \sum_{s} p_s(0) \sum_{k=0}^{k=K} \sum_{s'} r_{s'a}\, p_{s'a}(k)$$

State-Value

for each state s

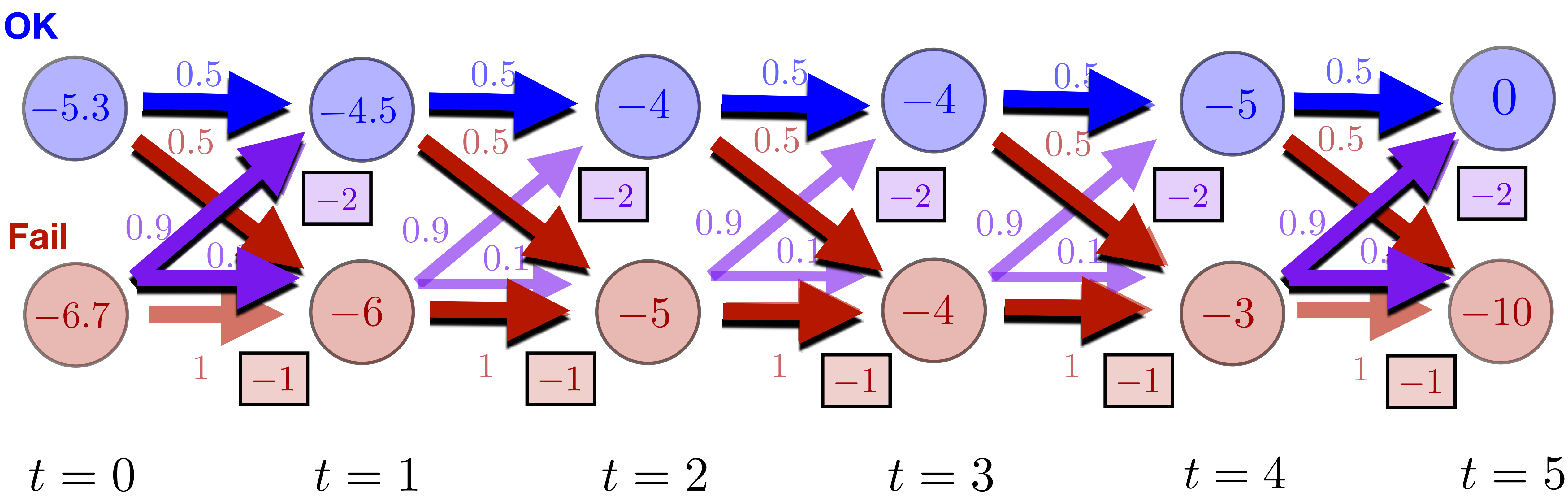
$$v_s(k-1) = \max_{a \in \mathcal{A}_s} q_{sa}(k-1)$$

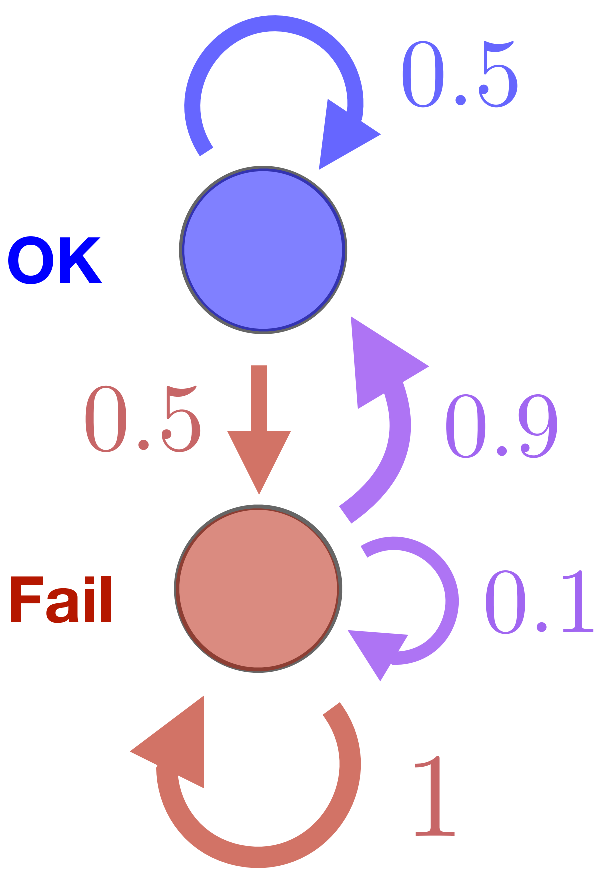
"Reward-to-go"

Bellman Equation

Dynamic Programming

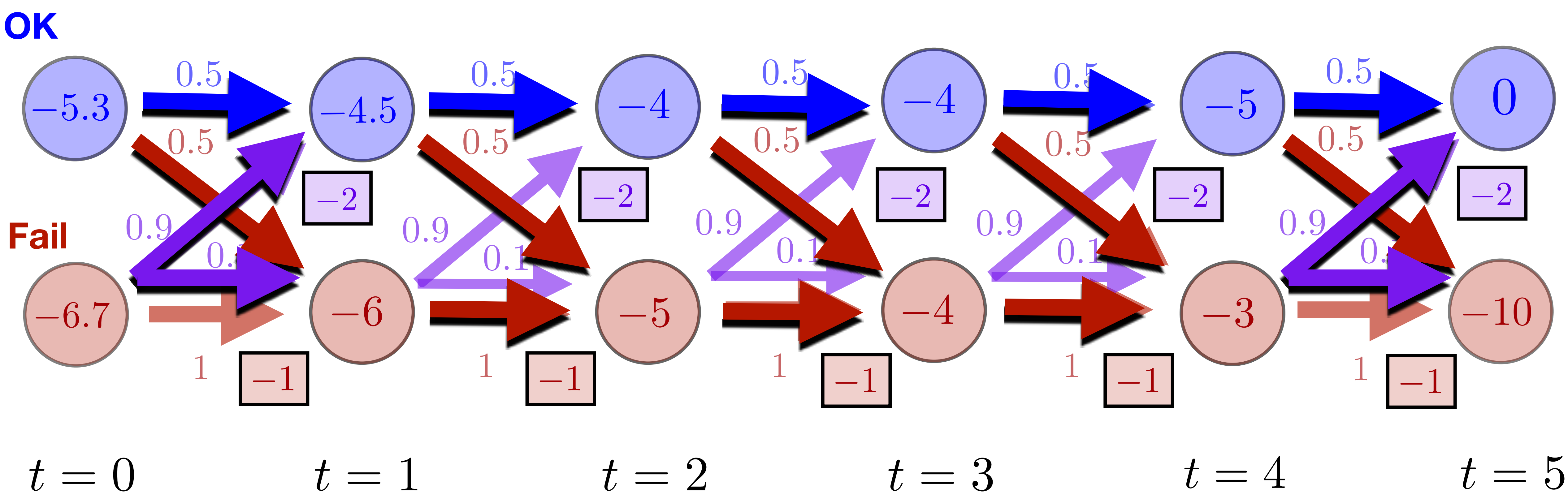
$$= \max_{a \in \mathcal{A}_s} \sum_{s'} v_{s'}(k)\, P(s'|s,a) + r_{sa}(k)$$

