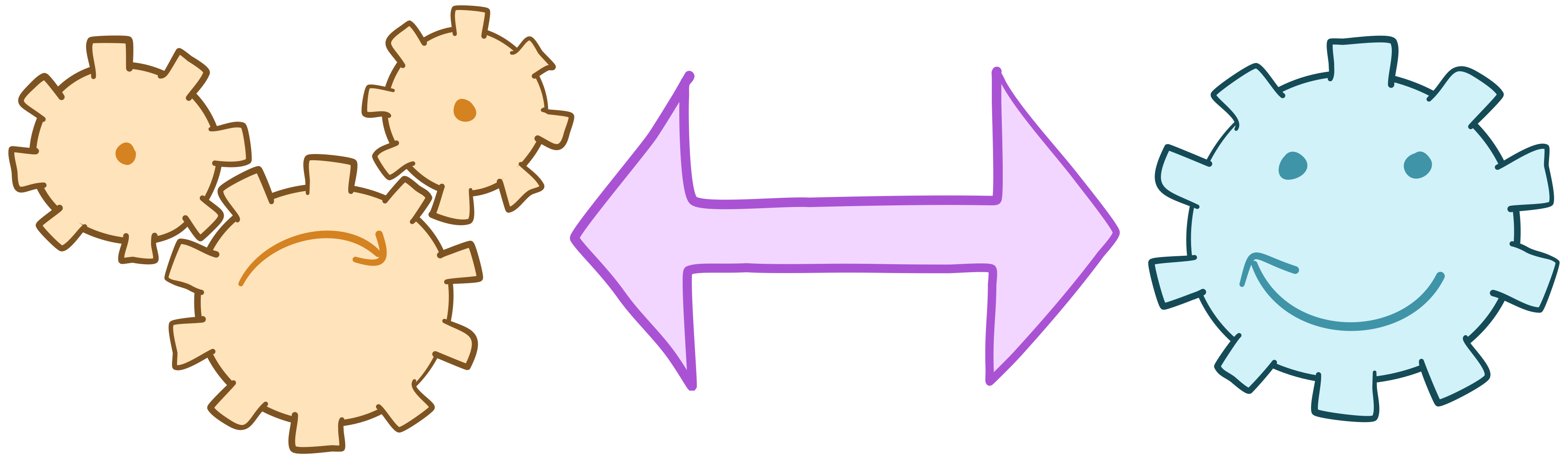
Coupling Techniques

for Complex Control Problems

Ziv Scully
Carnegie Mellon University

Sid Banerjee
Cornell University



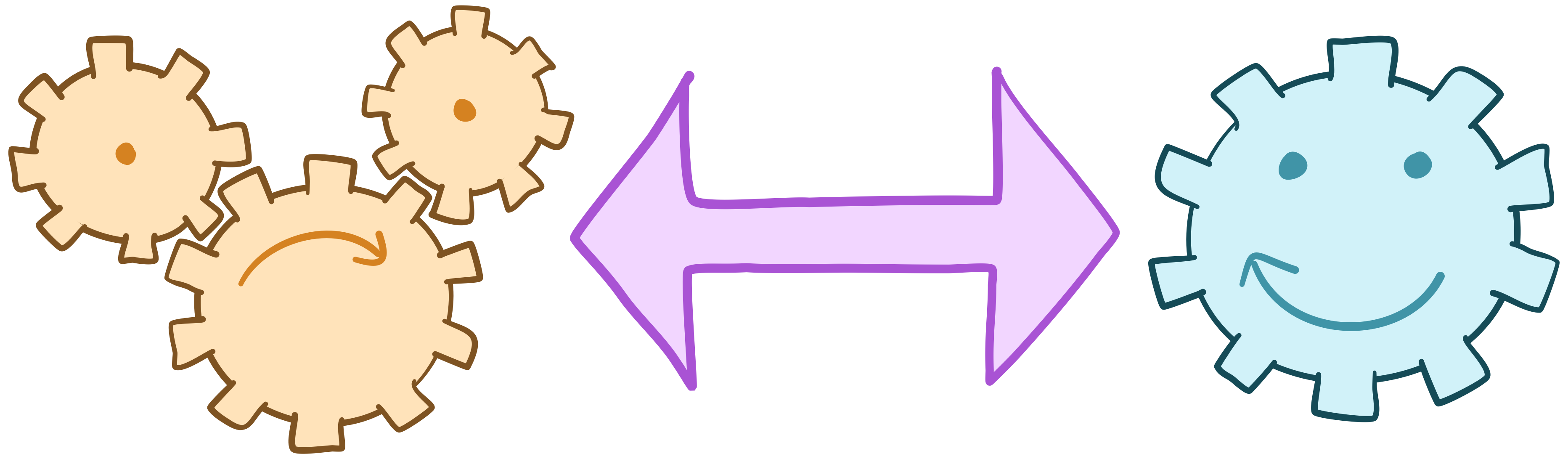
Coupling Techniques

for Complex Control Problems

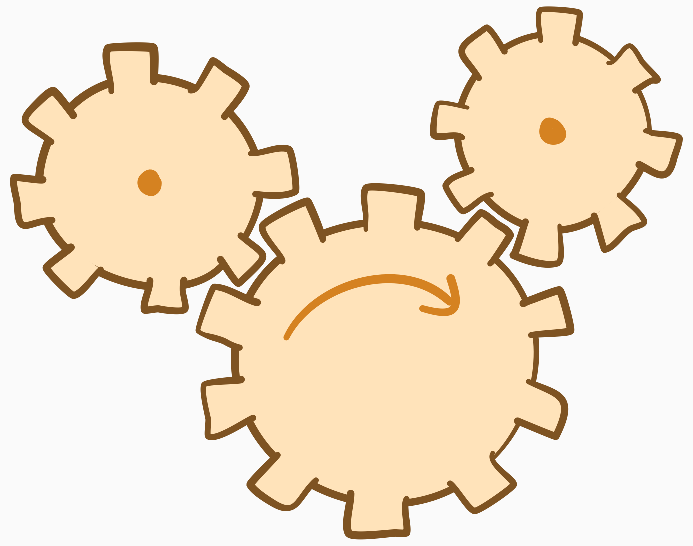
Part 1 of 2

Ziv Scully
Carnegie Mellon University

Sid Banerjee
Cornell University

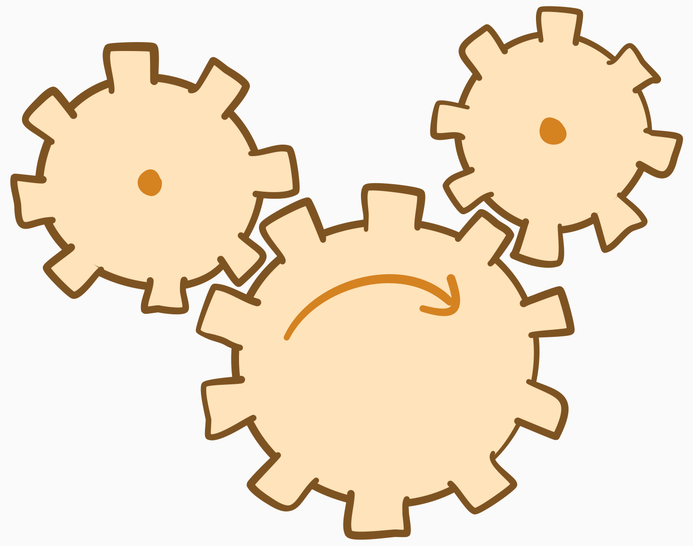


Complex stochastic systems



World is full of **complex systems**

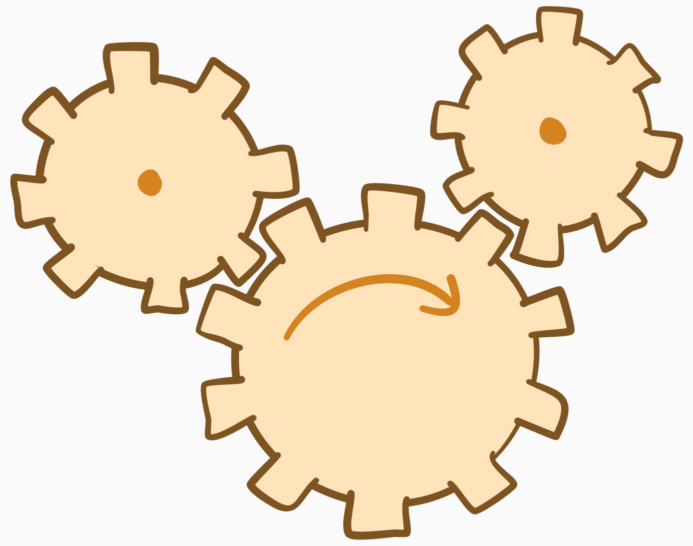
Complex stochastic systems



World is full of **complex systems**

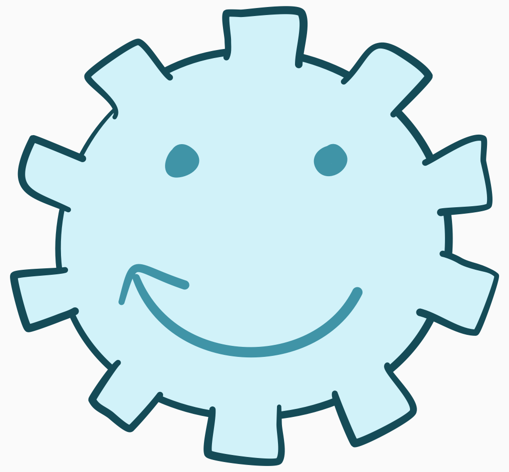
- Epidemics in social networks
- Inventory management
- Ride-sharing networks
- Multiserver queueing systems
- Load-balancing systems
- Many more…

Complex stochastic systems



World is full of **complex systems**

- Epidemics in social networks
- Inventory management
- Ride-sharing networks
- Multiserver queueing systems
- Load-balancing systems
- Many more…

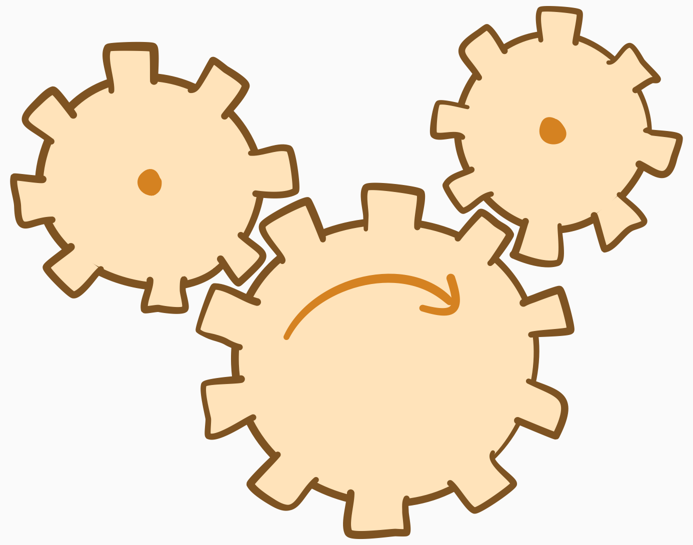


Coupling:

a way to analyze **complex systems** by working with related **easy systems**

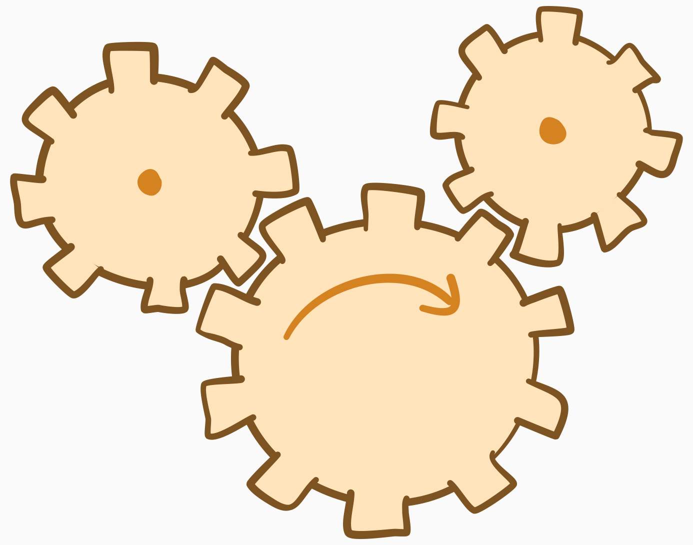
What is coupling?

complex system *X*



What is coupling?

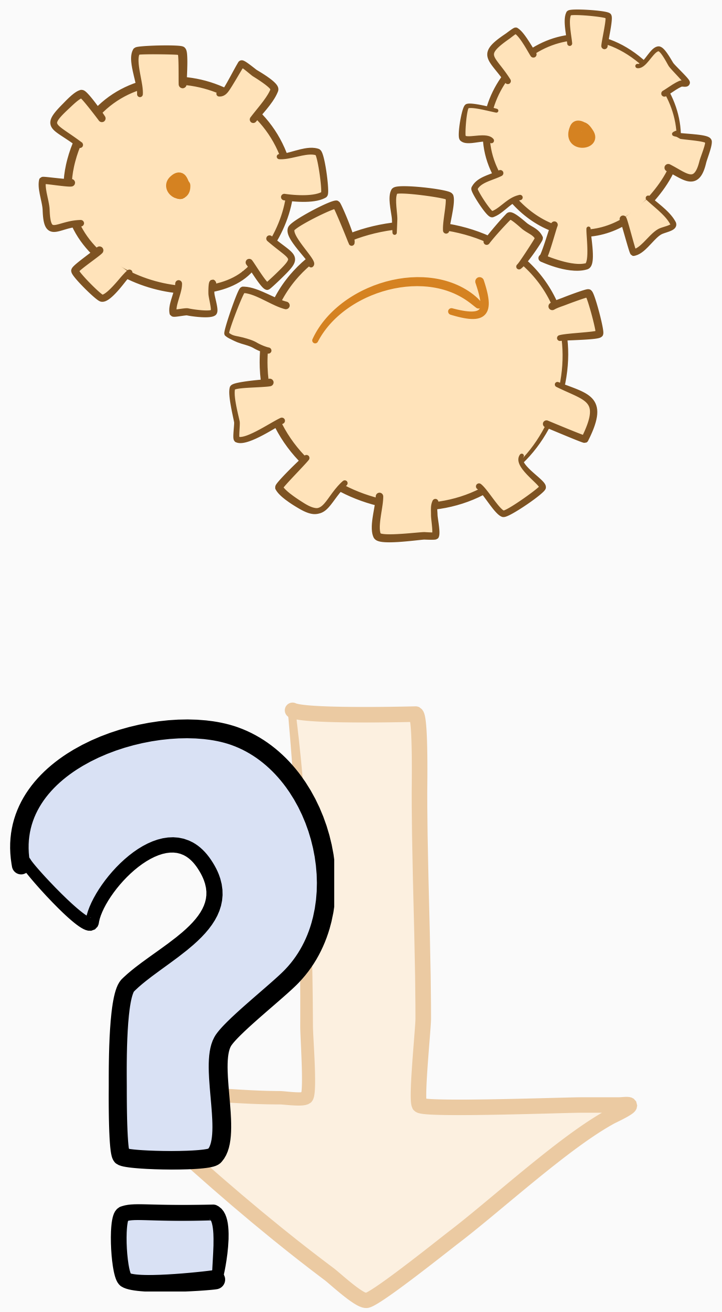
complex system *X*



Goal: answer a
question about ***X***
(approximate is okay)

What is coupling?

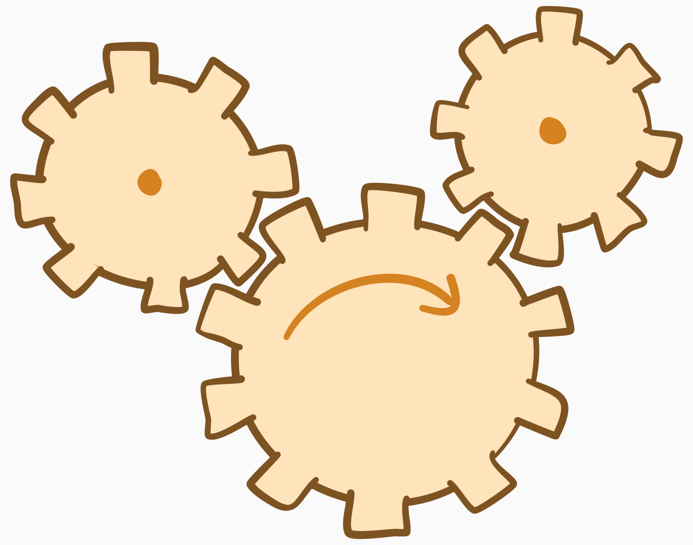
complex system *X*



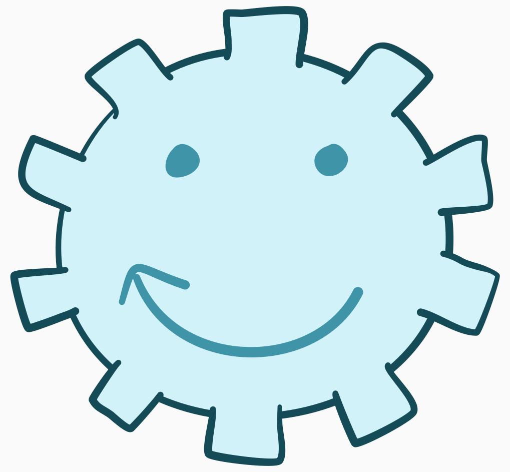
Goal: answer a question about *X*
(approximate is okay)

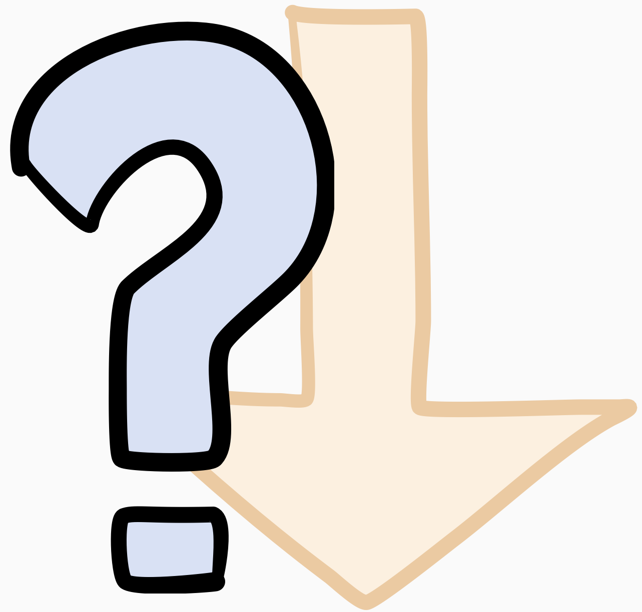
What is coupling?

complex system *X*



easy system *Y*

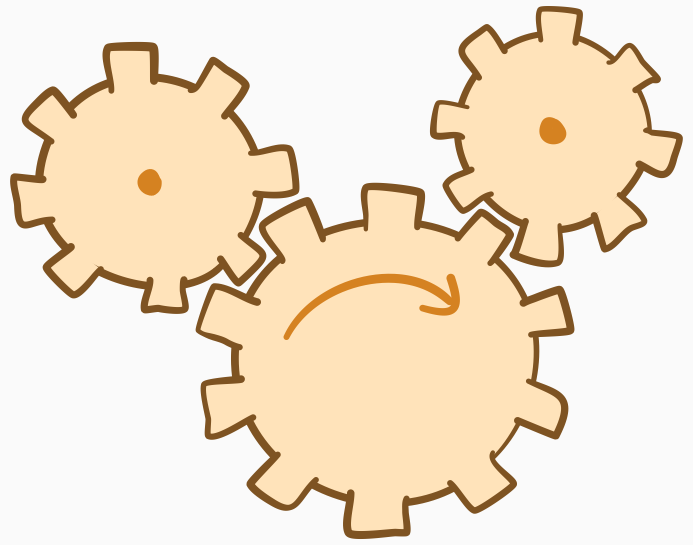


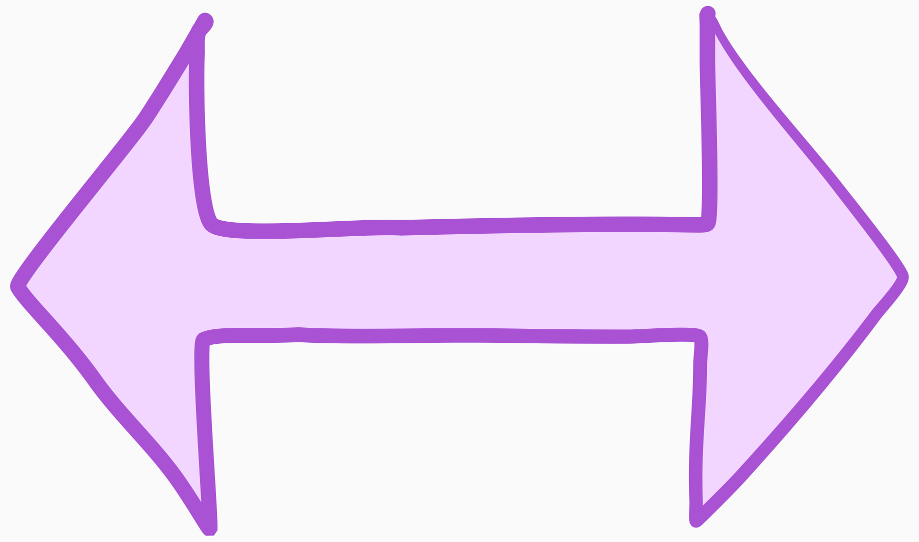


Goal: answer a question about *X*
(approximate is okay)

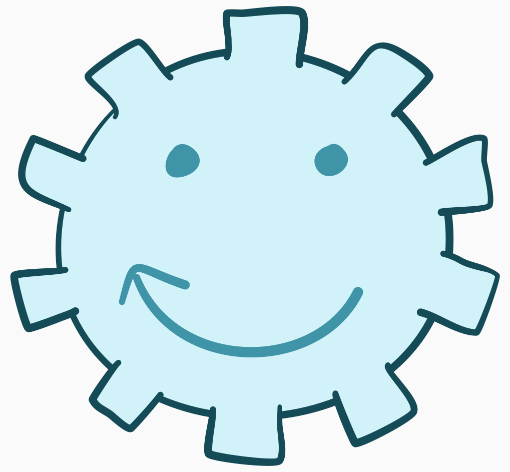
What is coupling?