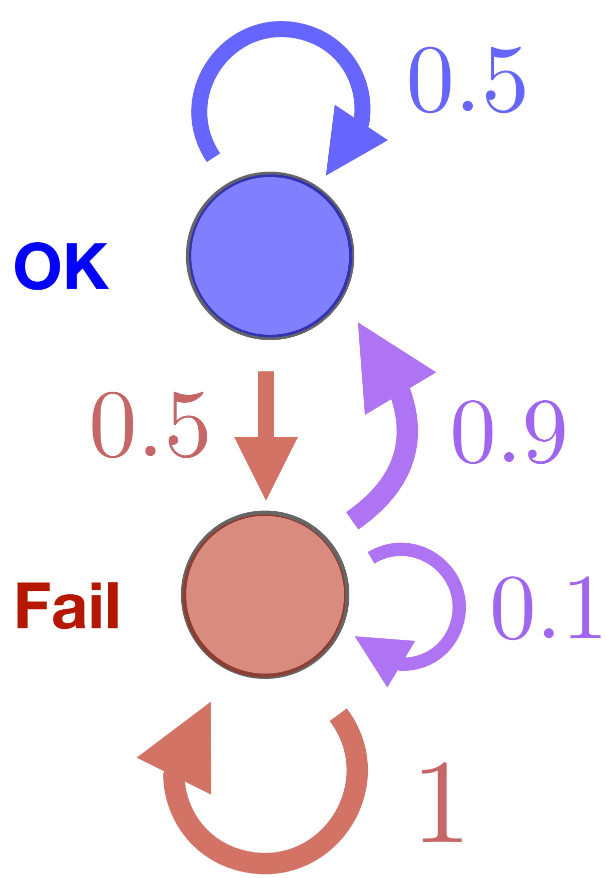




Bellman Equation

$$v_s(k-1) = \max_{a \in \mathcal{A}_s} \sum_{s'} v_{s'}(k)\, P(s'|s,a) + r_{sa}(k)$$



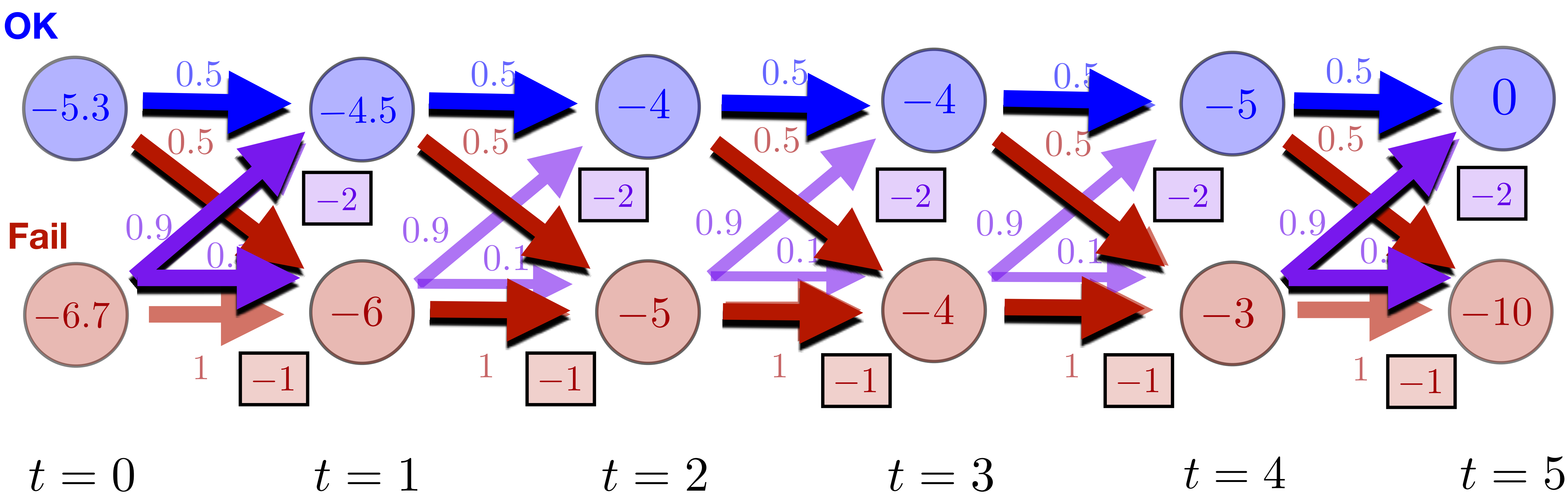


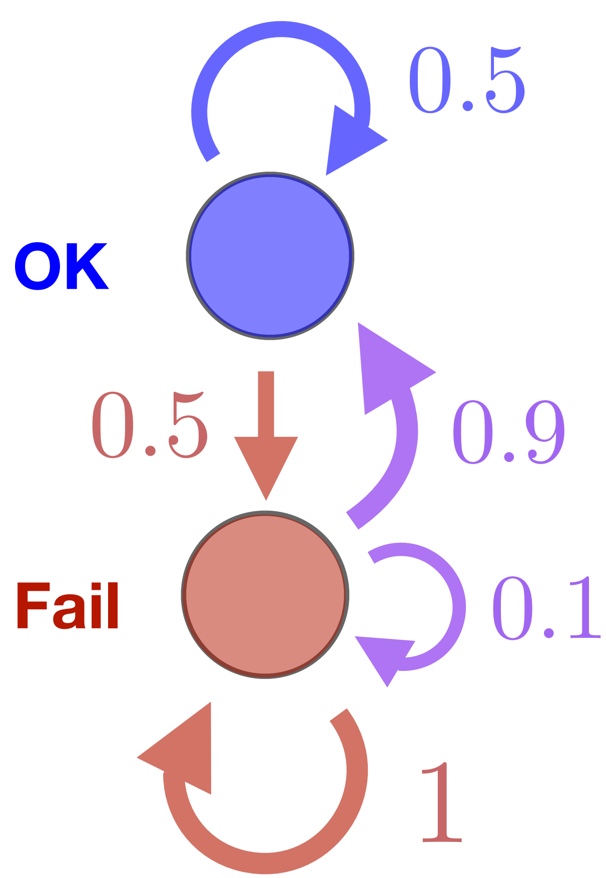
Bellman Equation

$$v_s(k-1) = \max_{a \in \mathcal{A}_s} \sum_{s'} v_{s'}(k)\, P(s'|s,a) + r_{sa}(k)$$