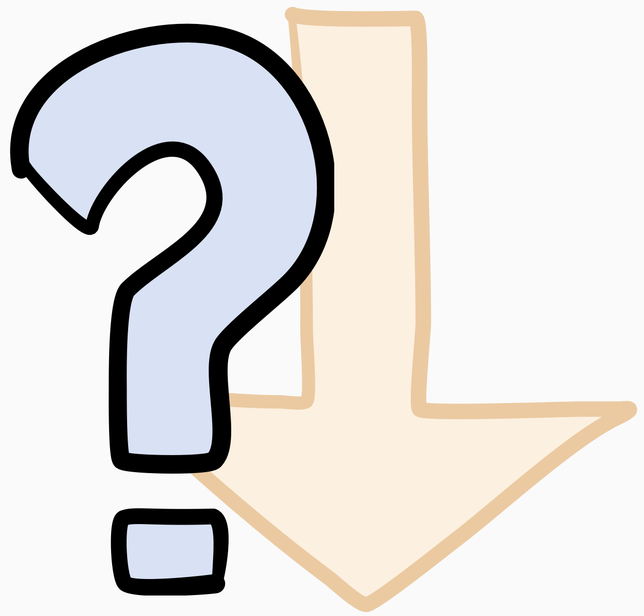
complex system *X*





easy system *Y*

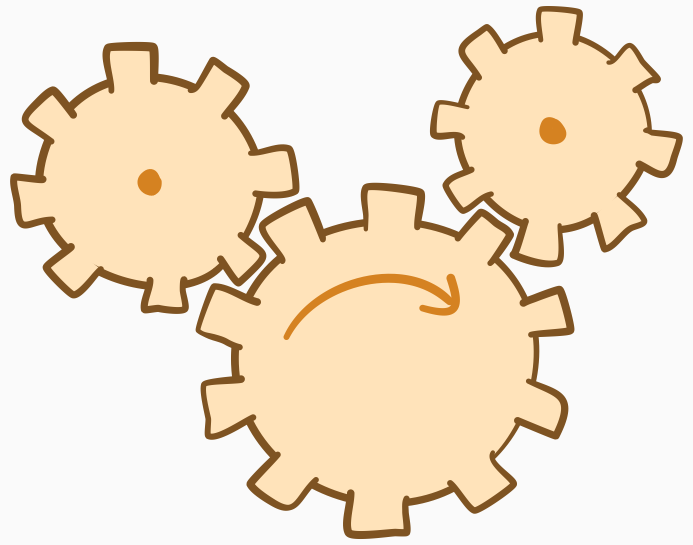


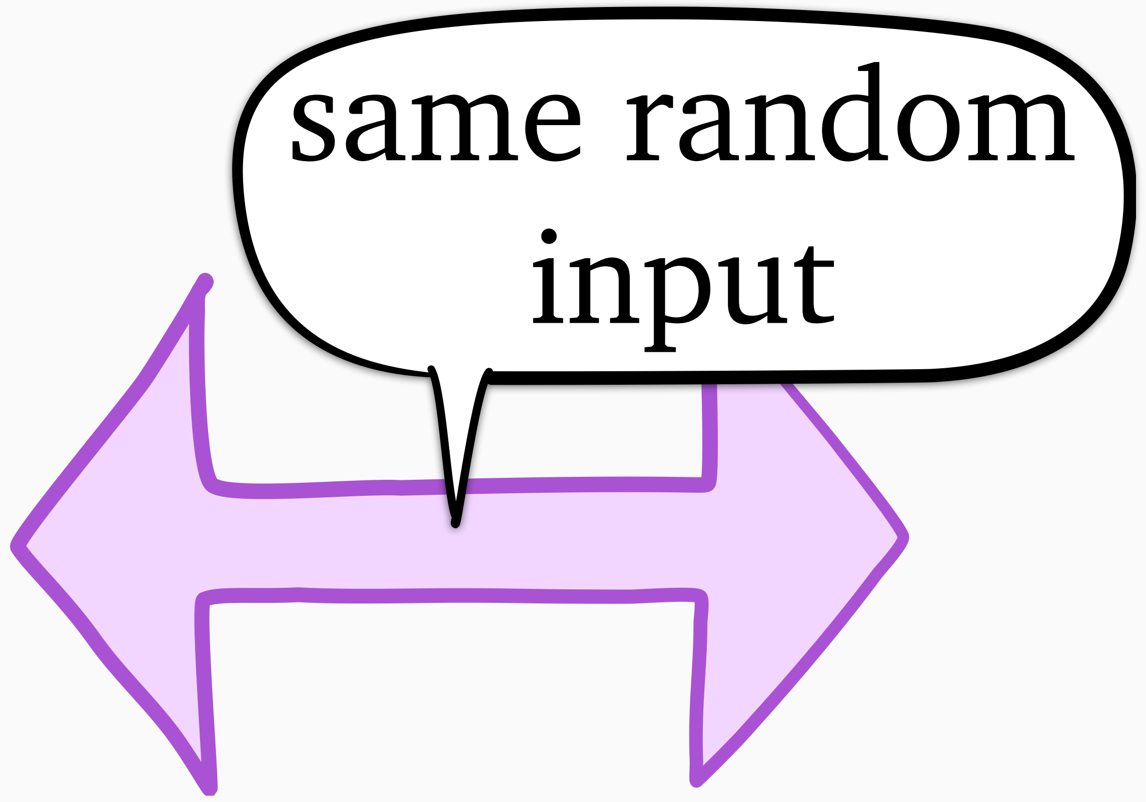


Goal: answer a question about *X*
(approximate is okay)

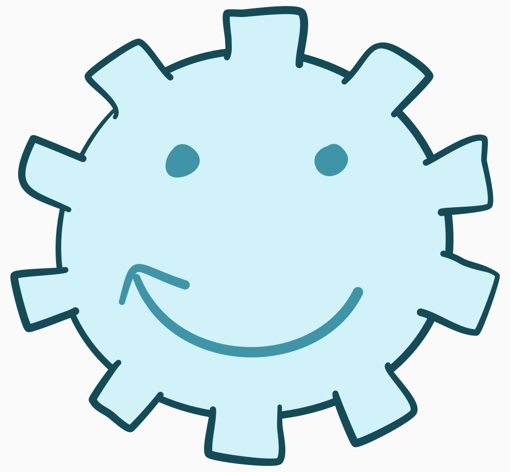
What is coupling?

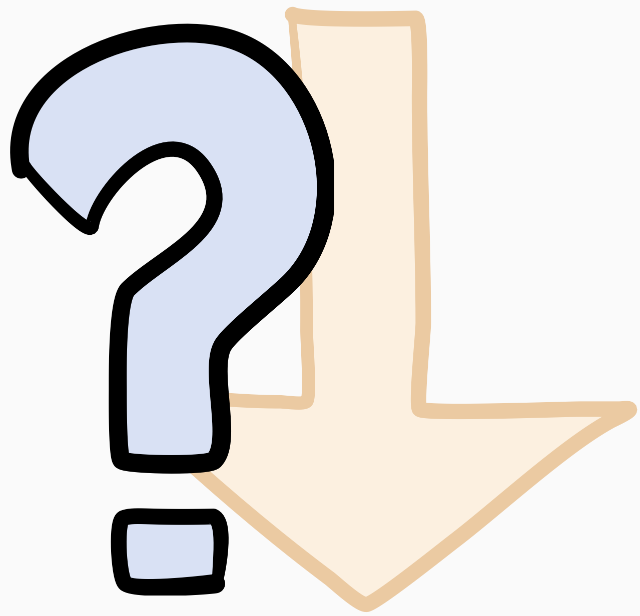
complex system X





easy system Y

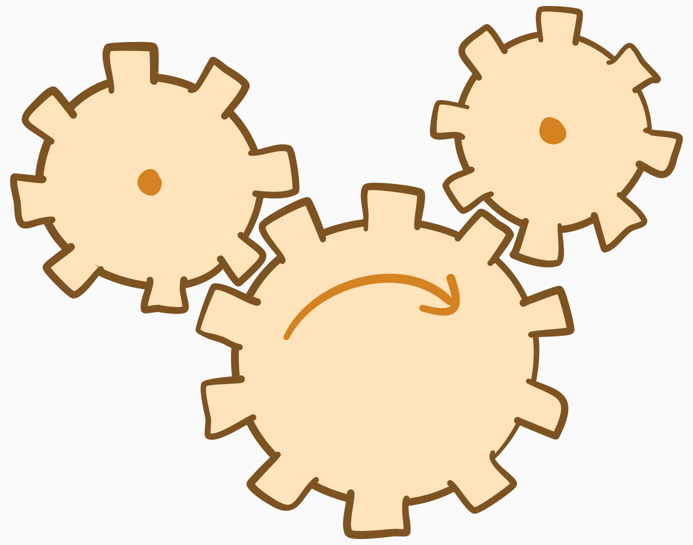


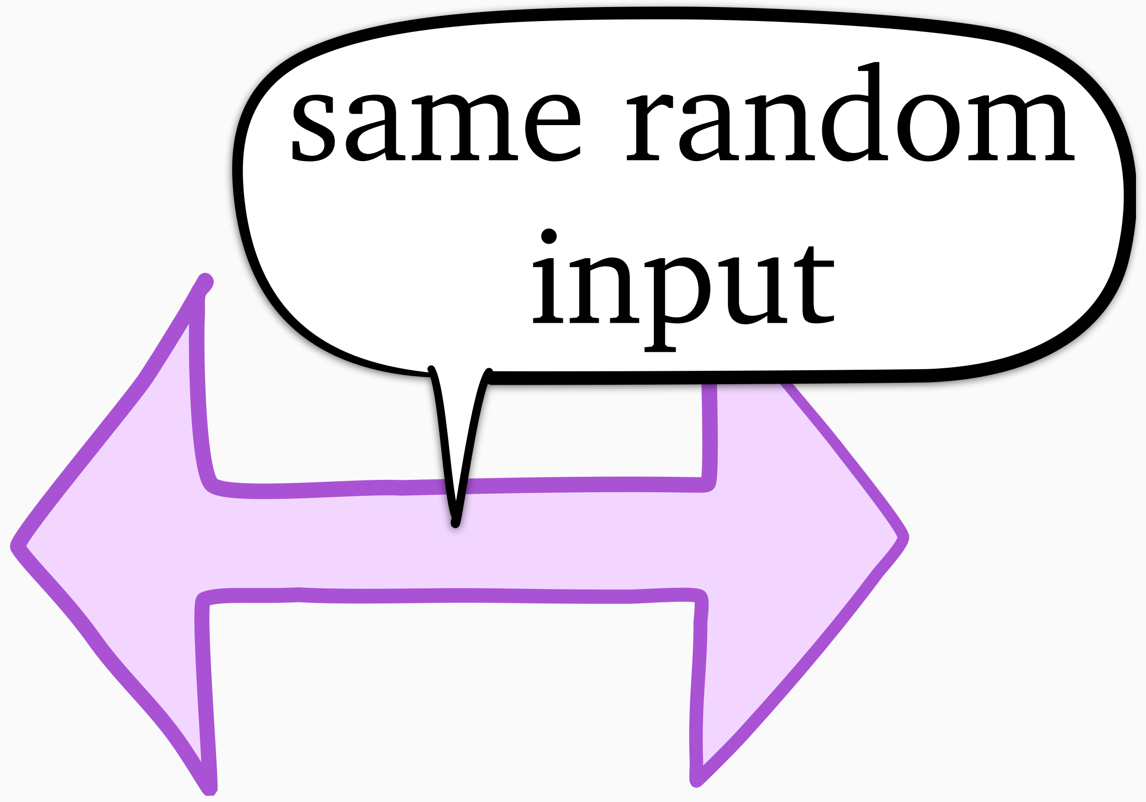


Goal: answer a question about X
(approximate is okay)

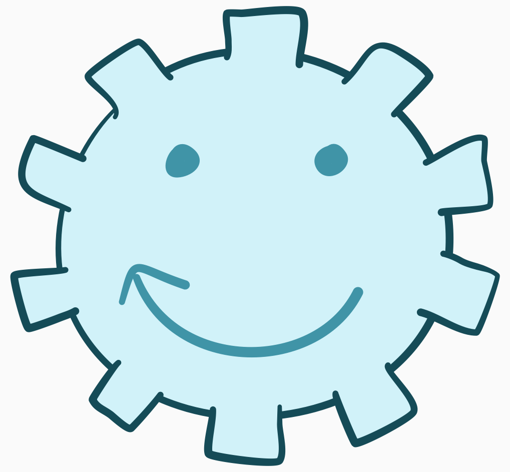
What is coupling?

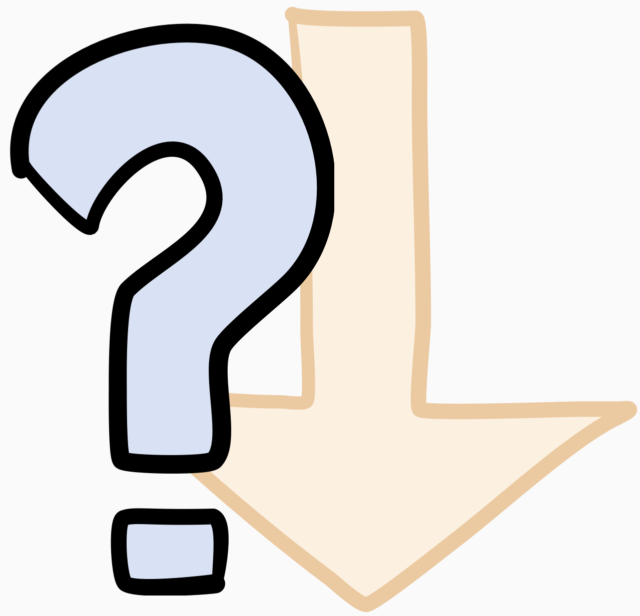
complex system X



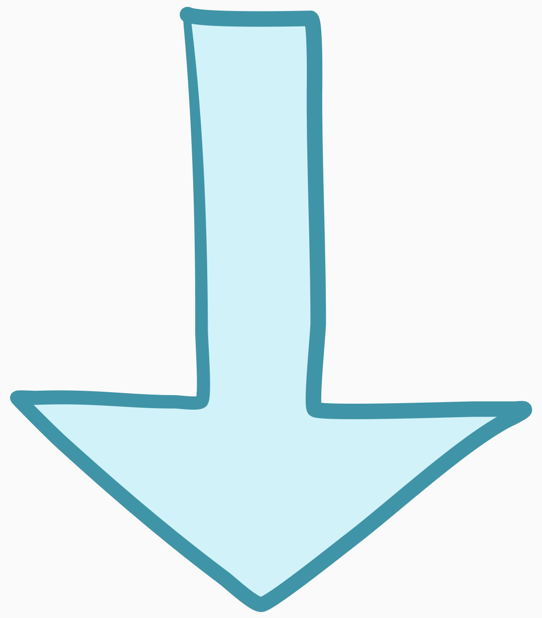


easy system Y



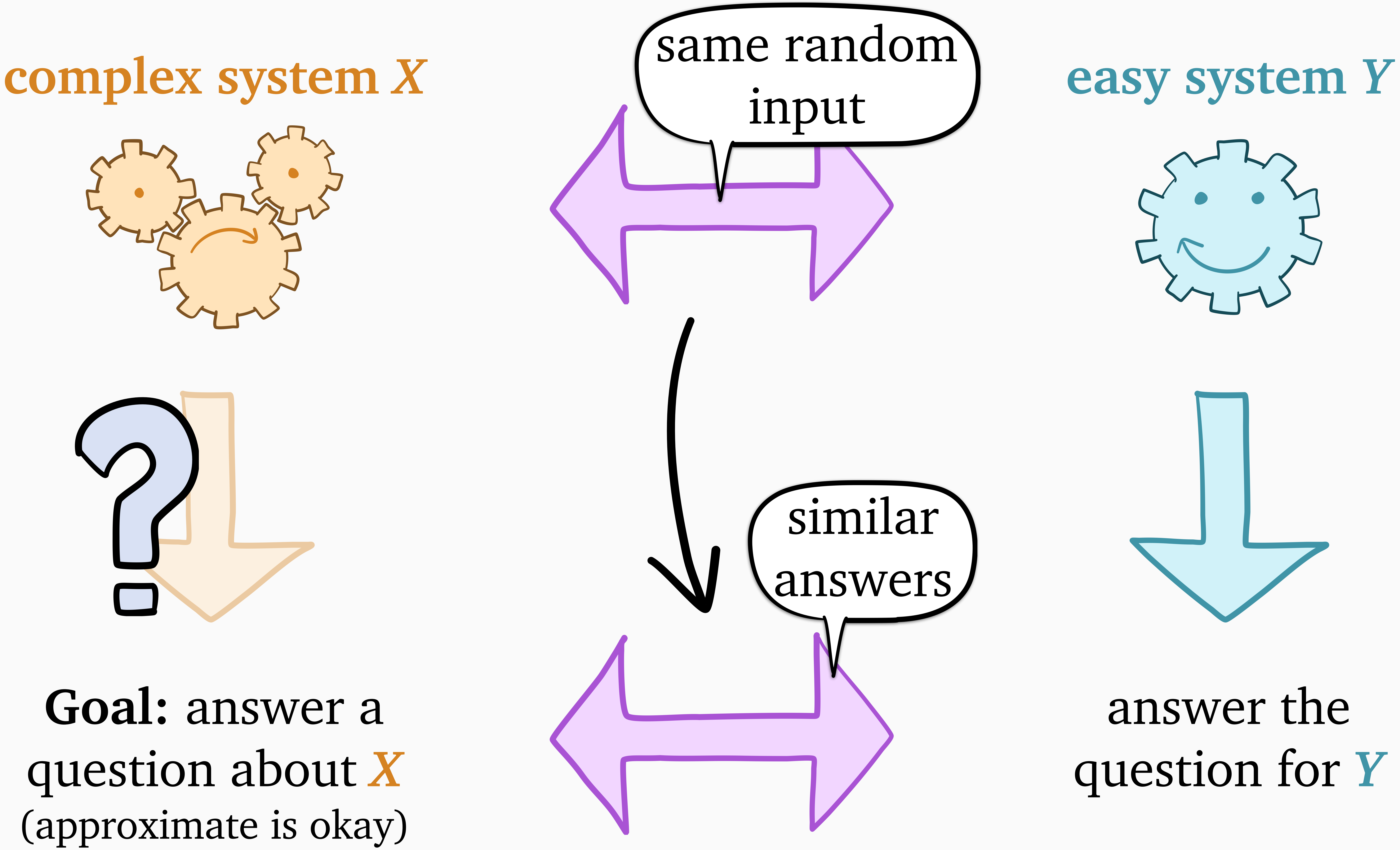


Goal: answer a question about X
(approximate is okay)



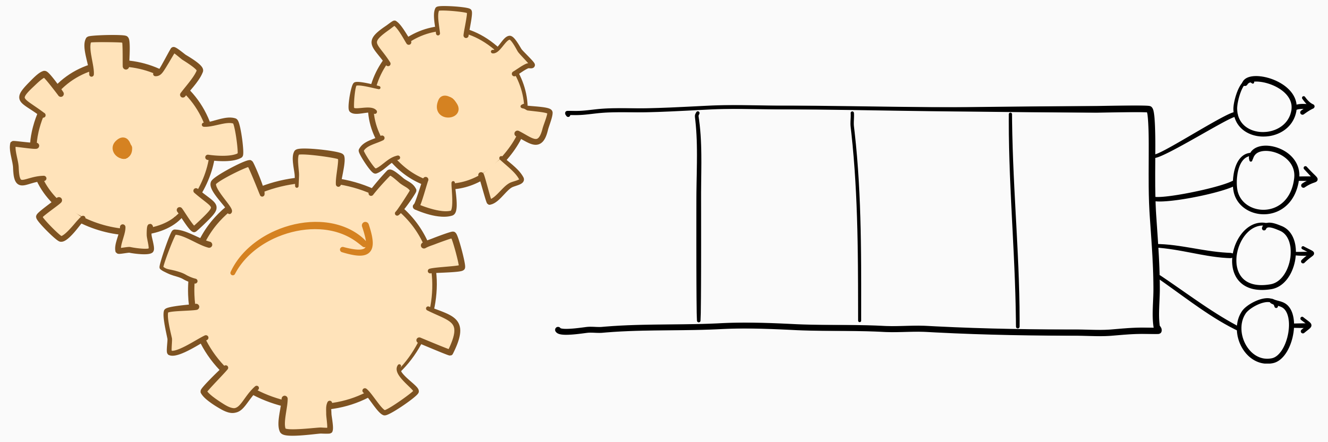
answer the question for Y

What is coupling?