iterative equation…

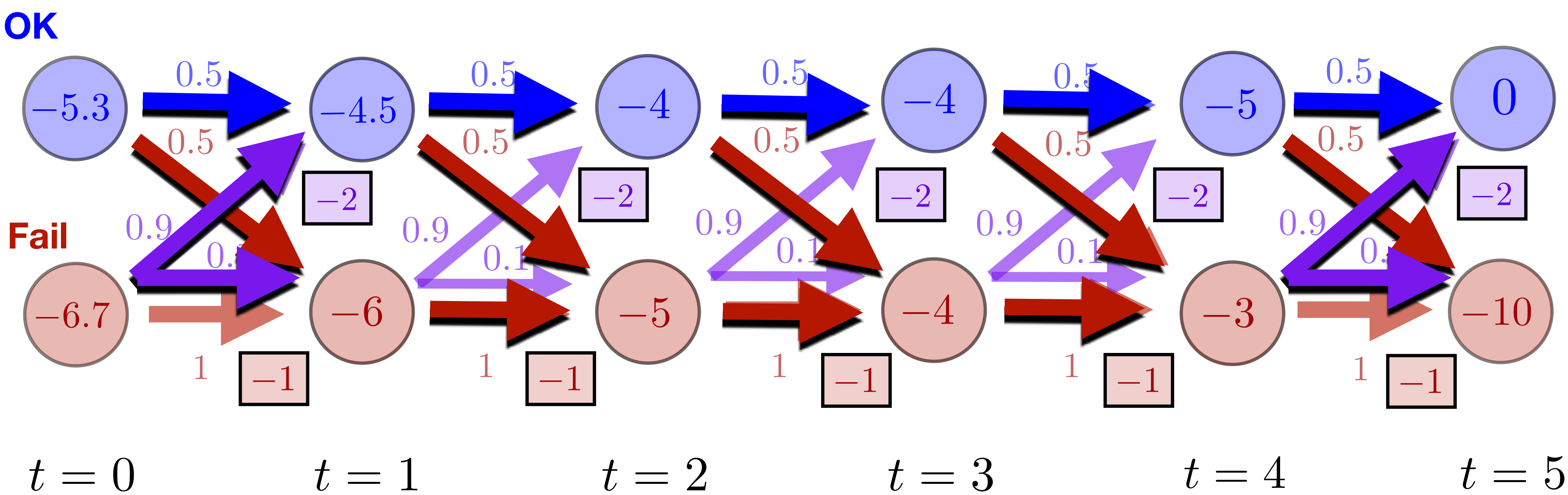
value doesn't necessarily converge…

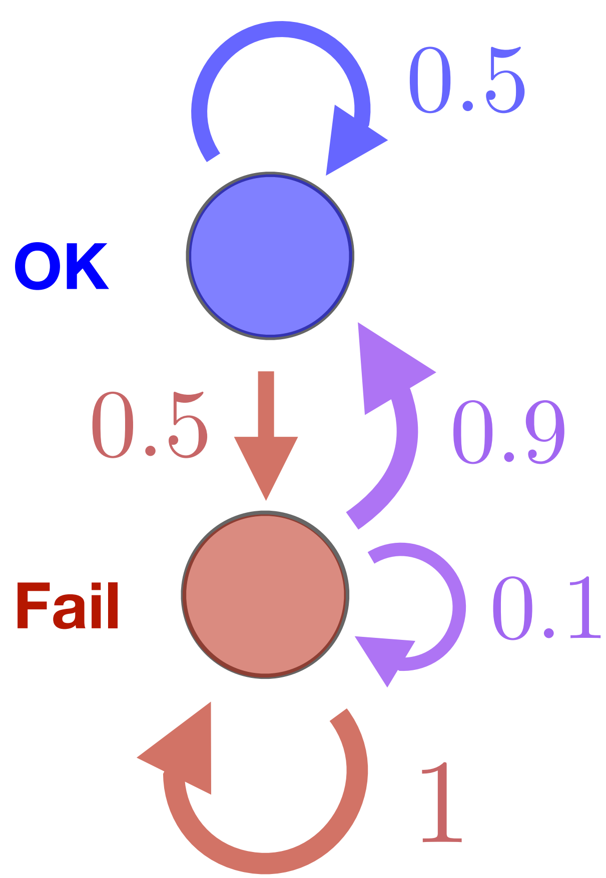




Discounted Bellman Equation $\gamma \in [0,1)$

$$v_s(k-1) = \max_{a \in \mathcal{A}_s} \sum_{s'} \gamma \, v_{s'}(k) \, P(s'|s,a) + r_{sa}(k)$$