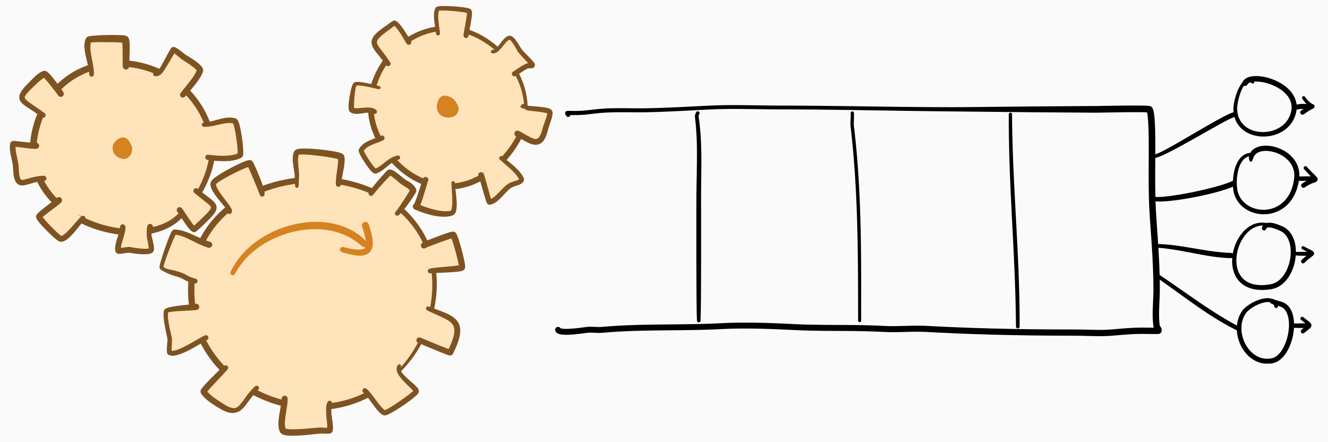
Example: M/M/k vs. M/M/1

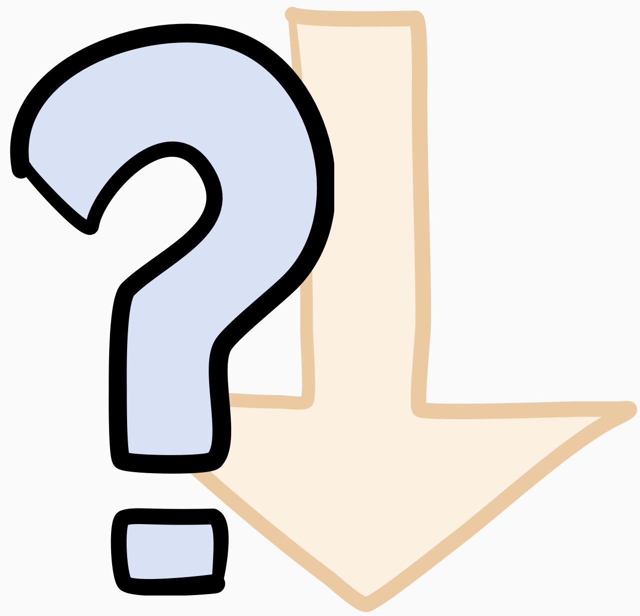
X = M/M/k



Example: M/M/k vs. M/M/1

$X = \text{M/M/}k$



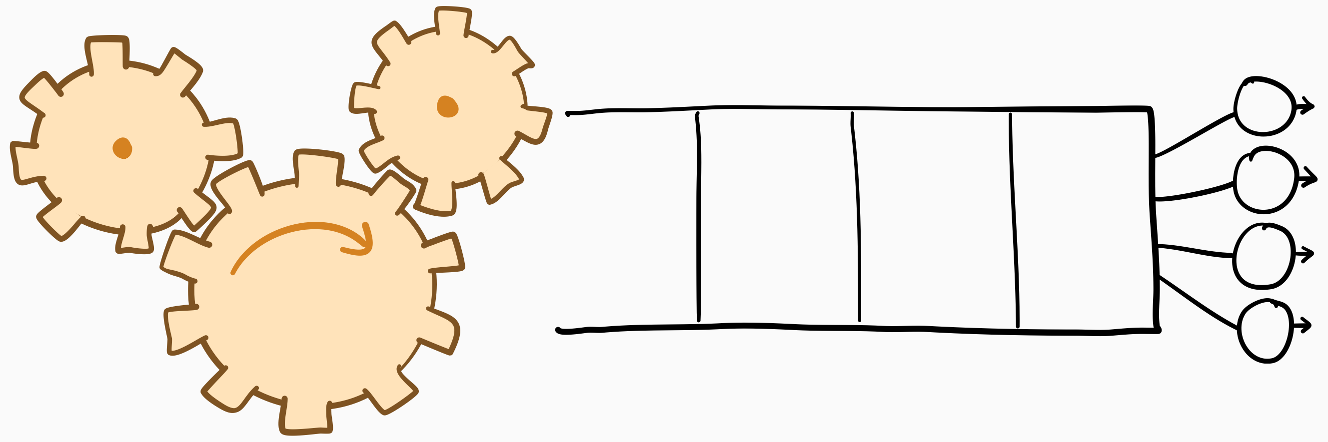


number of jobs

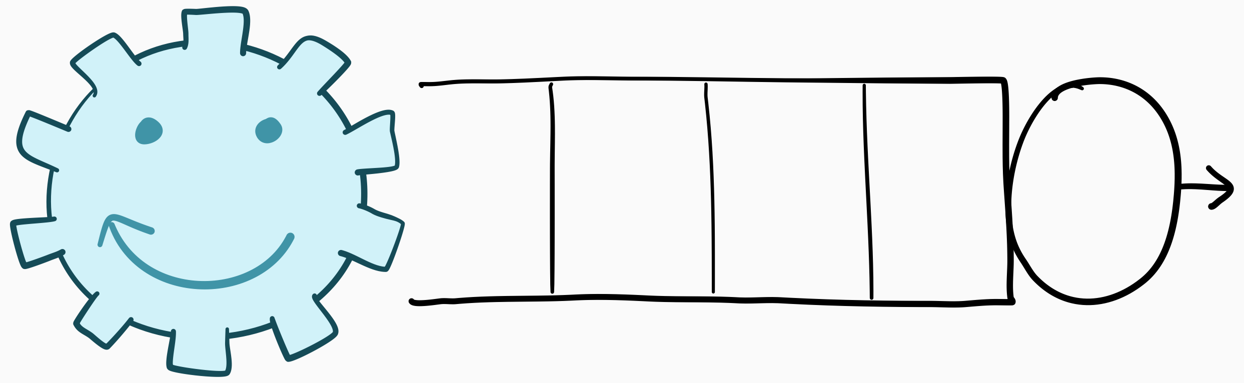
$\mathbf{E}[N_k] = ???$

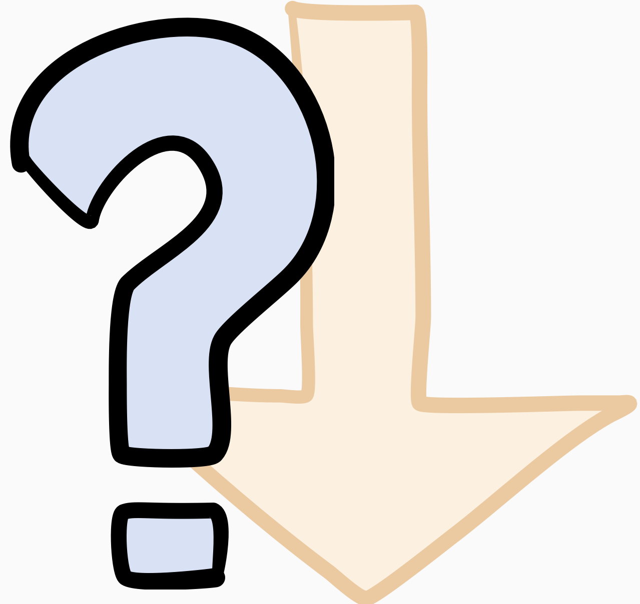
Example: M/M/k vs. M/M/1

X = M/M/k



Y = M/M/1



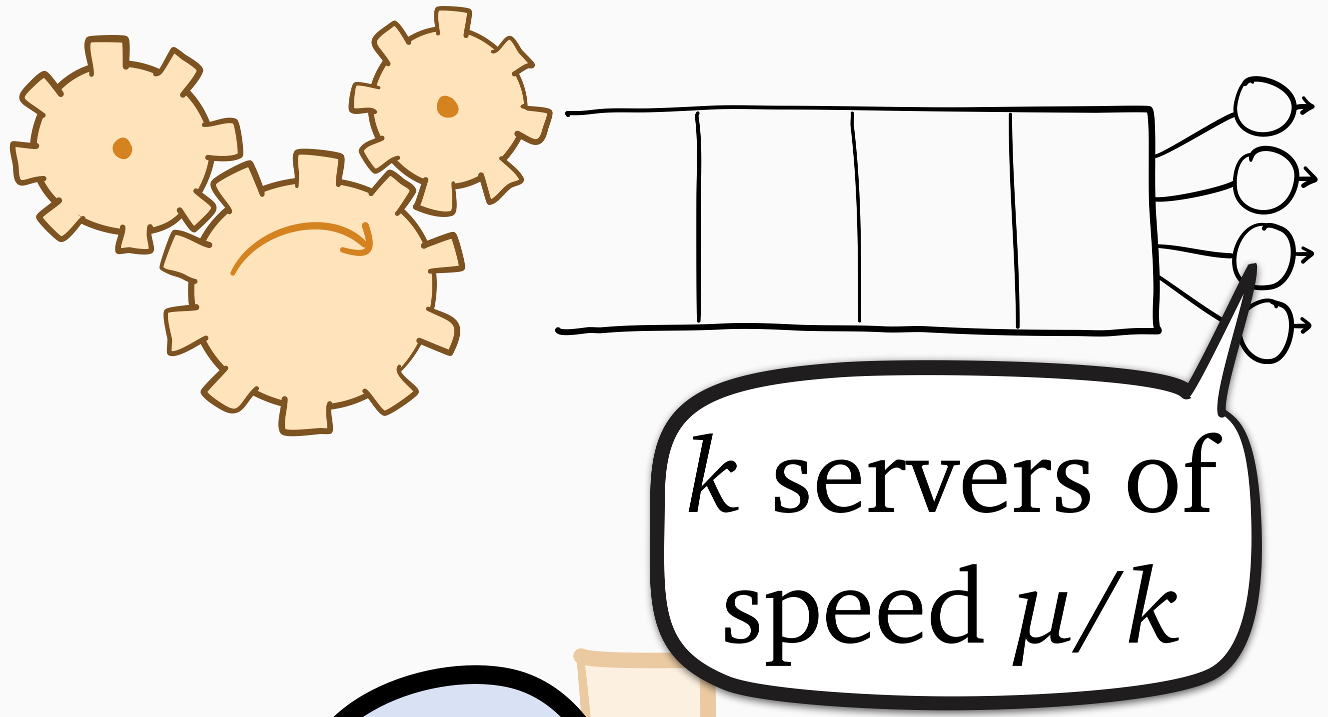


number of jobs

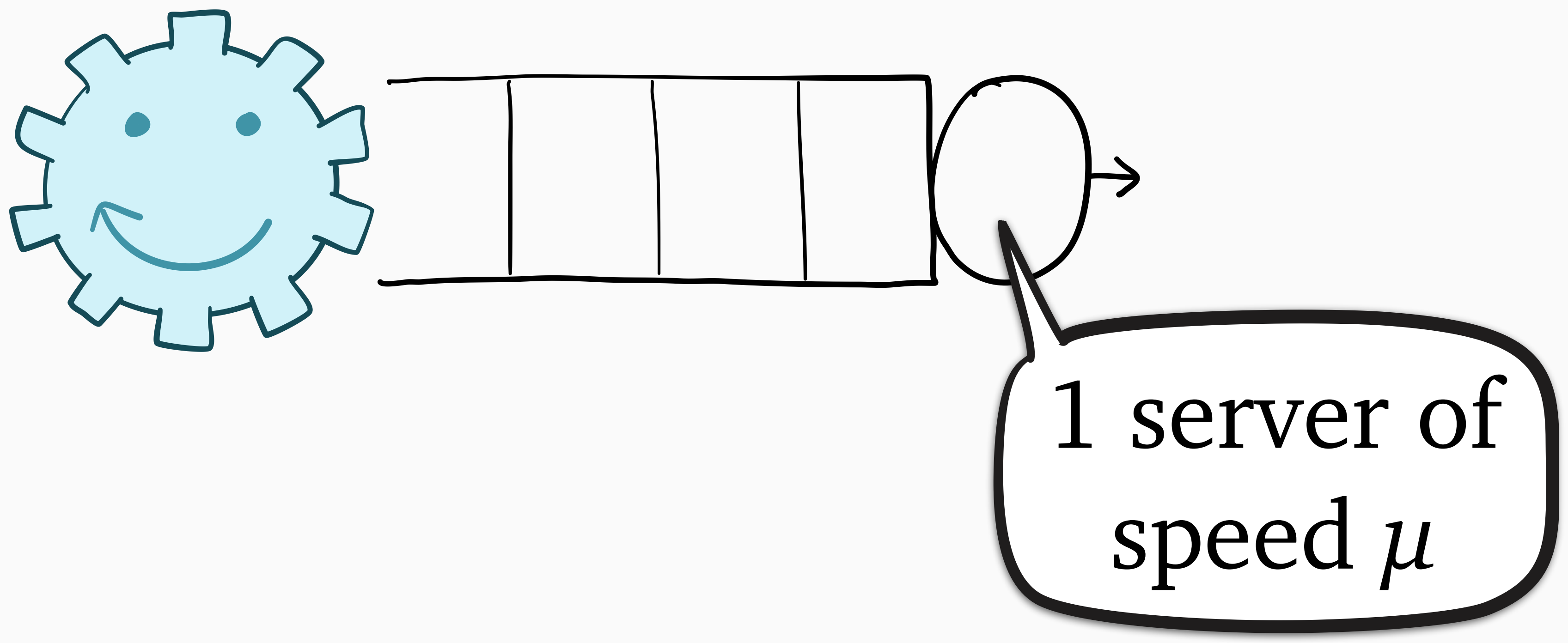
$$\mathbf{E}[N_k] = ???$$

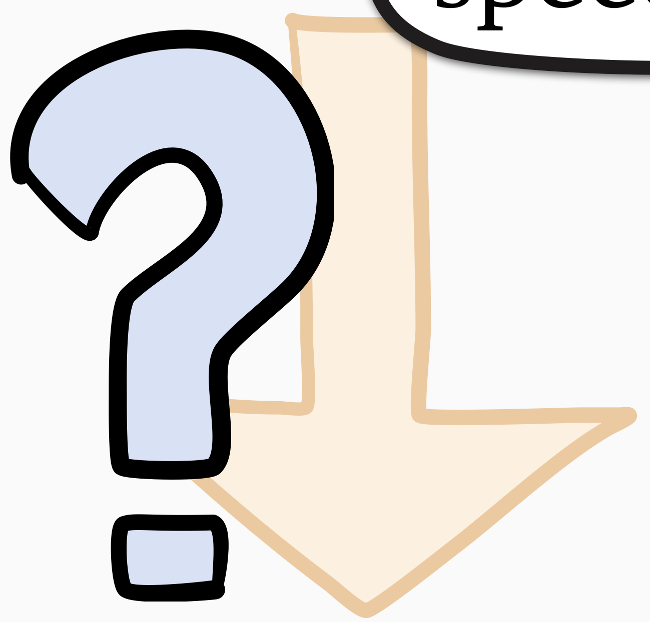
Example: M/M/k vs. M/M/1

$X =$ M/M/k



$Y =$ M/M/1

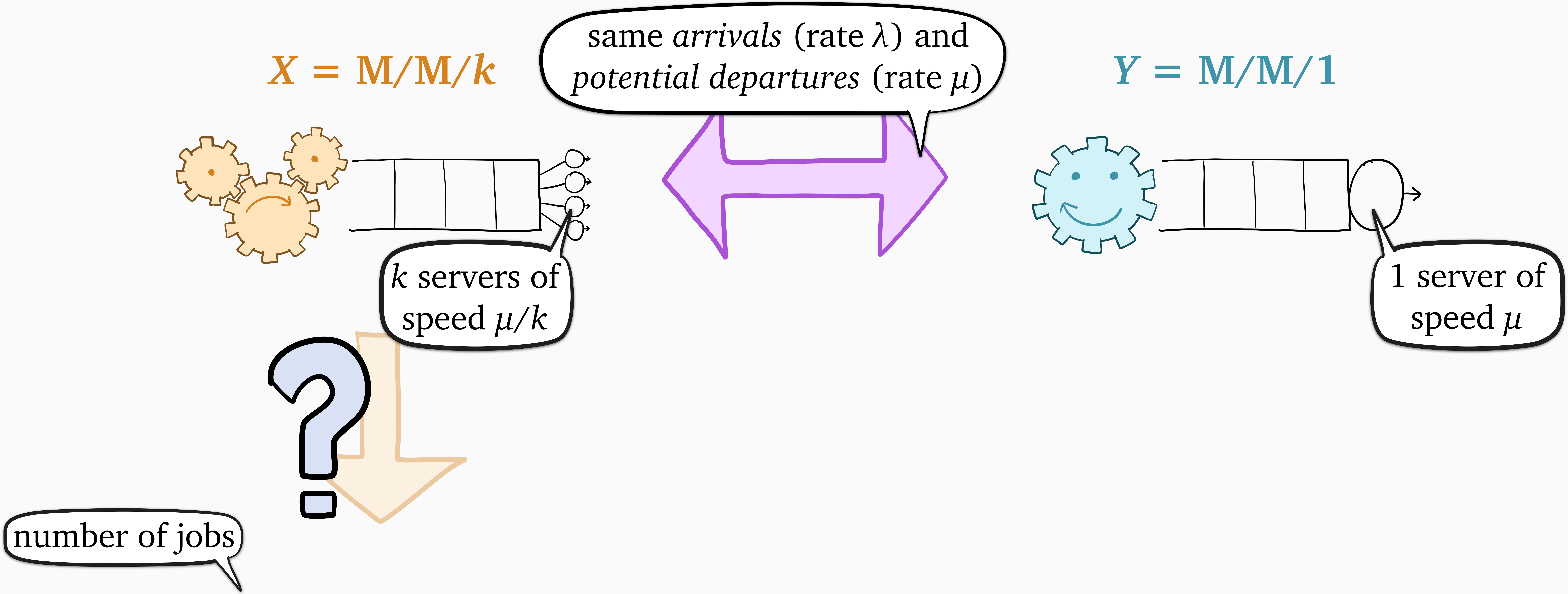




number of jobs

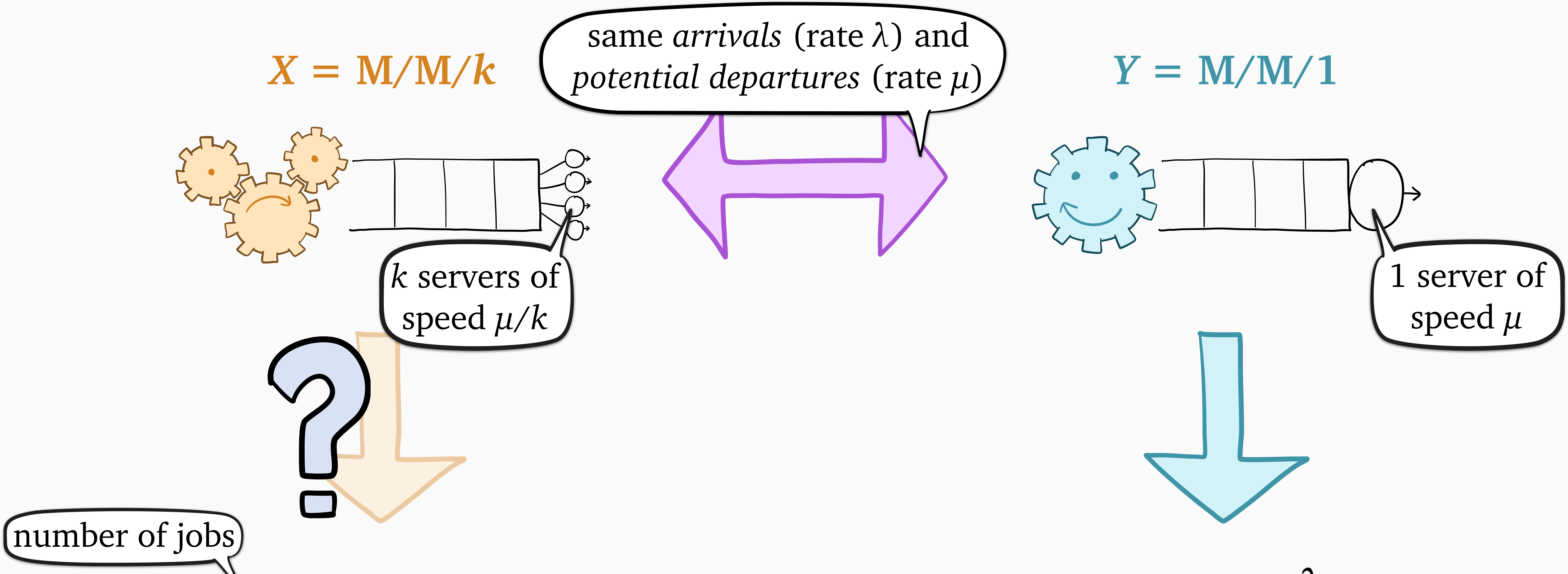
$\mathbf{E}[N_k] = ???$

Example: M/M/k vs. M/M/1



$$\mathbf{E}[N_k] = ???$$

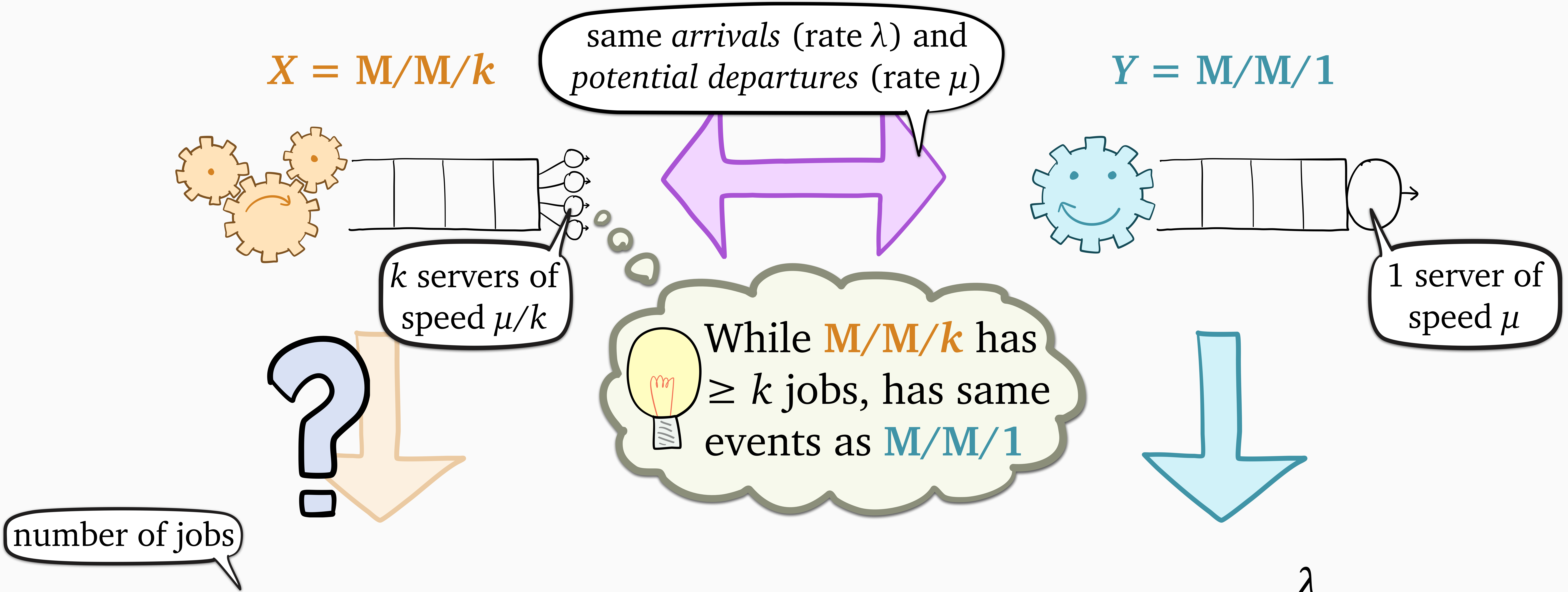
Example: M/M/k vs. M/M/1



$$\mathbf{E}[N_k] = ???$$

$$\mathbf{E}[N_1] = \frac{\lambda}{\mu - \lambda}$$

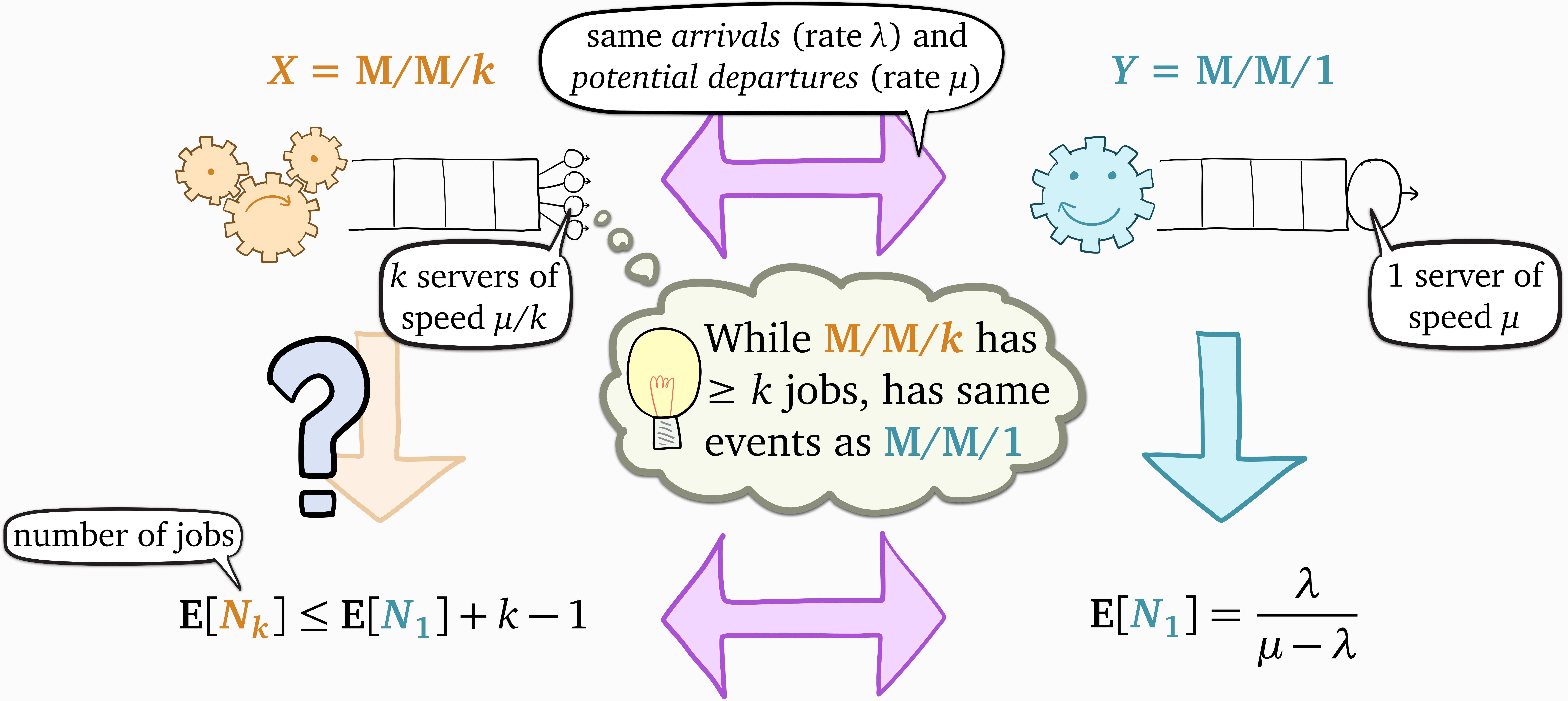
Example: M/M/k vs. M/M/1



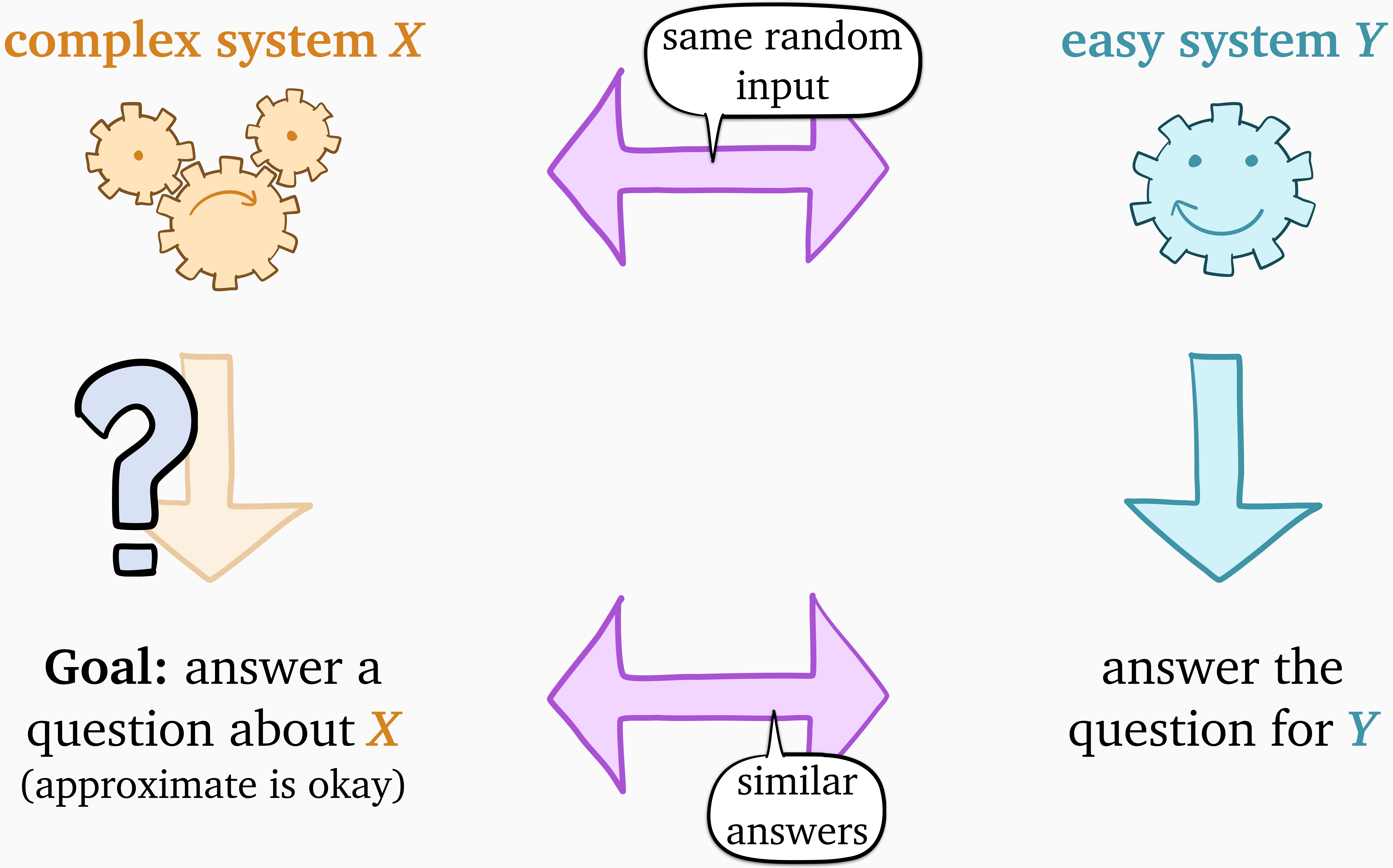
$\mathbf{E}[N_k] = ???$

$$\mathbf{E}[N_1] = \frac{\lambda}{\mu - \lambda}$$

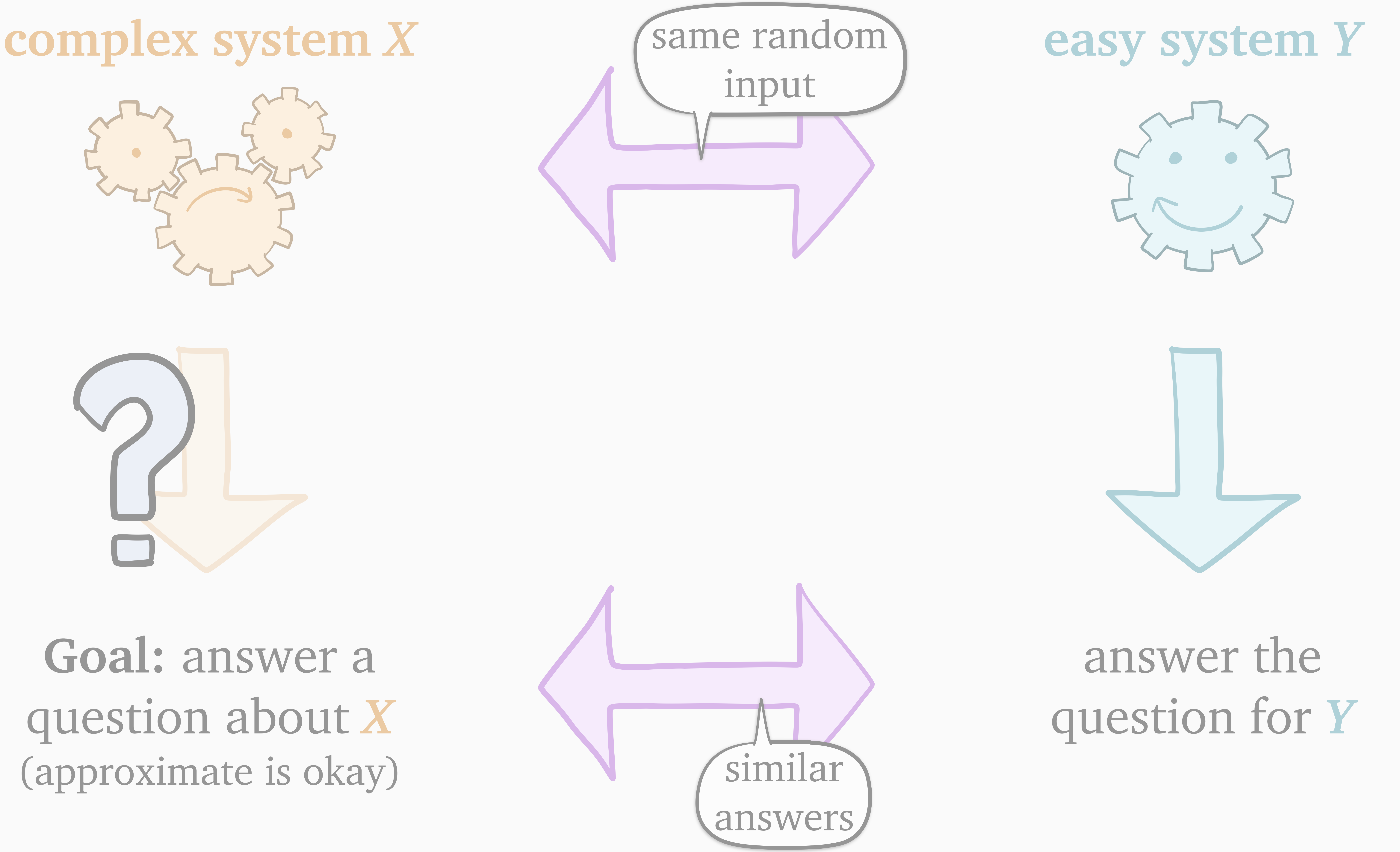
Example: M/M/k vs. M/M/1



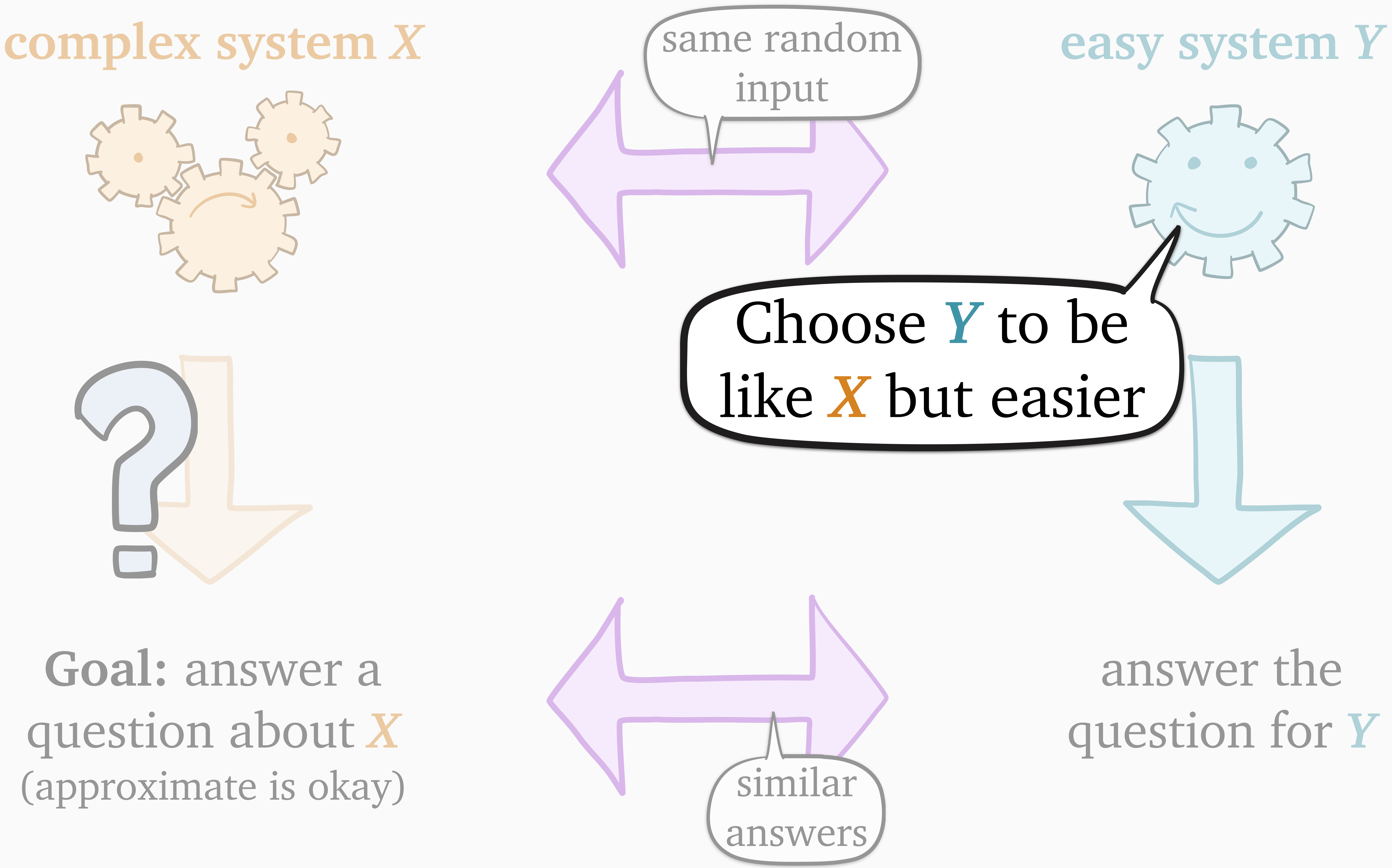
Key steps of coupling arguments