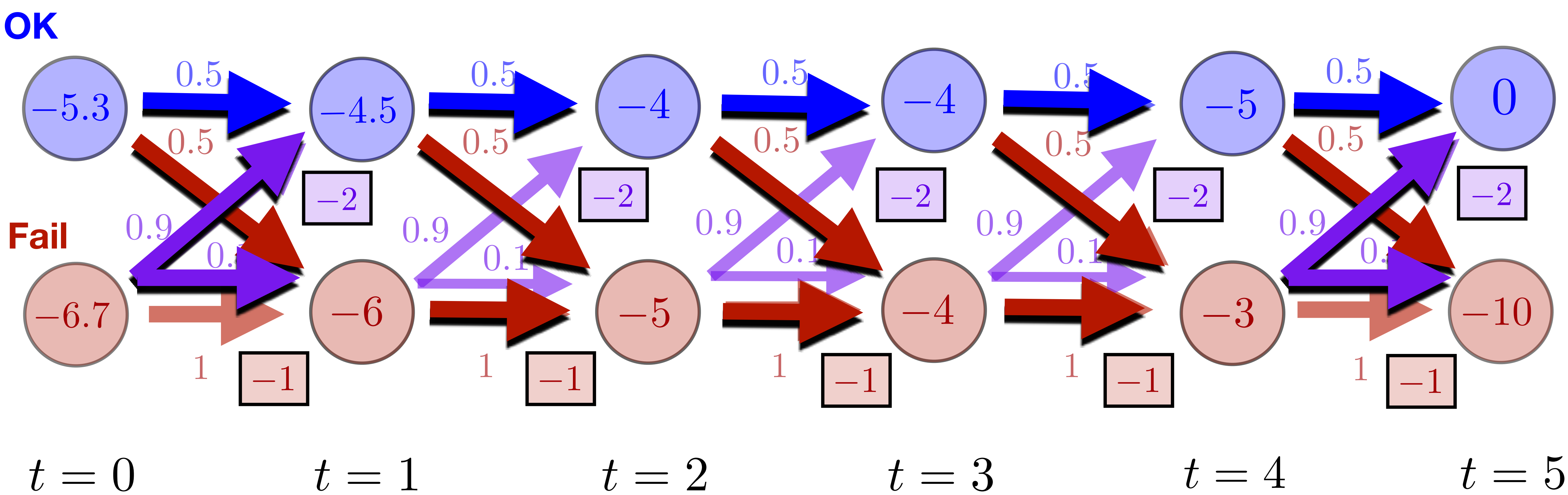


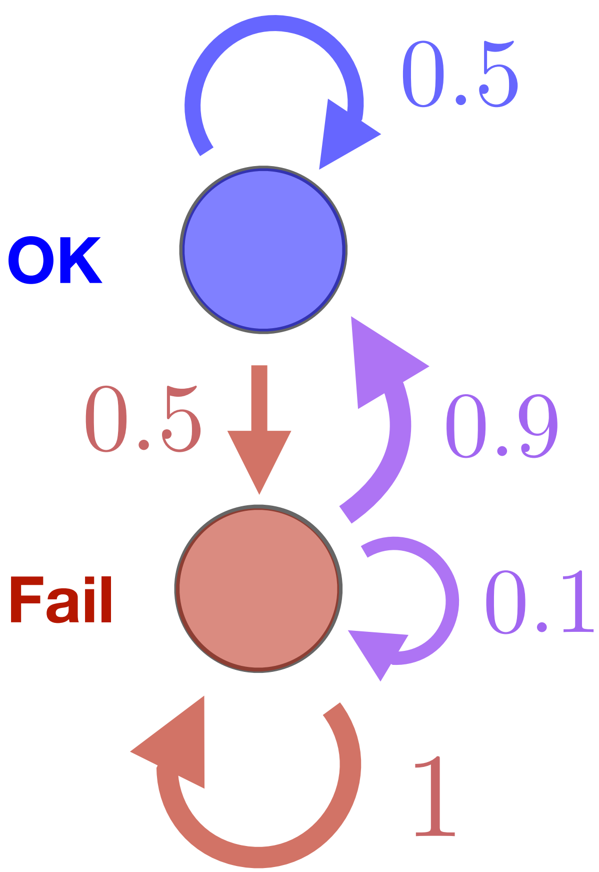
Discounted Bellman Equation $\gamma \in [0, 1)$

$$v_s(k-1) = \max_{a \in \mathcal{A}_s} \sum_{s'} \gamma \, v_{s'}(k) \, P(s'|s,a) + r_{sa}(k)$$

if $0 \leq \gamma < 1$ *we can always find* $v_s(\infty)$ *such that*

$$v_s(\infty) = \max_{a \in \mathcal{A}_s} \sum_{s'} \gamma \, v_{s'}(\infty) \, P(s'|s,a) + r_{sa}$$



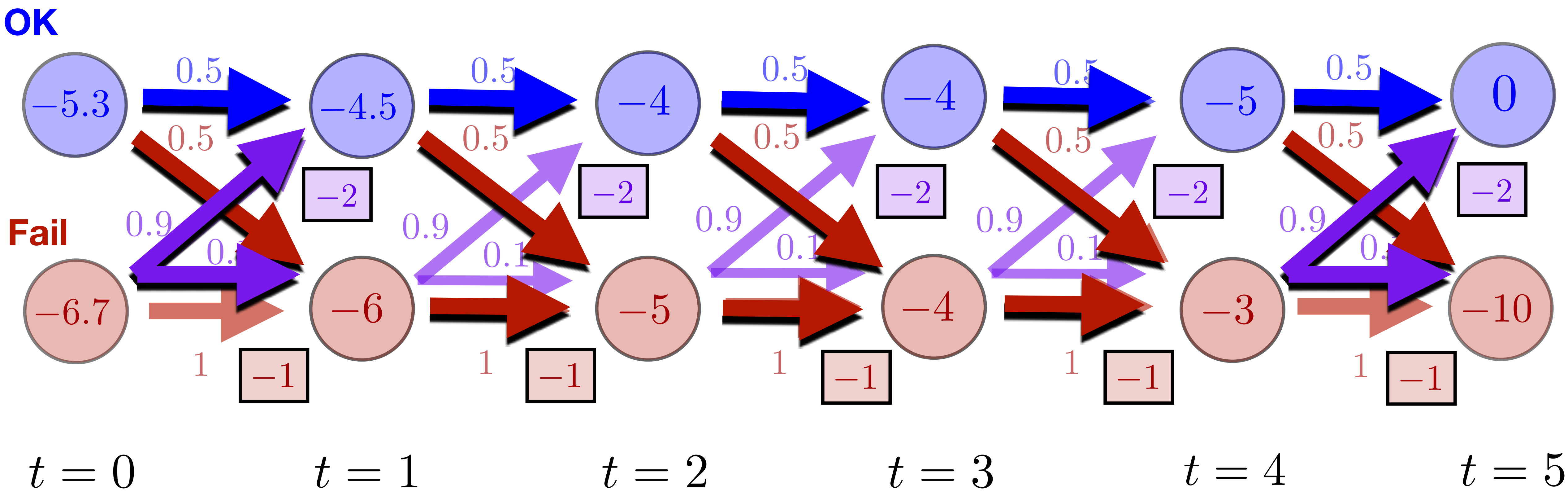


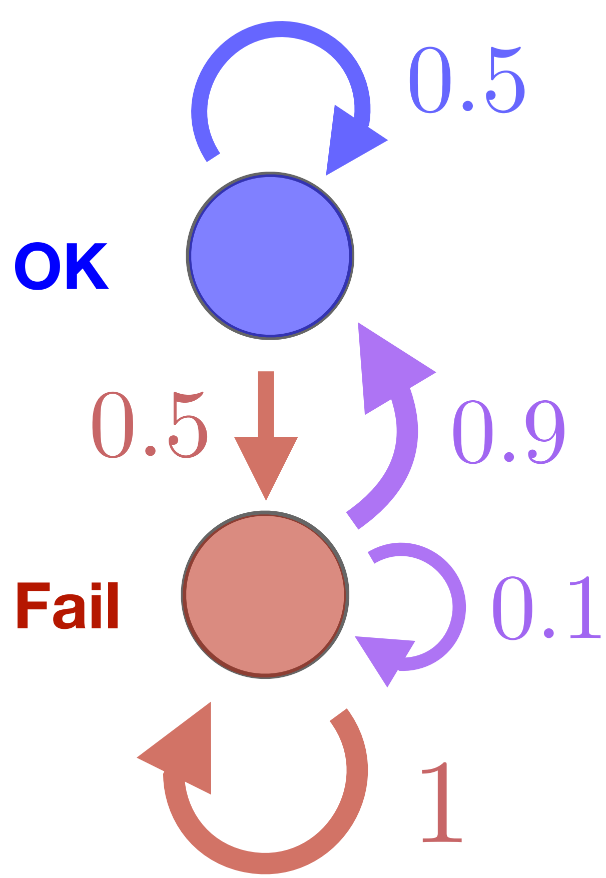
Discounted Bellman Equation $\gamma \in [0, 1)$

$$v_s(\infty) = \max_{a \in \mathcal{A}_s} \sum_{s'} \gamma \, v_{s'}(\infty) \, P(s'|s,a) + r_{sa}$$