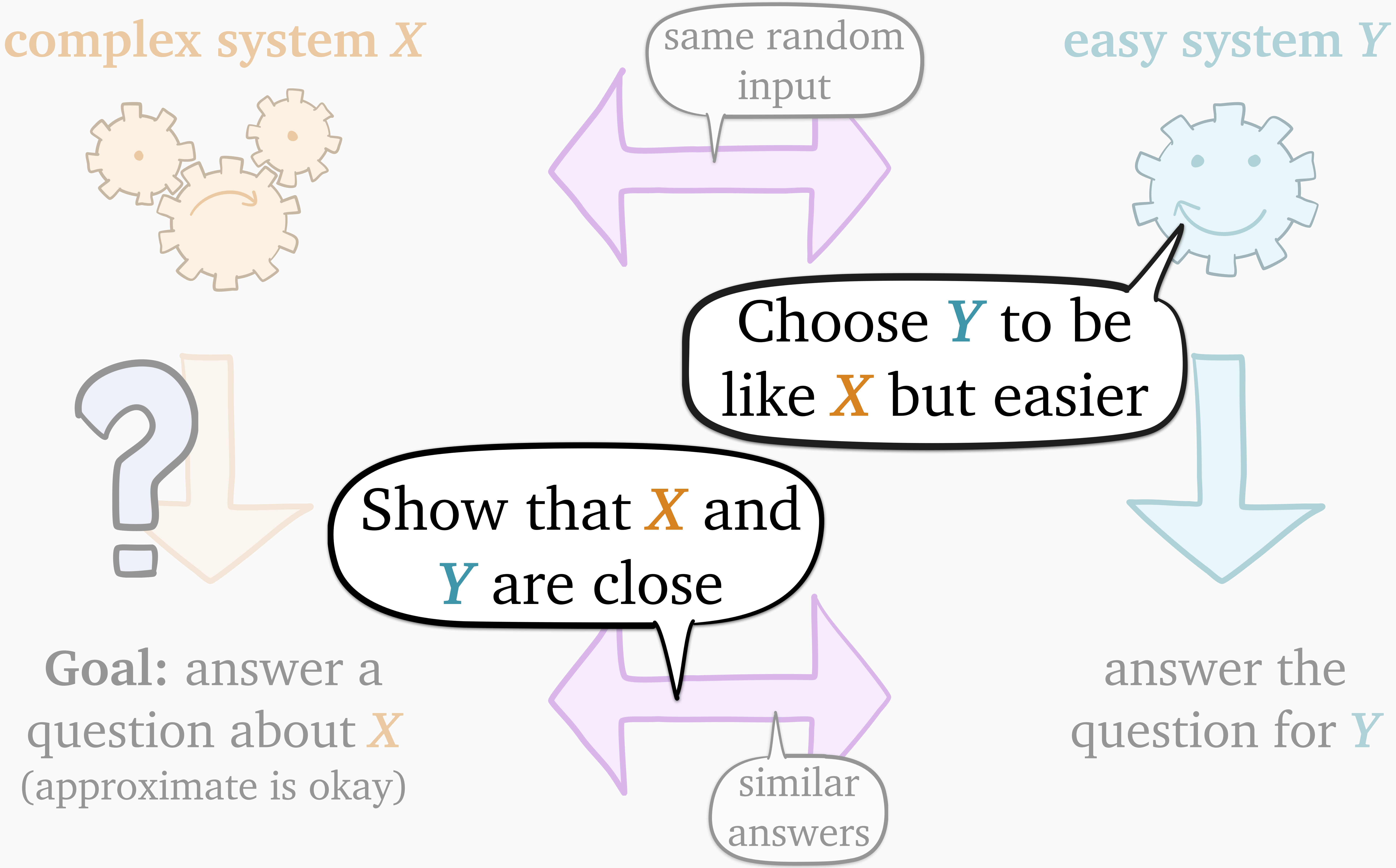
Key steps of coupling arguments



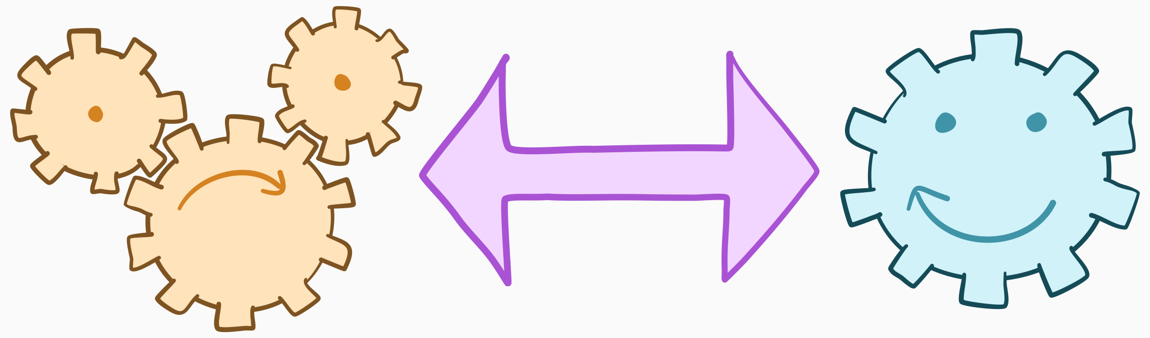
Key steps of coupling arguments



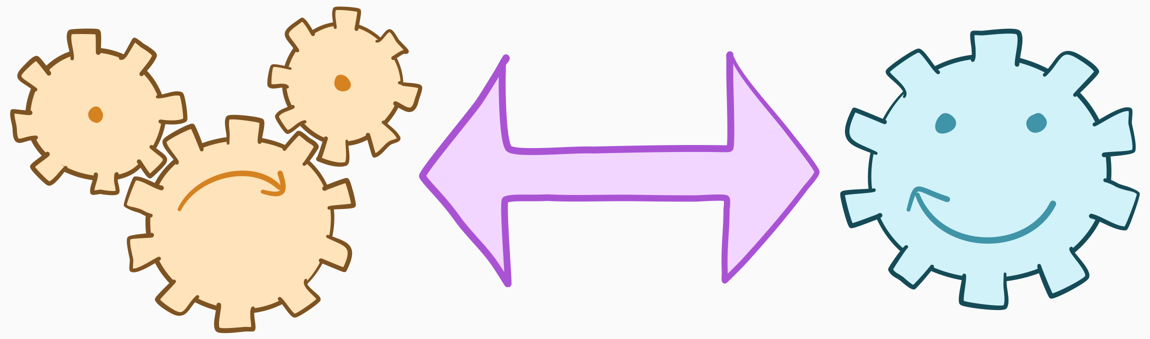
Key steps of coupling arguments



Classifying coupling techniques

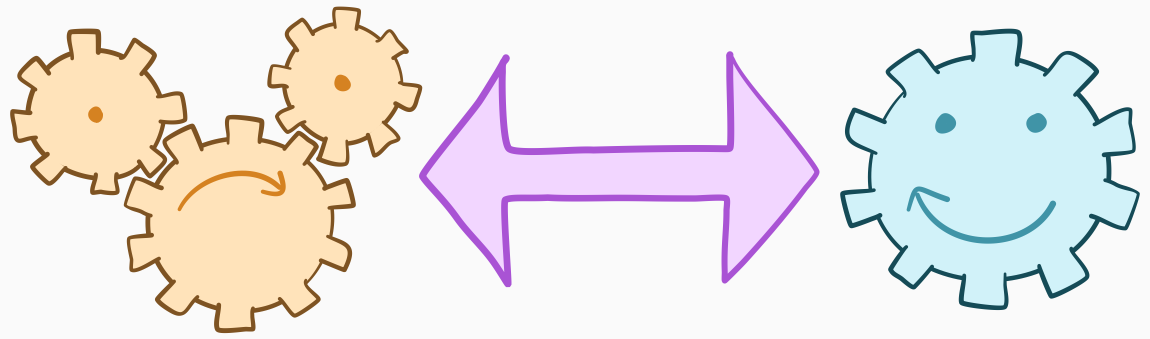


Classifying coupling techniques



How does ***Y*** make ***X*** easier?

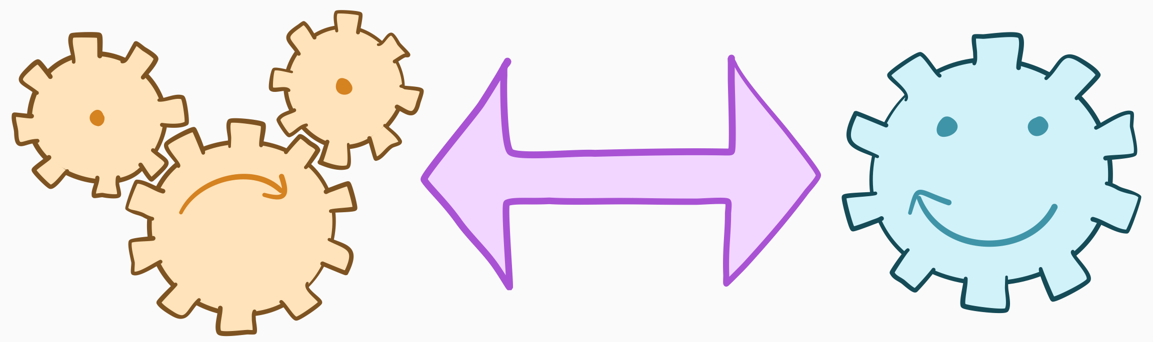
Classifying coupling techniques



In what sense are X and Y close?

How does Y make X easier?

Classifying coupling techniques

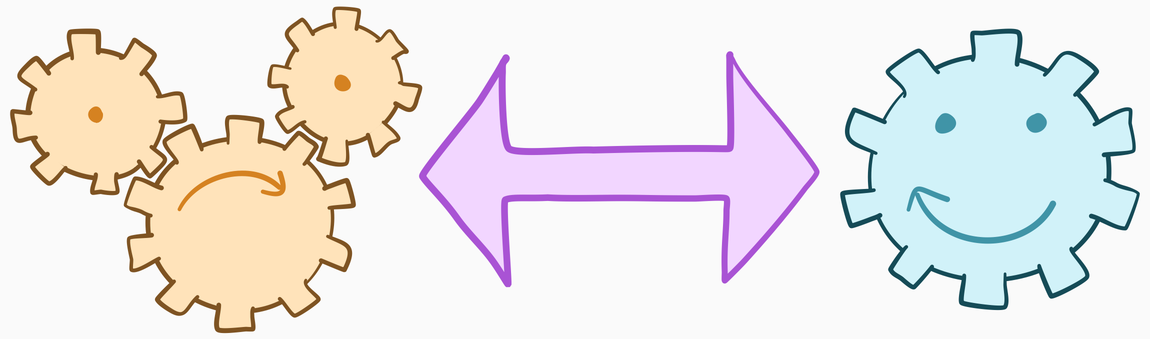


1. More information

In what sense are X and Y close?

How does Y make X easier?

Classifying coupling techniques



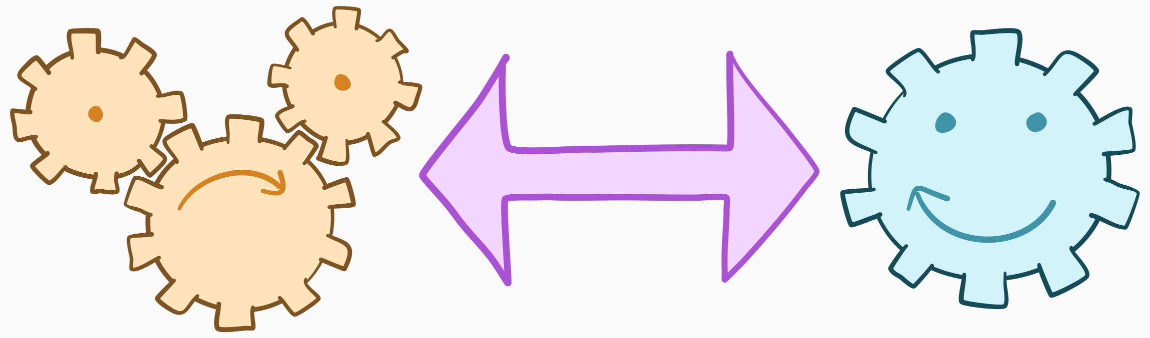
1. More information

In what sense are X and Y close?

2. Fewer constraints

How does Y make X easier?

Classifying coupling techniques

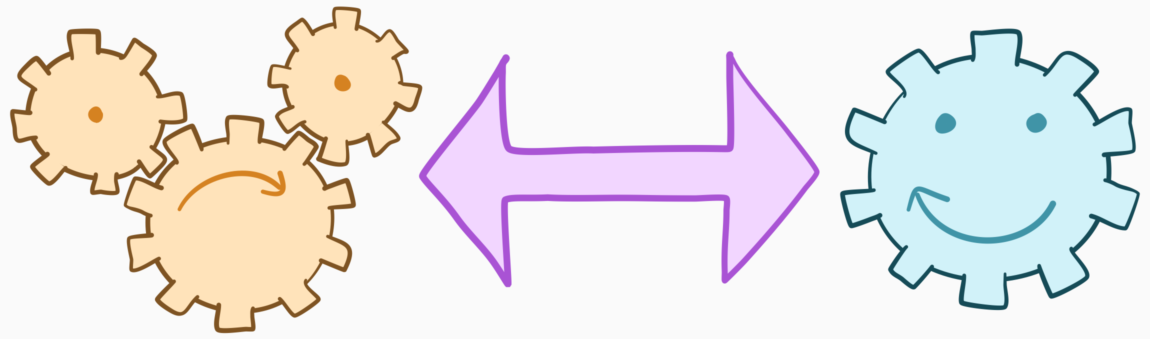


1. More information
2. Fewer constraints
3. Simpler dynamics

In what sense are *X* and *Y* close?

How does *Y* make *X* easier?

Classifying coupling techniques



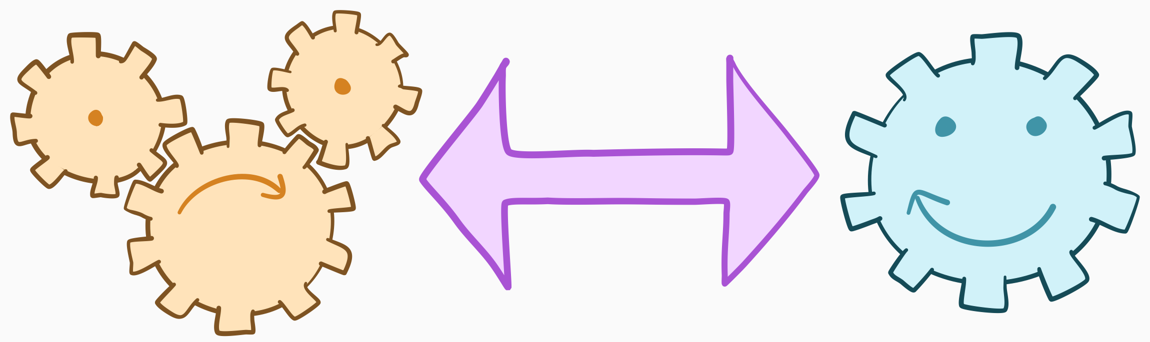
A. Every sample path

1. More information
2. Fewer constraints
3. Simpler dynamics

In what sense are X and Y close?

How does Y make X easier?