Discounted State-Value Meaning

$$v_s(\infty) = \frac{1}{1-\gamma} \sum_{k=0}^{\infty} \sum_{s'} \gamma^k \, r_{s'a} \, p_{s'a}(k) \,, \qquad p_s(0) = 1$$



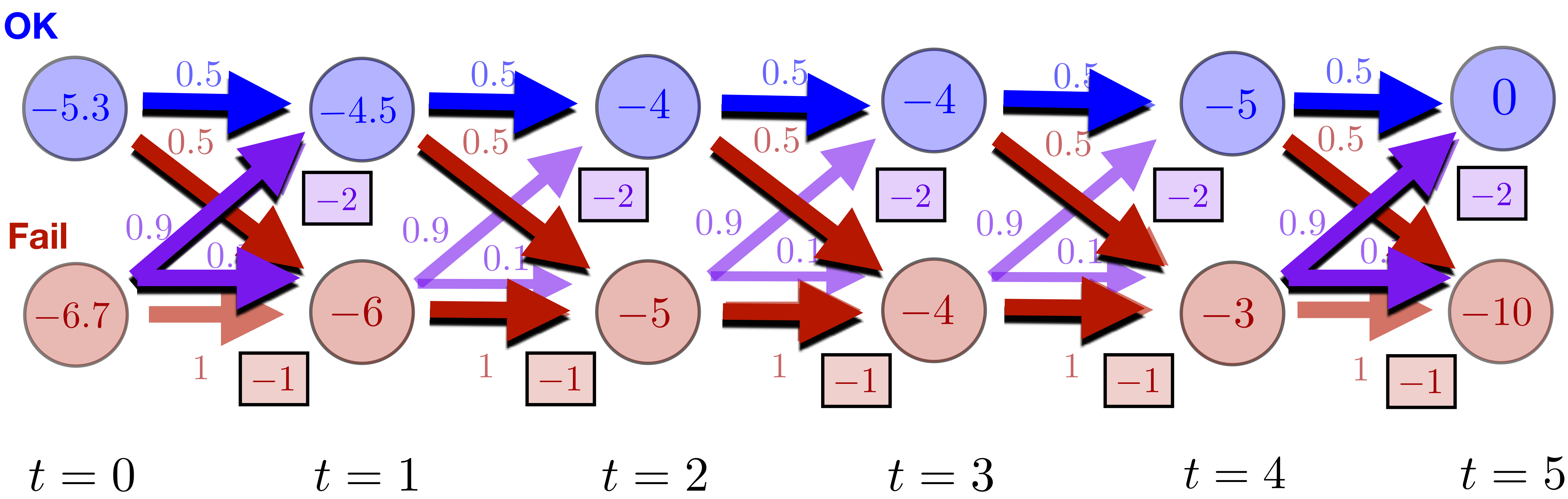


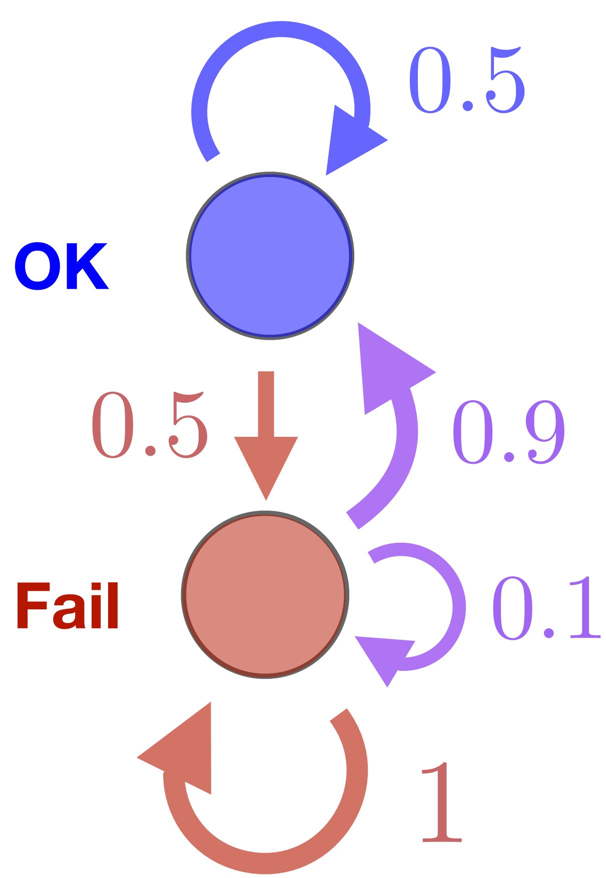
Discounted Bellman Equation $\gamma \in [0, 1)$

$$v_s(\infty) = \max_{a \in \mathcal{A}_s} \sum_{s'} \gamma \, v_{s'}(\infty) \, P(s'|s,a) + r_{sa}$$

Discounted Reward

$$R_\gamma = \sum_s p_s(0) \; v_s(\infty)$$



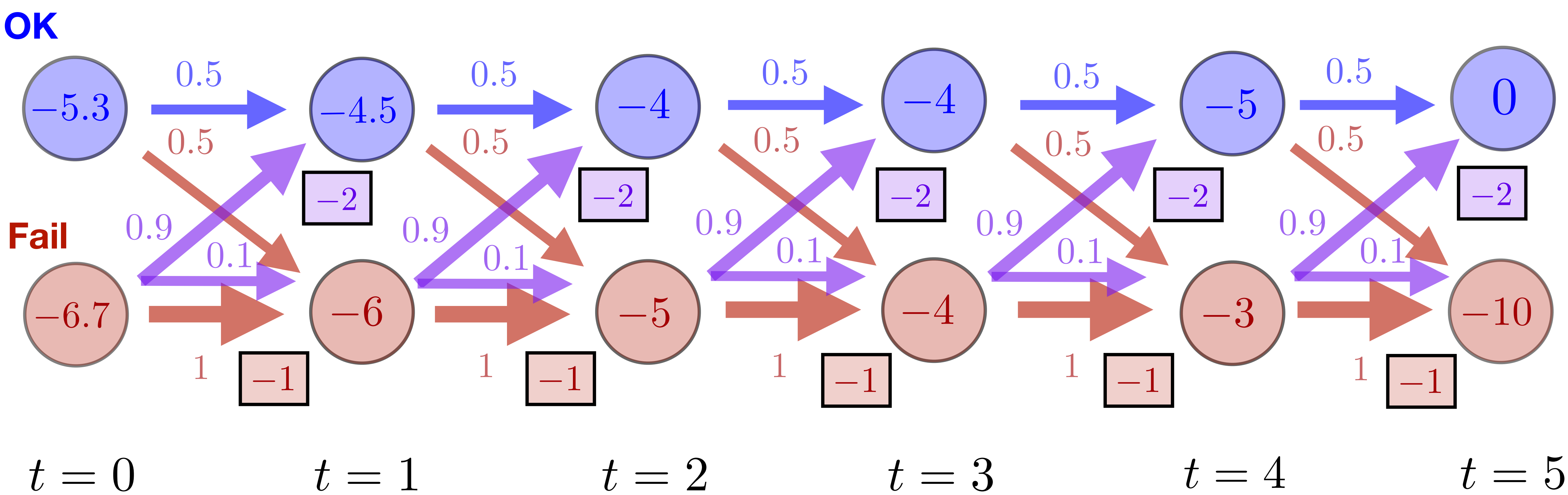


Discounted Bellman Equation $\gamma \in [0, 1)$

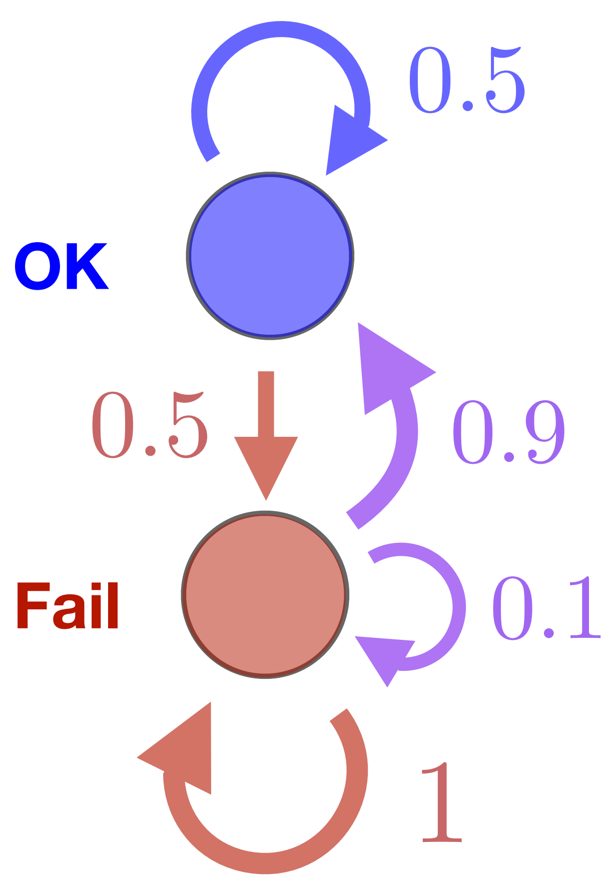
$$v_s(\infty) = \max_{a \in \mathcal{A}_s} \sum_{s'} \gamma \, v_{s'}(\infty) P(s'|s, a) + r_{sa}$$

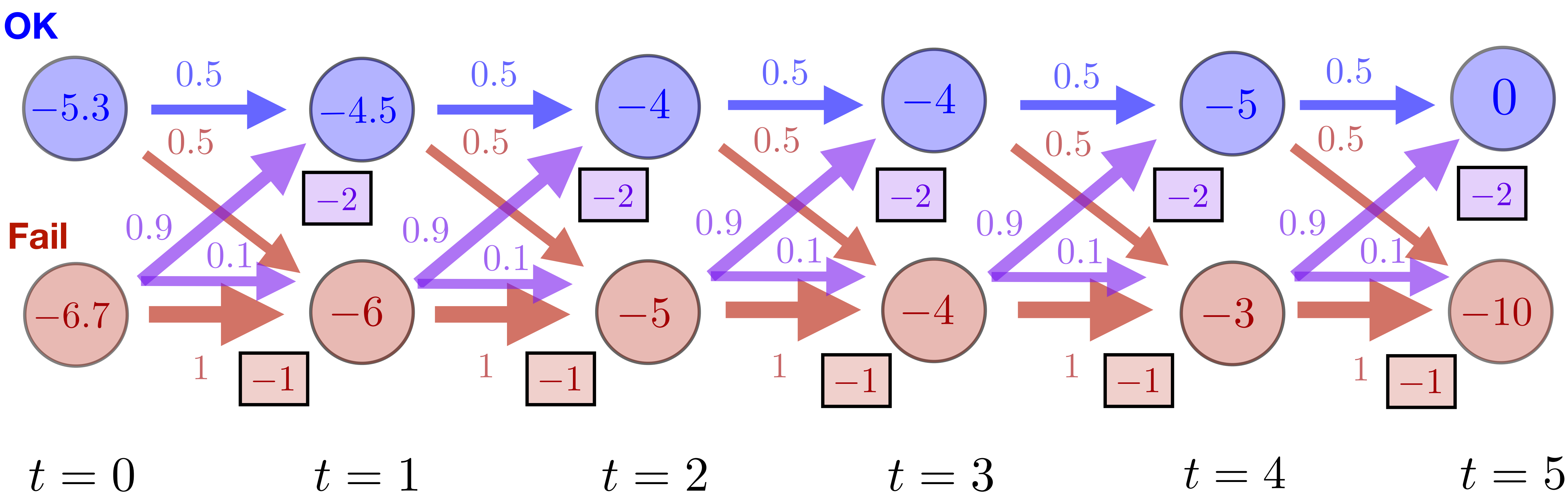
Discounted Reward

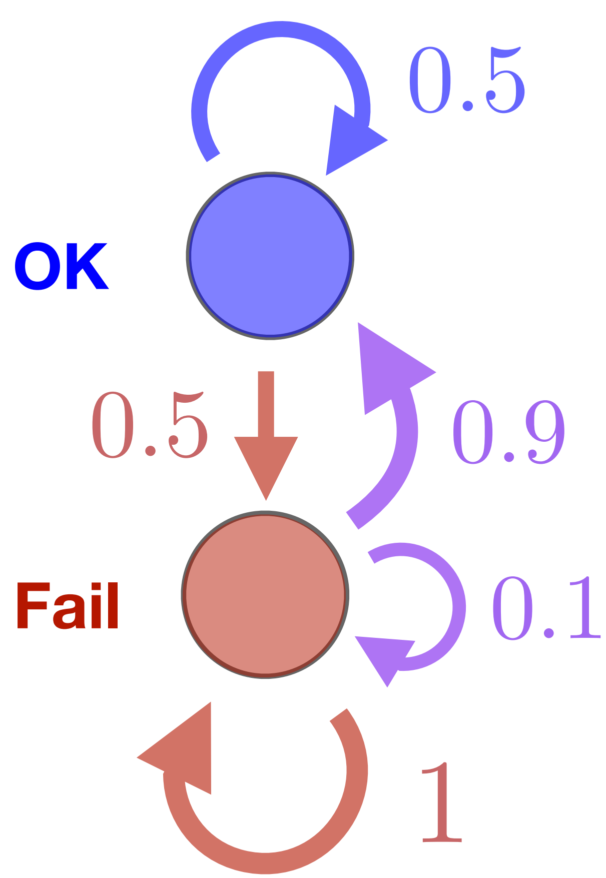
$$R_\gamma = \sum_s p_s(0) \left(\frac{1}{1-\gamma} \sum_{k=0}^{\infty} \sum_{s'} \gamma^k \, r_{s'a} \, p_{s'a}(k) \right)$$