Classifying coupling techniques



A. Every sample path

B. Steady-state distribution

1. More information

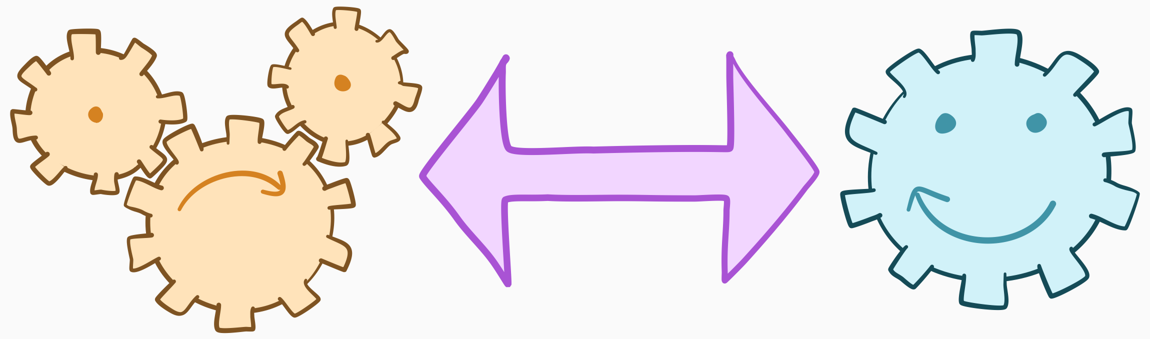
2. Fewer constraints

3. Simpler dynamics

In what sense are X and Y close?

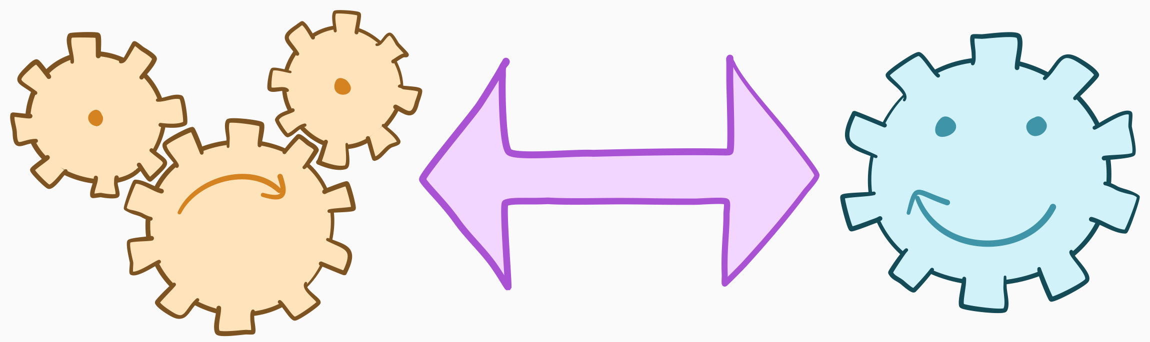
How does Y make X easier?

Classifying coupling techniques



	A. Every sample path	B. Steady-state distribution
1. More information	A1	B1
2. Fewer constraints	A2	B2
3. Simpler dynamics	A3	B3

Classifying coupling techniques



	A. Every sample path	B. Steady-state distribution
1. More information	A1	B1
2. Fewer constraints	A2 M/M/k vs. M/M/1	B2
3. Simpler dynamics	A3	B3

Overview

Part 1

Part 2

Overview

Part 1

Part 2

Survey 1:
Sample-Path Coupling

Overview

Part 1

Survey 1:
Sample-Path Coupling

Part 2

Survey 2:
Steady-State Coupling

Overview

Part 1

Survey 1:
Sample-Path Coupling

In-Depth Study 1:
Online Resource Allocation

Part 2

Survey 2:
Steady-State Coupling

Overview

Part 1

Survey 1:
Sample-Path Coupling

In-Depth Study 1:
Online Resource Allocation

Part 2

Survey 2:
Steady-State Coupling

In-Depth Study 2:
Gittins in the M/G/k

Overview

Part 1

Survey 1:
Sample-Path Coupling

In-Depth Study 1:
Online Resource Allocation

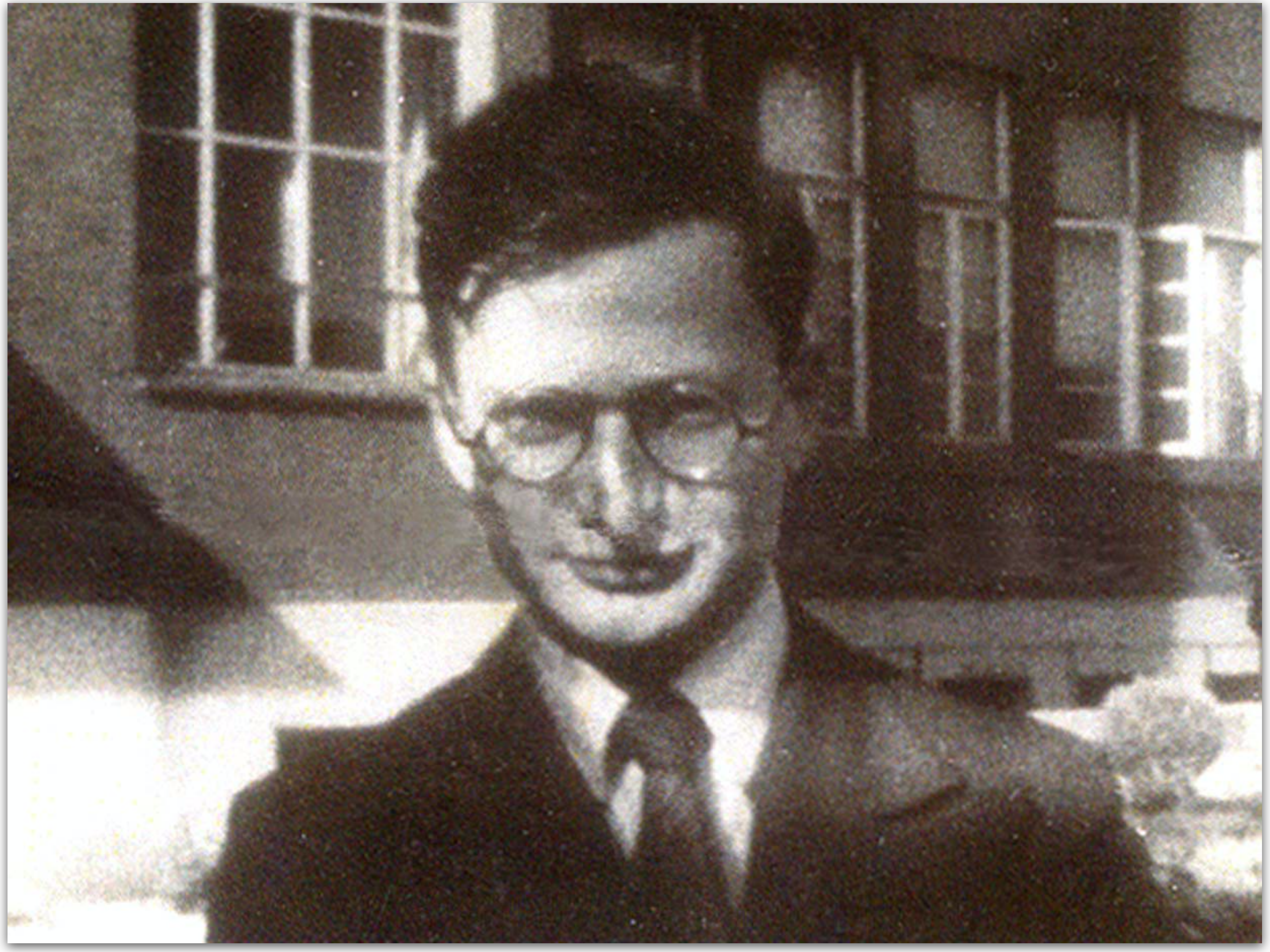
Part 2

Survey 2:
Steady-State Coupling

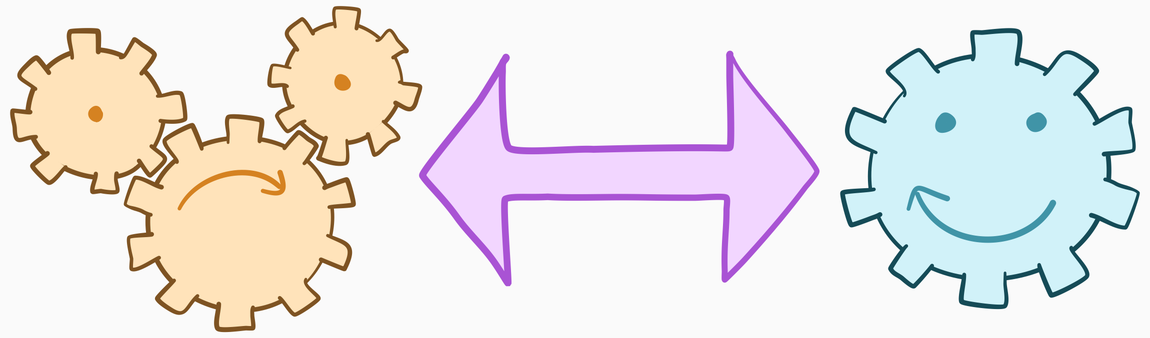
In-Depth Study 2:
Gittins in the M/G/k

Survey 1:
Sample-Path Coupling

Historical Note: Wolfgang Döblin

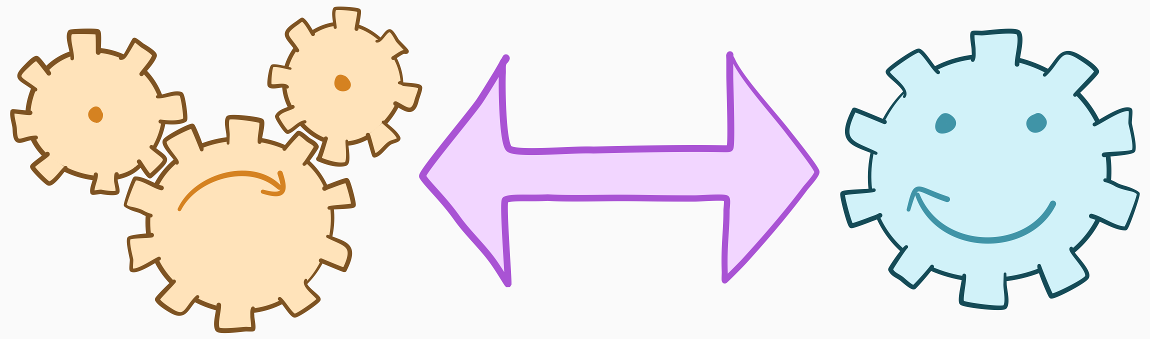


Classifying coupling techniques



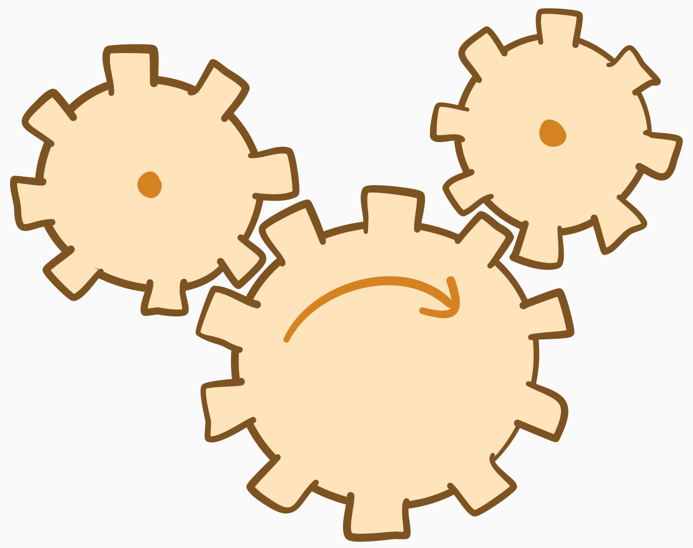
	A. Every sample path	B. Steady-state distribution
1. More information	A1	B1
2. Fewer constraints	A2 M/M/k vs. M/M/1	B2
3. Simpler dynamics	A3	B3

Classifying coupling techniques

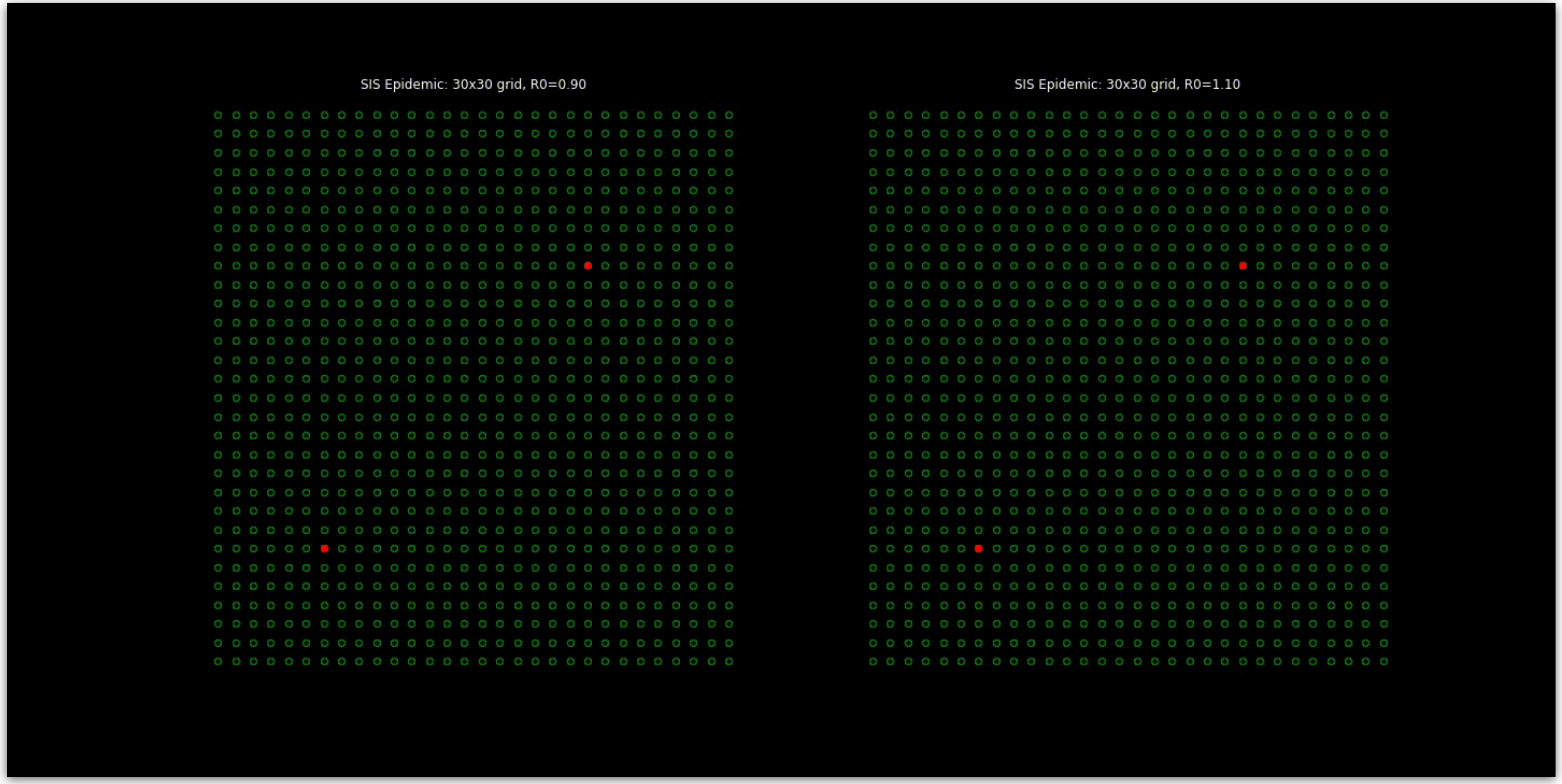


	A. Every sample path	B. Steady-state distribution
1. More information	A1	B1
2. Fewer constraints	A2 M/M/k vs. M/M/1	B2
3. Simpler dynamics	A3	B3

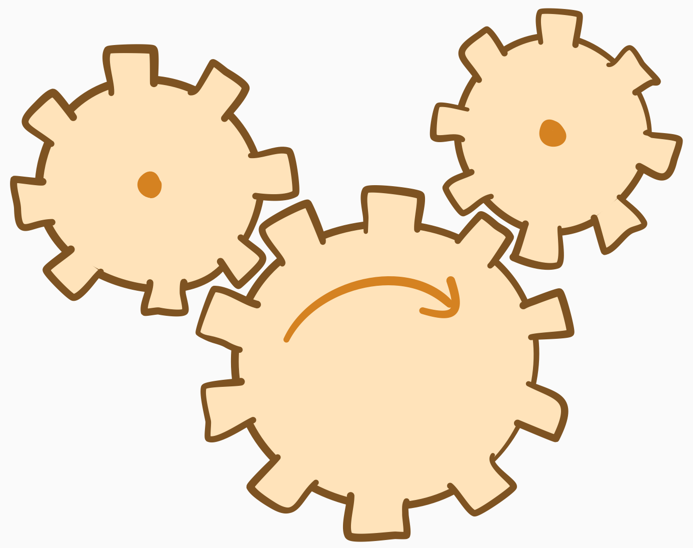
SIS epidemic model



- Set of people connected via social network graph G
- Each person has *infection level* in $\{0, 1\}$
- *Healing:* at each node, $1 \to 0$ at rate μ
- *Infection:* at each node, $0 \to 1$ at rate $\lambda \cdot$(# of neighboring 1's)



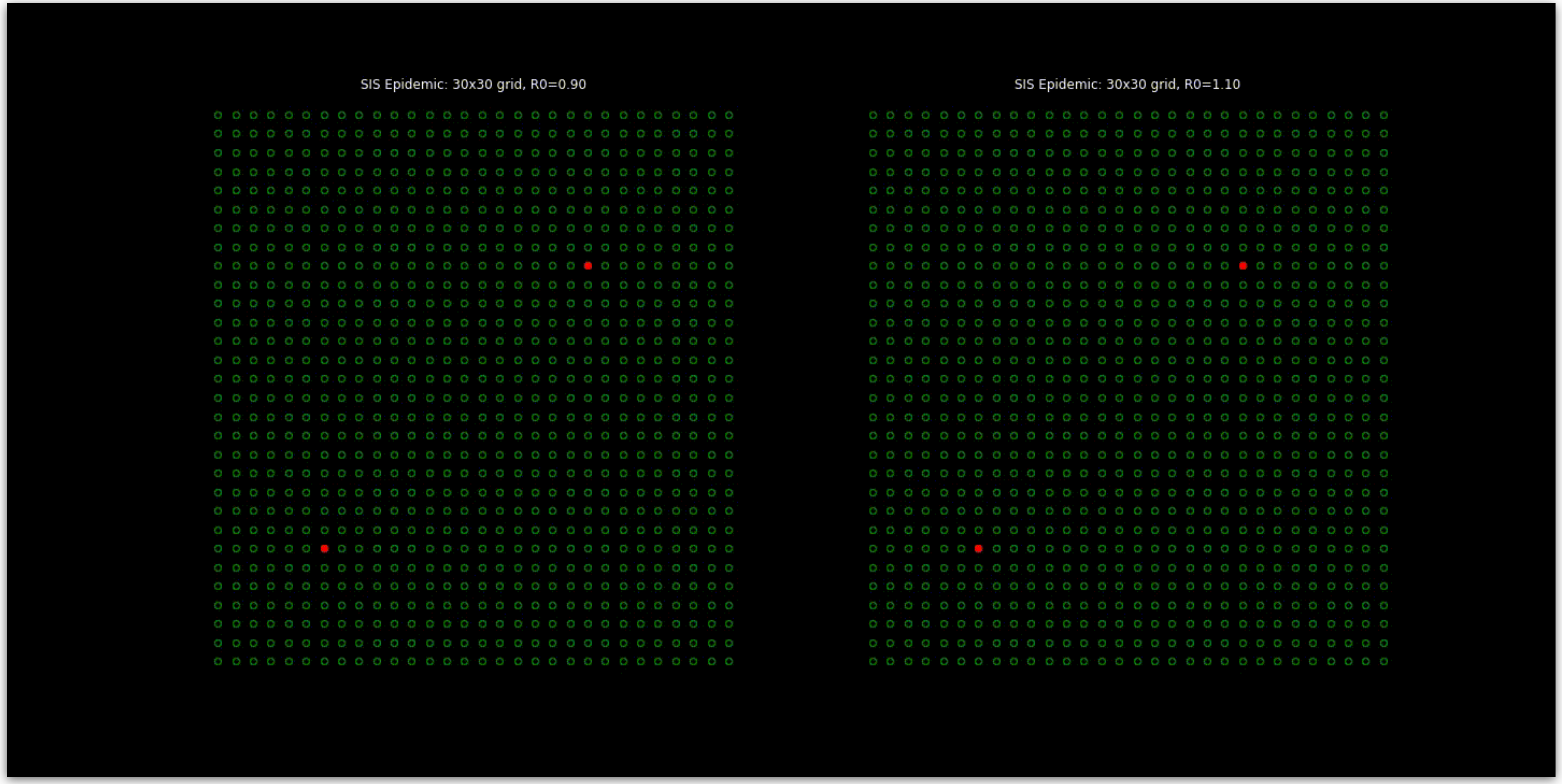
SIS epidemic model



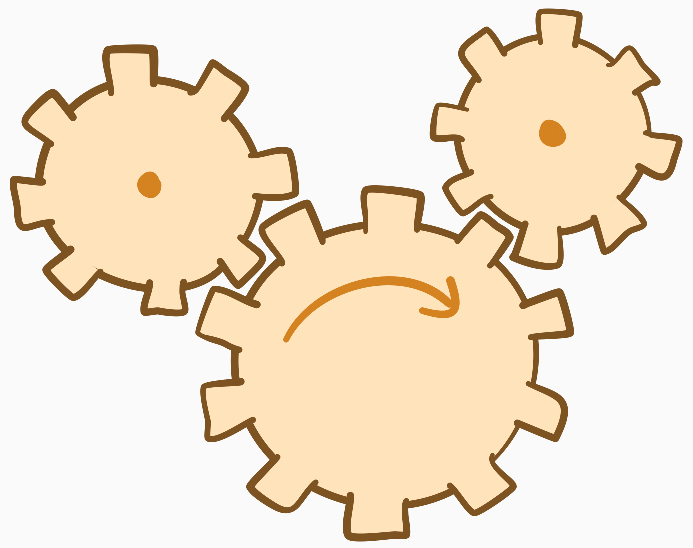
- Set of people connected … graph G
- Each person has *infection level* in {0, 1}
- *Healing:* at each node, $1 \to 0$ at rate μ
- *Infection:* at each node, $0 \to 1$ at rate $\lambda \cdot$(# of neighboring 1's)

susceptible

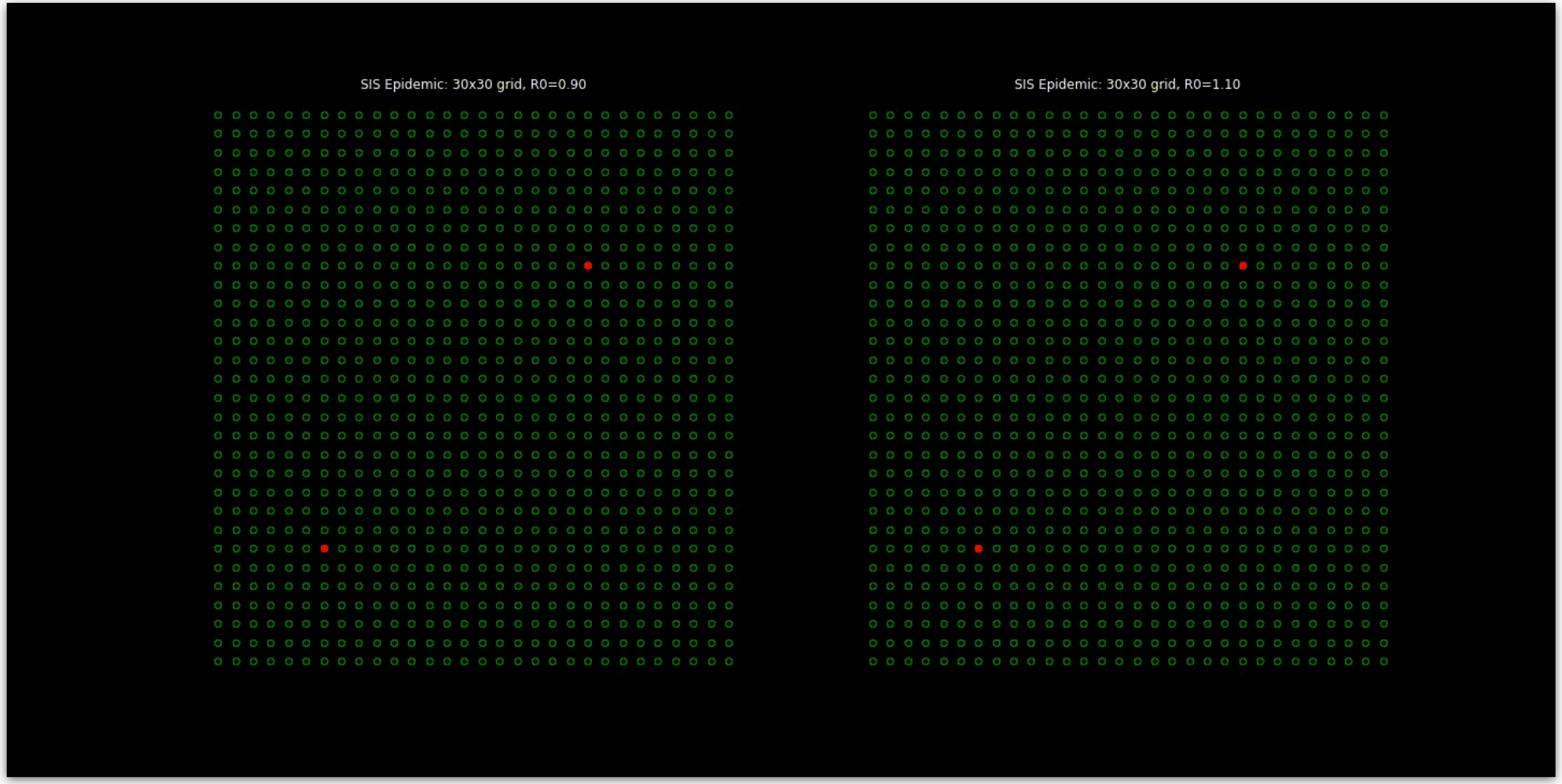
infected



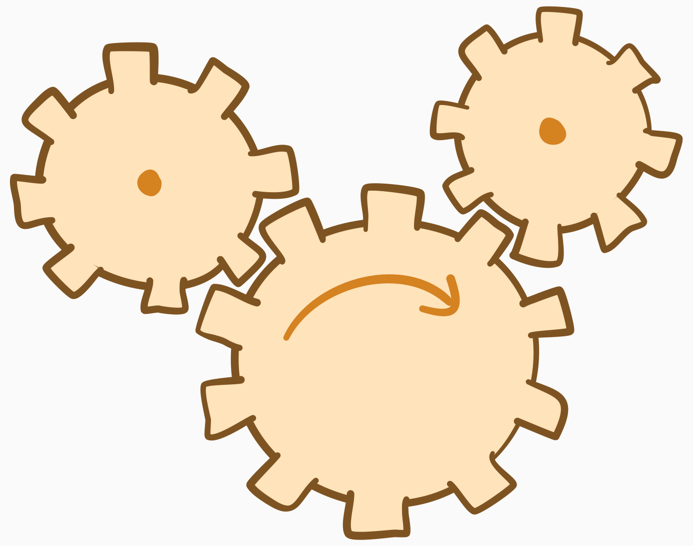
SIS epidemic model

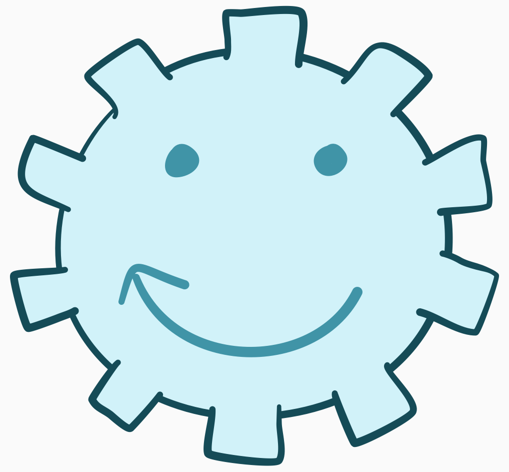


- Set of people connected net graph G susceptible infected
- Each person has *infection level* in {0, 1}
- *Healing:* at each node, $1 \to 0$ at rate μ
- *Infection:* at each node, $0 \to 1$ at rate $\lambda \cdot$(# of neighboring 1's)



SIS epidemic: lower bound

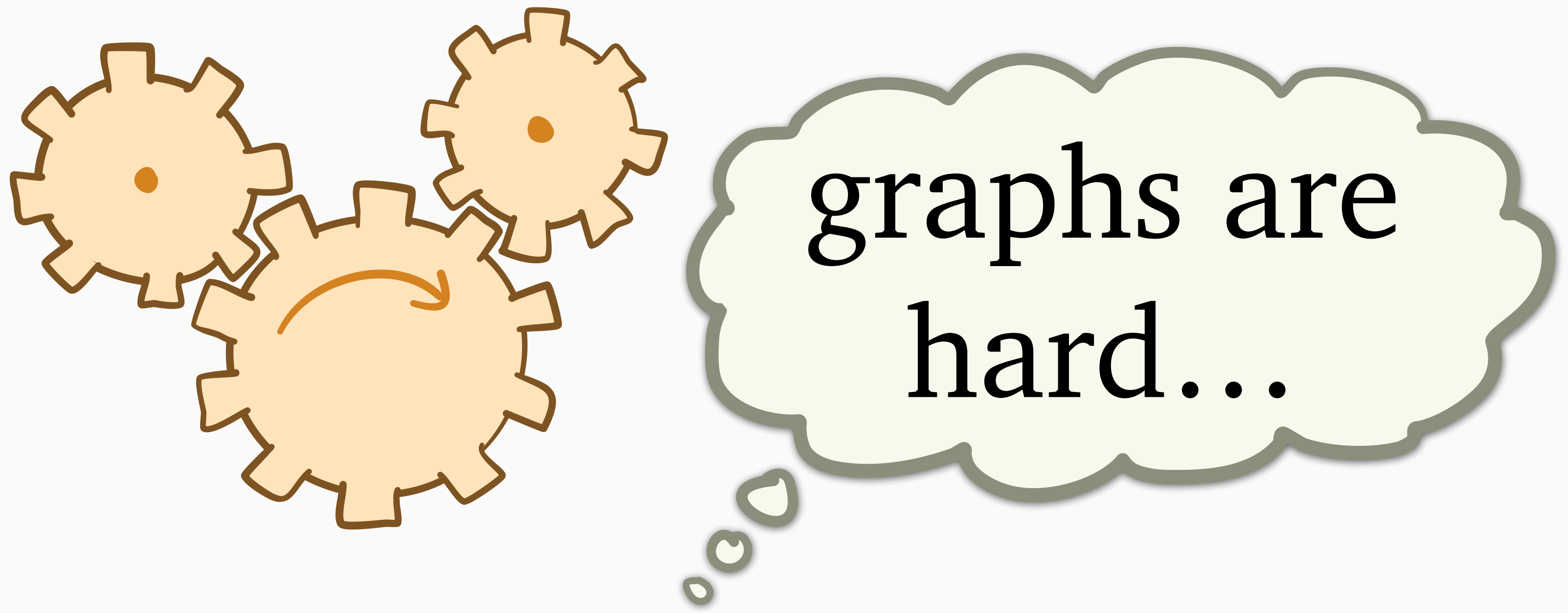


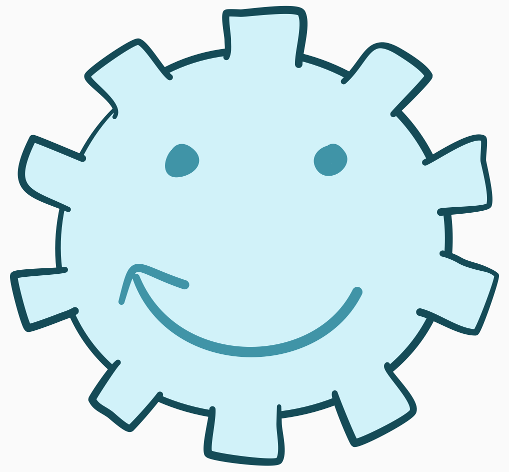


- Social network graph G
- Each person has *infection level* in $\{0, 1\}$
- *Healing:* at each node, $1 \to 0$ at rate μ
- *Infection:* at each node, $0 \to 1$ at rate $\lambda \cdot$(# of neighboring 1's)

Ganesh, Massoulié and Towsley: The effect of network topology on the spread of epidemics (2005)

SIS epidemic: lower bound

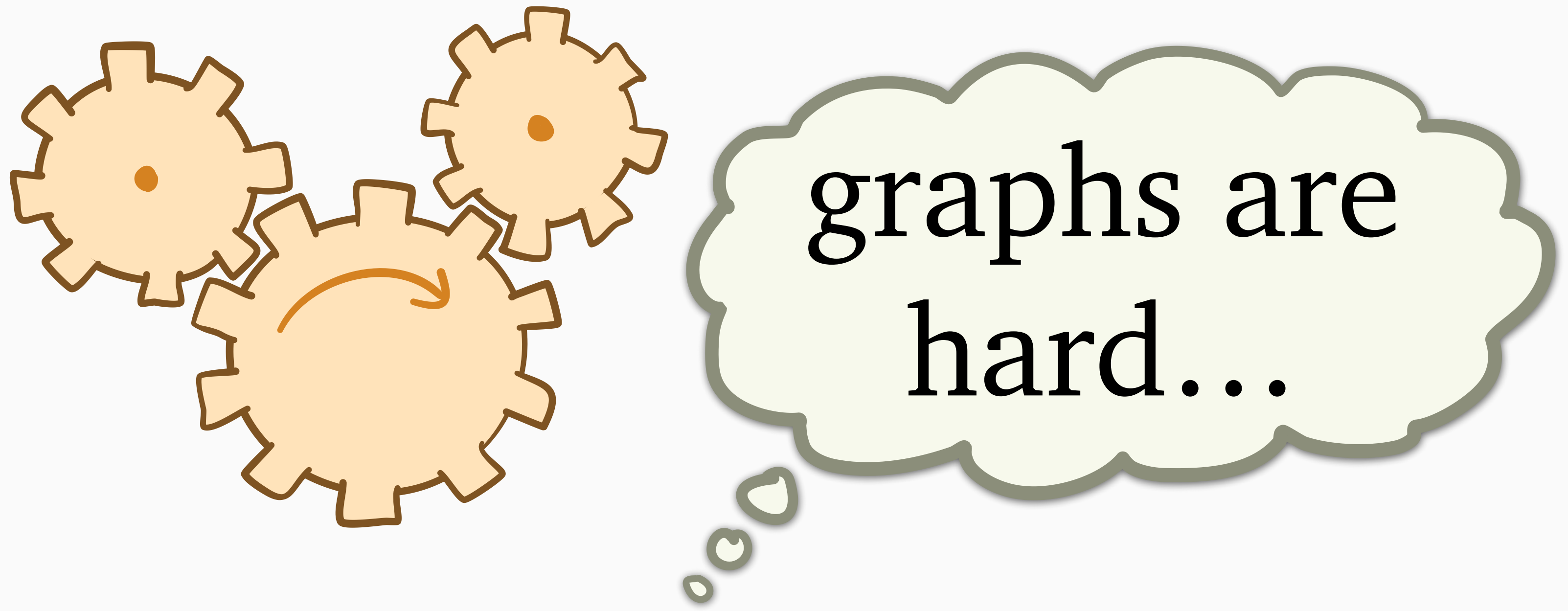




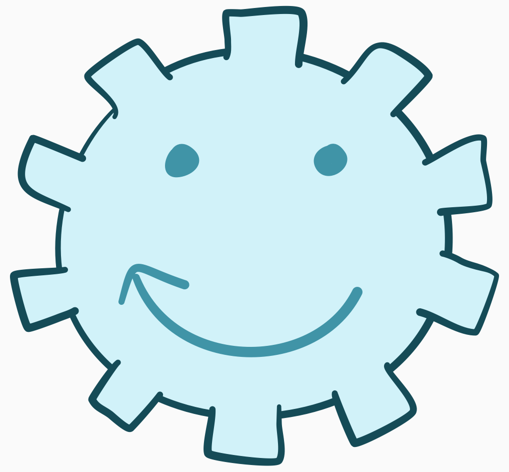
- Social network graph G
- Each person has *infection level* in {0, 1}
- *Healing:* at each node, $1 \to 0$ at rate μ
- *Infection:* at each node, $0 \to 1$ at rate $\lambda \cdot$(# of neighboring 1's)

Ganesh, Massoulié and Towsley: The effect of network topology on the spread of epidemics (2005)

SIS epidemic: lower bound



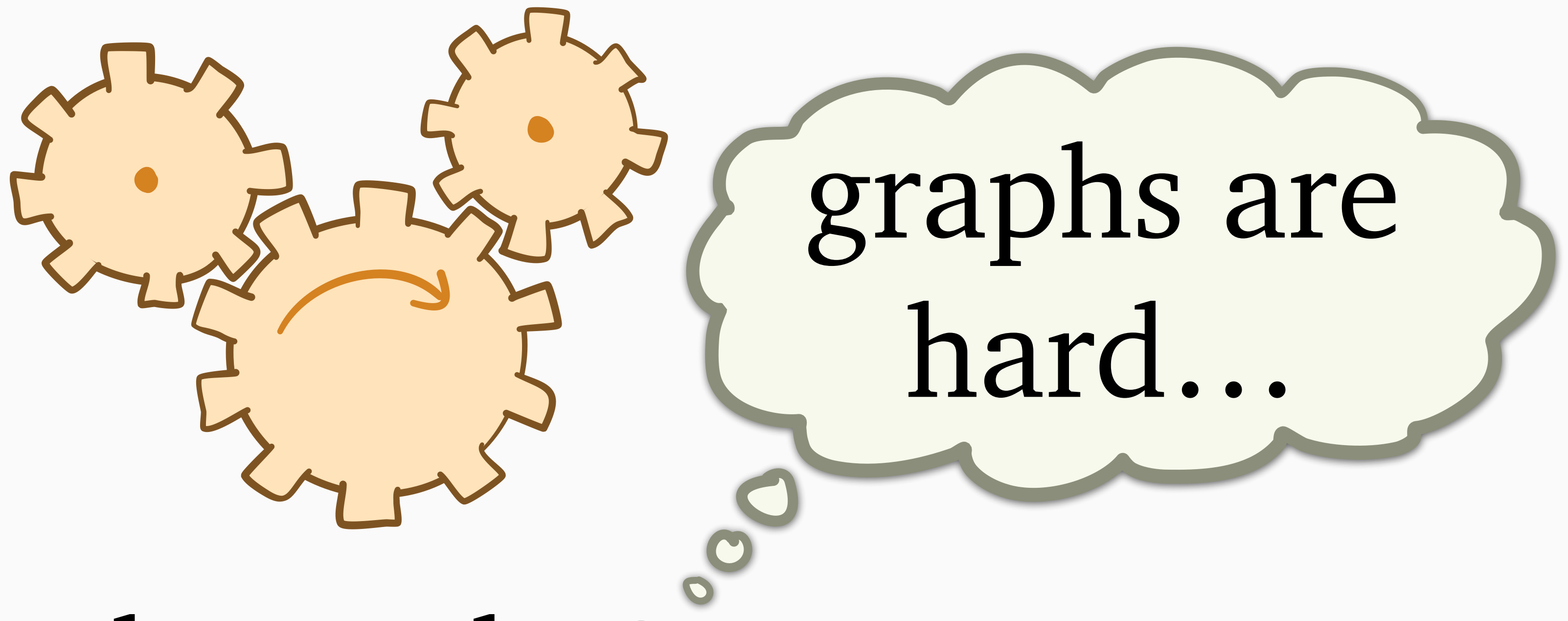
- Social network graph G
- Each person has *infection level* in {0, 1}
- *Healing:* at each node, $1 \rightarrow 0$ at rate μ
- *Infection:* at each node, $0 \rightarrow 1$ at rate $\lambda \cdot$(# of neighboring 1’s)