Connection to Markov Chains







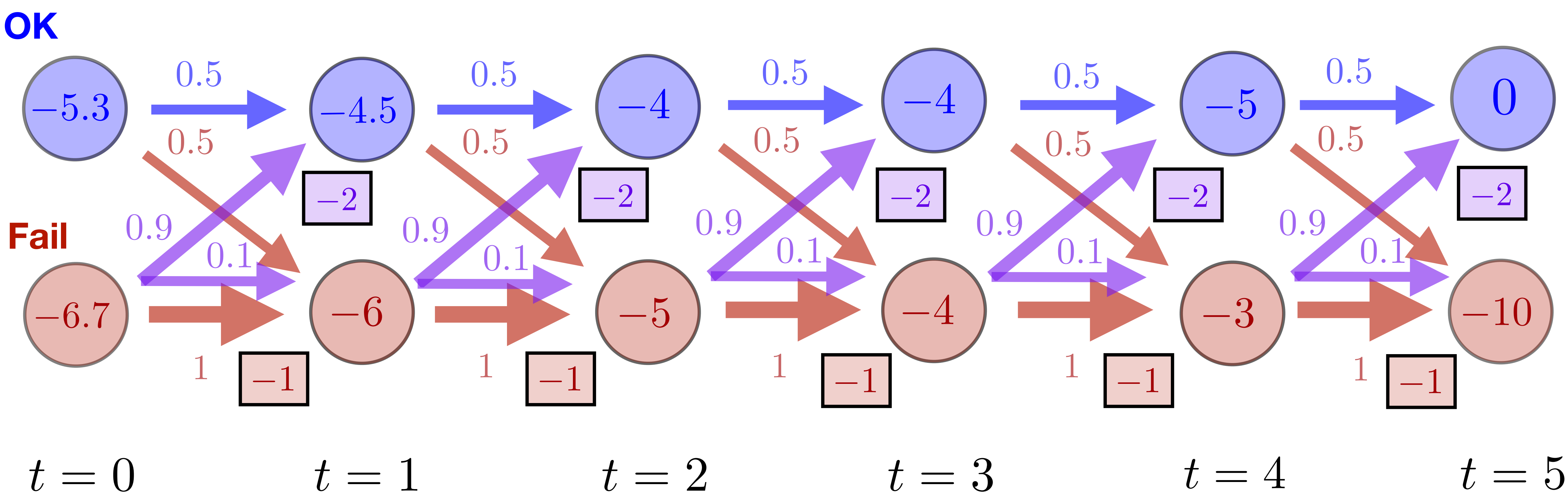
Connection to Markov Chains

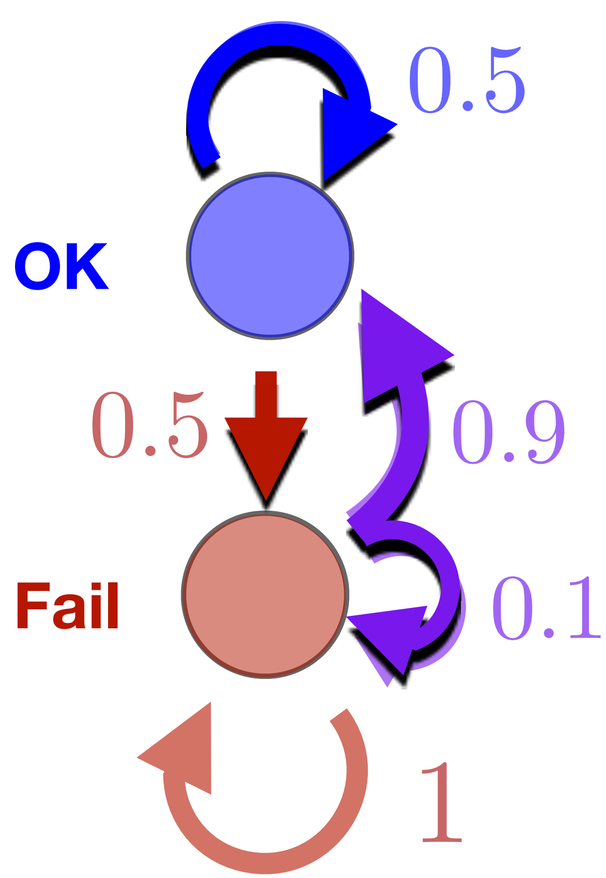
state 1 | state 2

$$\mathbf{P} = \left[\begin{array}{c|cc} 0.5 & 0.9 & 0 \\ 0.5 & 0.1 & 1 \end{array}\right]$$

action 1 | action 1 | action 2

probability transition kernel





Connection to Markov Chains

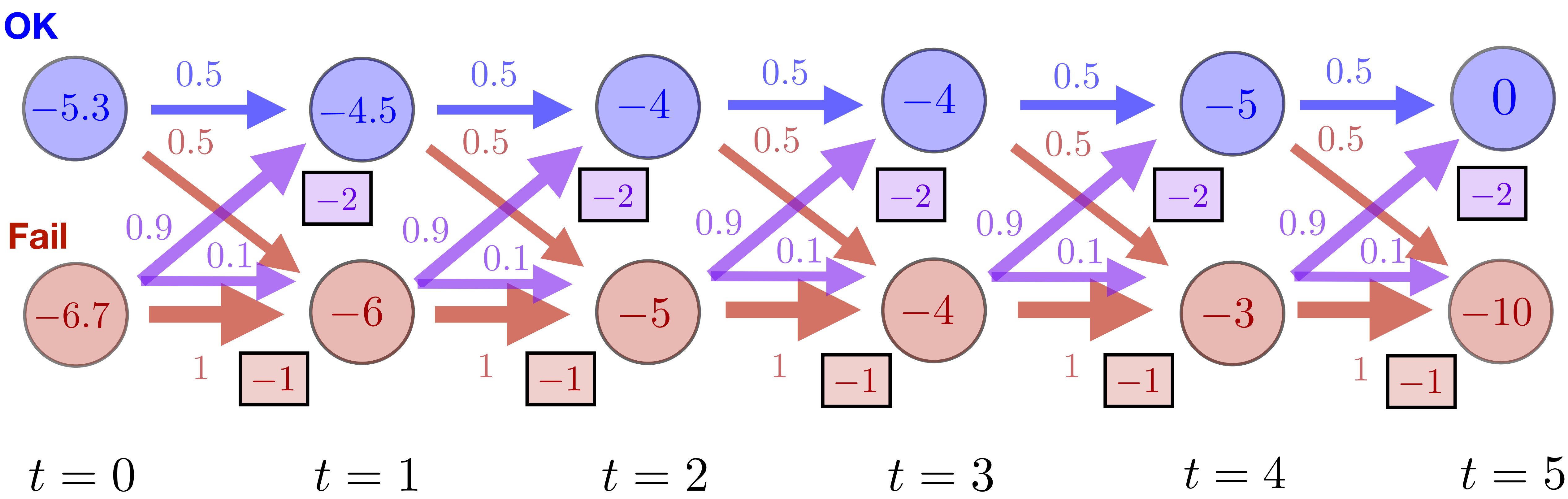
$$\mathbf{P} = \left[\begin{array}{c|cc} 0.5 & 0.9 & 0 \\ 0.5 & 0.1 & 1 \end{array}\right]$$

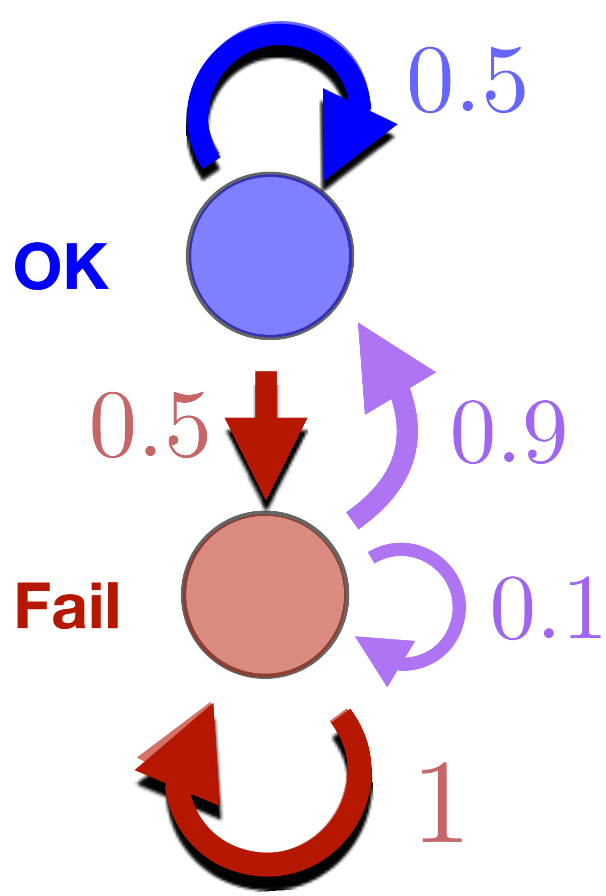
state 1 (action 1) | state 2 (action 1, action 2)

probability transition kernel

$$M = \left[\begin{array}{c|c} 0.5 & 0.9 \\ 0.5 & 0.1 \end{array}\right]$$

“choosing a policy determines a Markov chain”





Connection to Markov Chains

	state 1	state 2	
	action 1	action 1	action 2

$$\mathbf{P} = \left[\begin{array}{c|cc} 0.5 & 0.9 & 0 \\ 0.5 & 0.1 & 1 \end{array}\right]$$

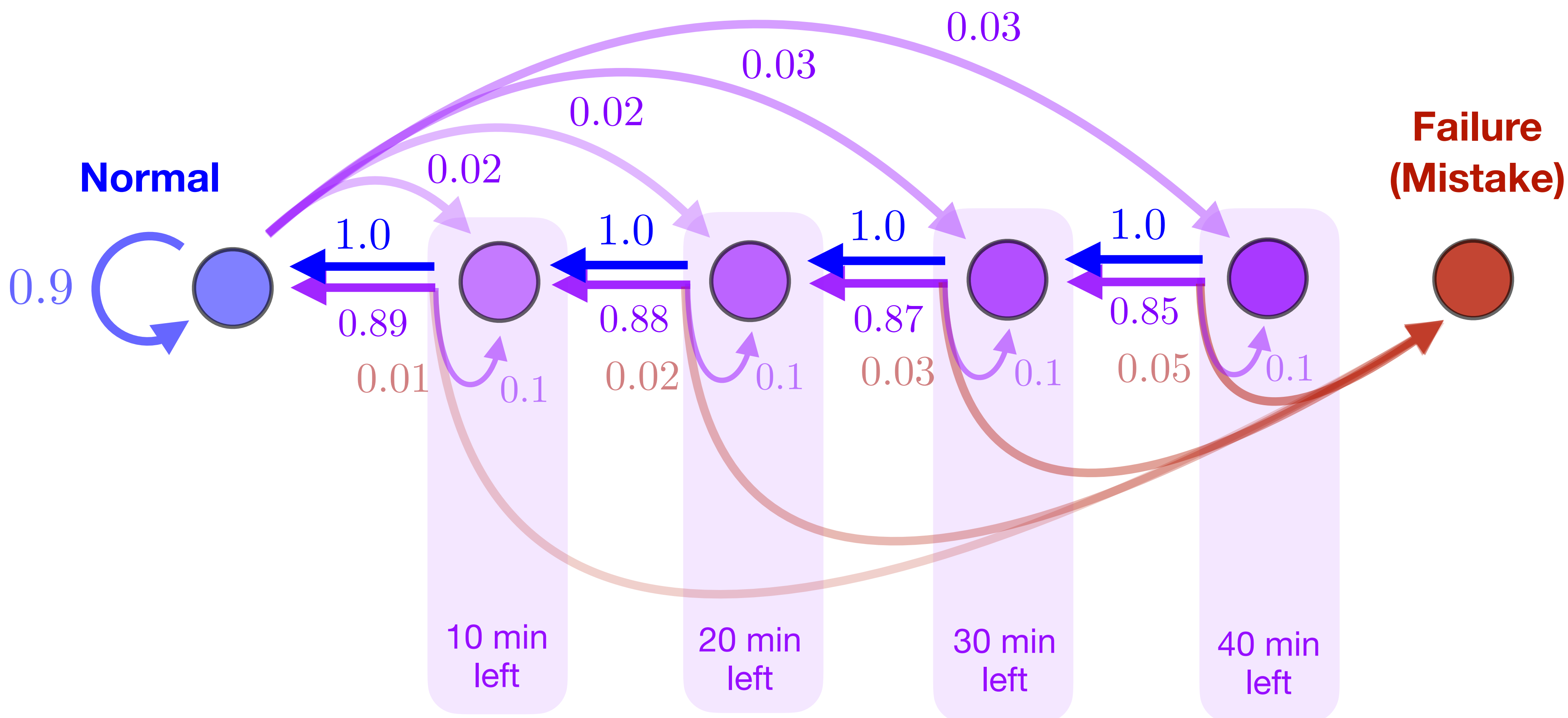
probability transition kernel

$$M = \left[\begin{array}{c|c} 0.5 & 0 \\ 0.5 & 1 \end{array}\right]$$

"choosing a policy determines a Markov chain"

Markov Decision Process (MDP) Models

5. *Modeling human decision making as an policy in an MDP*
6. *Compute optimal policy (over time) to keep failure risk below some threshold*



Markov Decision Process Models

5. *Modeling human decision making as an policy in an MDP*
6. *Compute optimal policy (over time) to keep failure risk below some threshold*