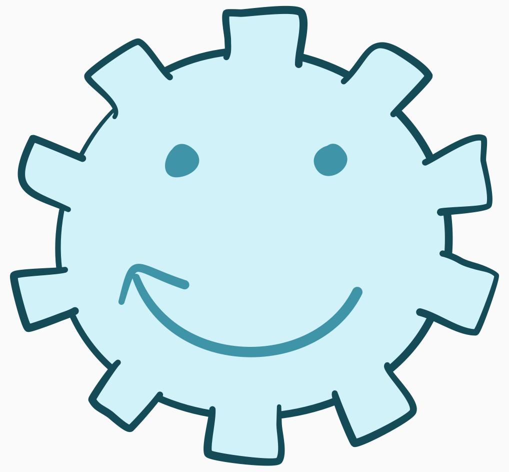
- Track *number*, not *set*, of infections

Ganesh, Massoulié and Towsley: The effect of network topology on the spread of epidemics (2005)

SIS epidemic: lower bound



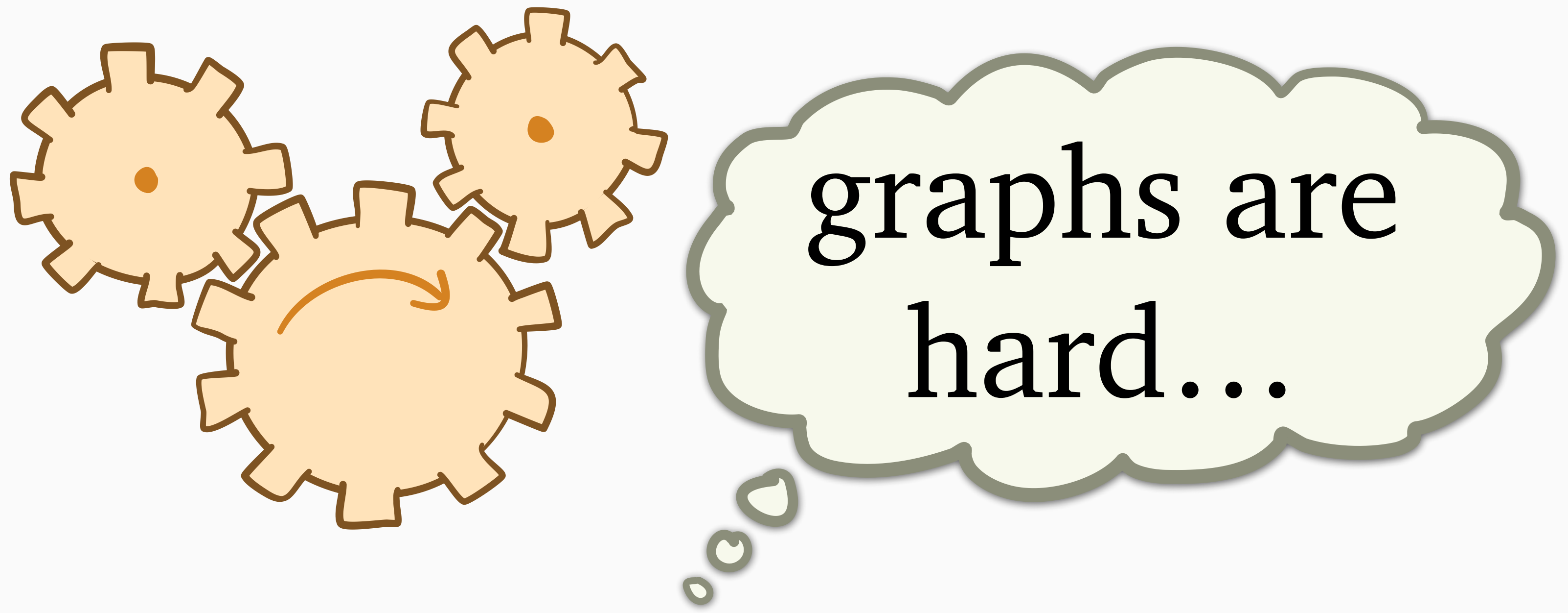
- Social network graph G
- Each person has *infection level* in $\{0, 1\}$
- *Healing:* at each node, $1 \to 0$ at rate μ
- *Infection:* at each node, $0 \to 1$ at rate $\lambda \cdot$(# of neighboring 1's)



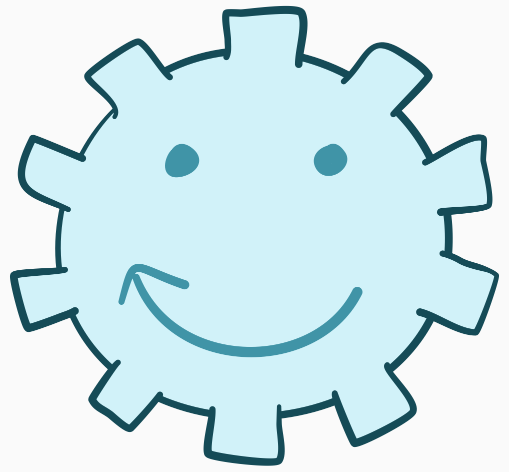
- Track *number*, not *set*, of infections
- k = number of people infected

Ganesh, Massoulié and Towsley: The effect of network topology on the spread of epidemics (2005)

SIS epidemic: lower bound



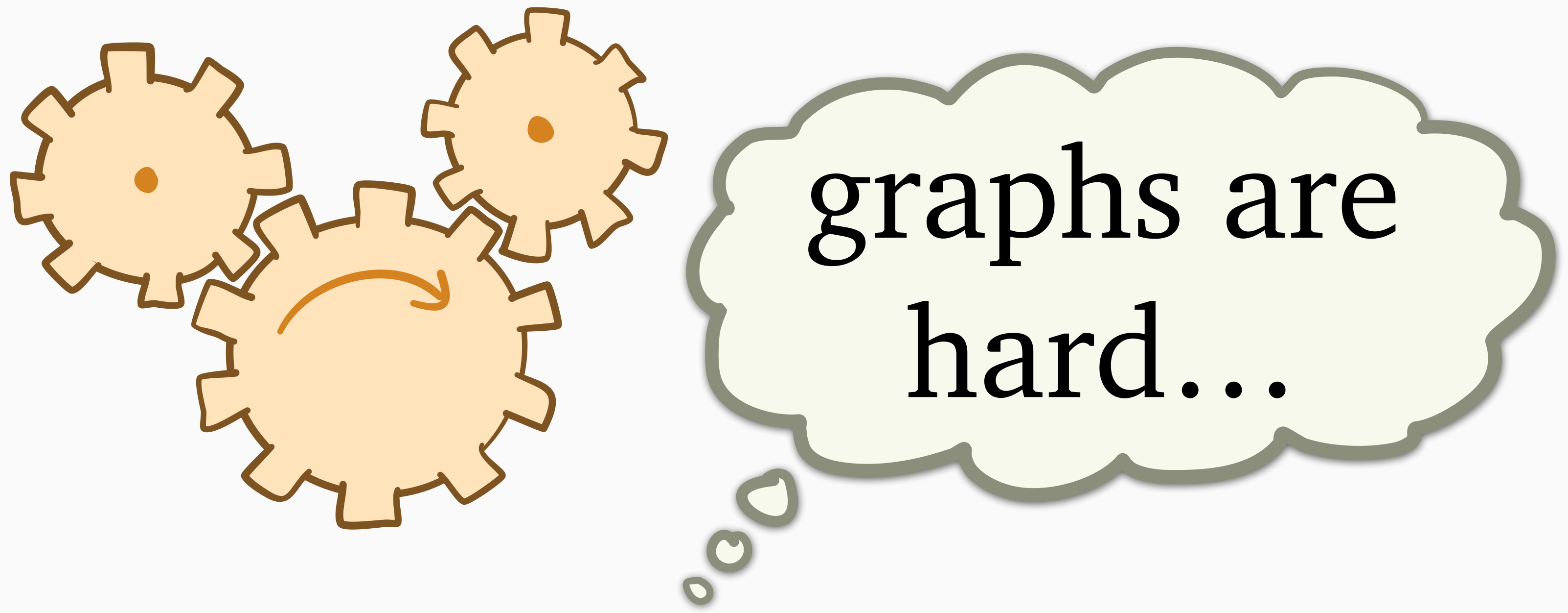
- Social network graph G
- Each person has *infection level* in $\{0, 1\}$
- *Healing:* at each node, $1 \to 0$ at rate μ
- *Infection:* at each node, $0 \to 1$ at rate $\lambda \cdot$(# of neighboring 1's)



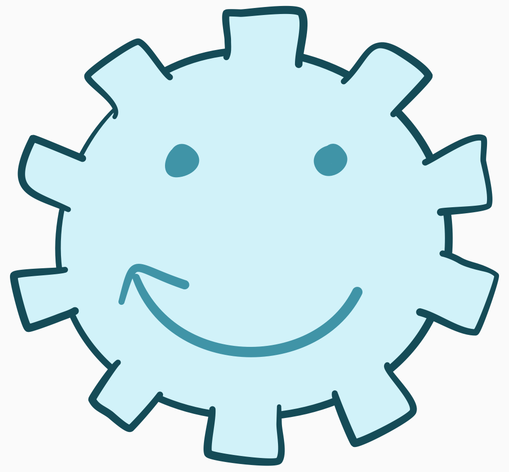
- Track *number*, not *set*, of infections
- k = number of people infected
- *Healing:* $k \to k - 1$ at rate μk

Ganesh, Massoulié and Towsley: The effect of network topology on the spread of epidemics (2005)

SIS epidemic: lower bound



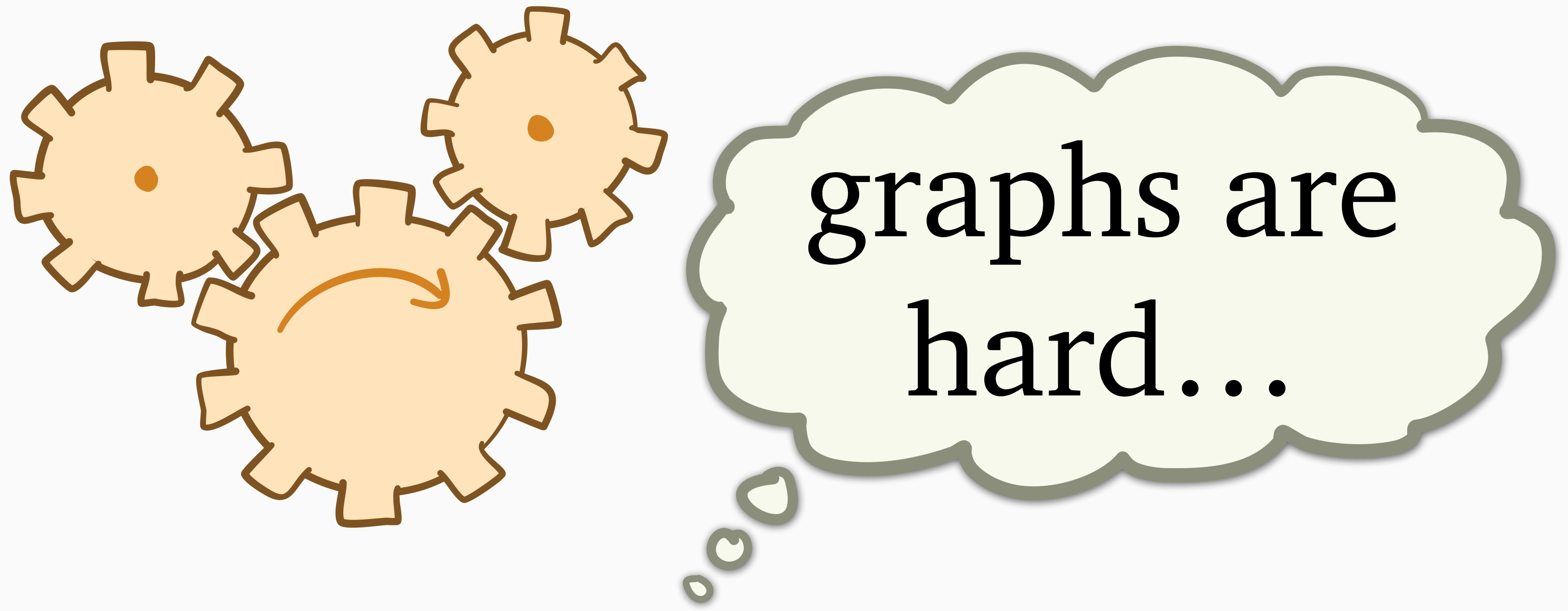
- Social network graph G
- Each person has *infection level* in $\{0, 1\}$
- *Healing:* at each node, $1 \to 0$ at rate μ
- *Infection:* at each node, $0 \to 1$ at rate $\lambda \cdot$(# of neighboring 1's)



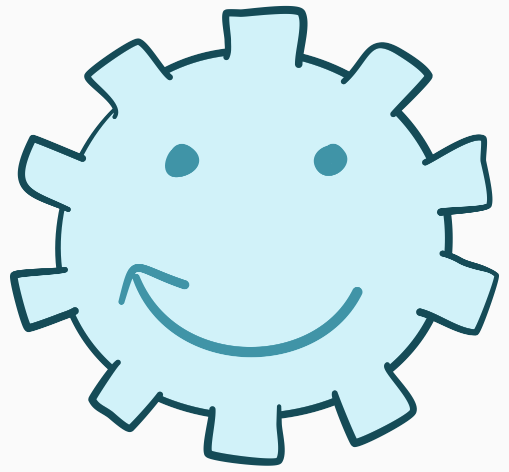
- Track *number*, not *set*, of infections
- k = number of people infected
- *Healing:* $k \to k - 1$ at rate μk
- *Infection:* $k \to k + 1$ at rate $\lambda \eta(G)$

Ganesh, Massoulié and Towsley: The effect of network topology on the spread of epidemics (2005)

SIS epidemic: lower bound



- Social network graph G
- Each person has *infection level* in $\{0, 1\}$
- *Healing:* at each node, $1 \to 0$ at rate μ
- *Infection:* at each node, $0 \to 1$ at rate $\lambda \cdot$(# of neighboring 1's)

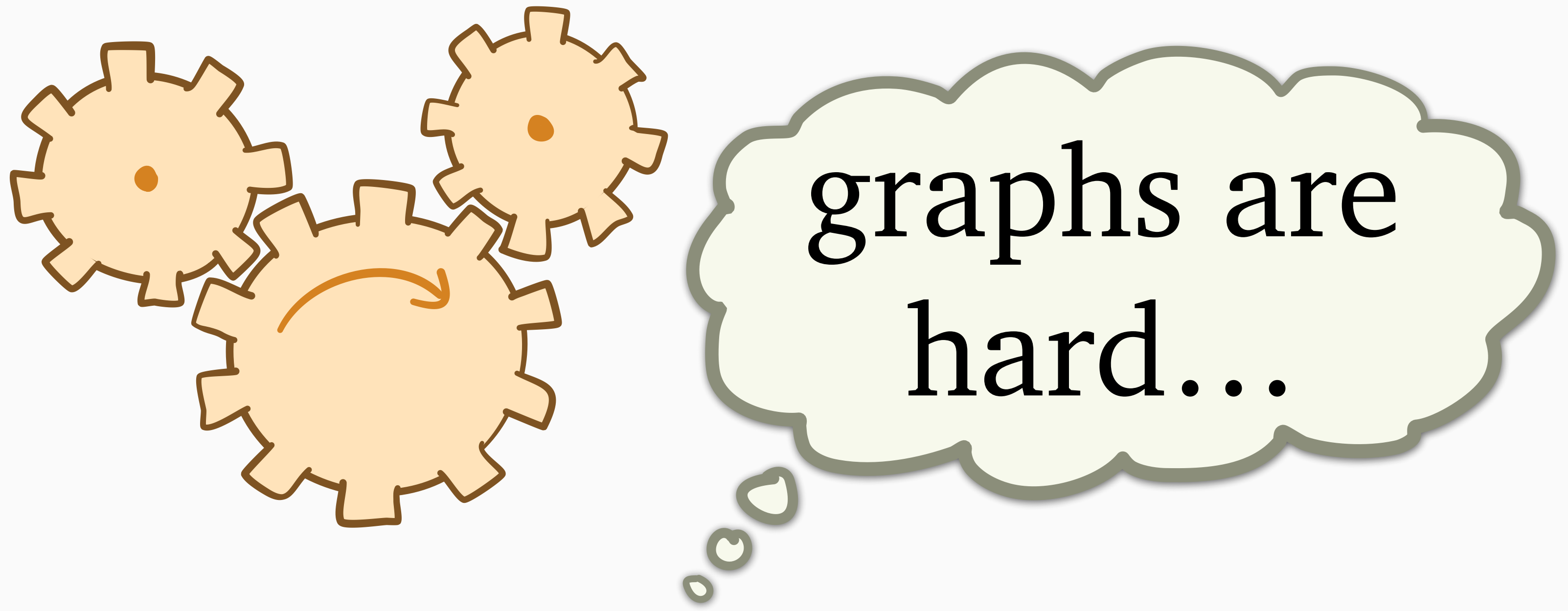


- Track *number*, not *set*, of infections
- k = number of people infected
- *Healing:* $k \to k - 1$ at rate μk
- *Infection:* $k \to k + 1$ at rate $\lambda \eta(G)$

$$\eta(G) = \min_{\text{vertex set } A} \frac{\#\text{ edges from } A \text{ to } \overline{A}}{|A|}$$

Ganesh, Massoulié and Towsley: The effect of network topology on the spread of epidemics (2005)

SIS epidemic: lower bound

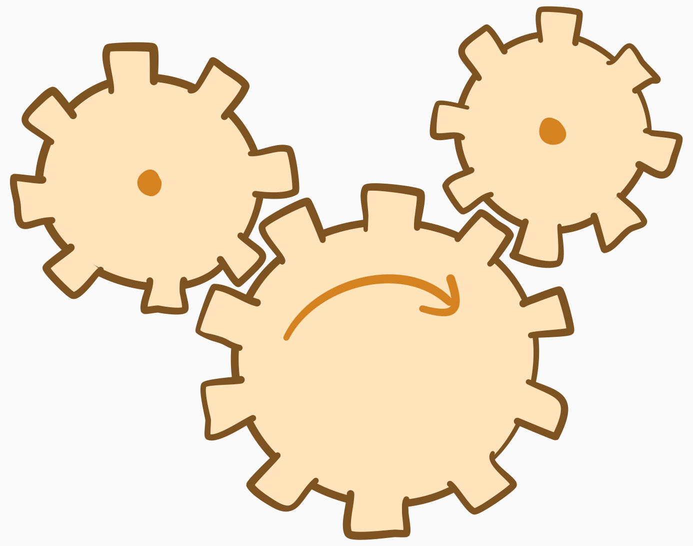


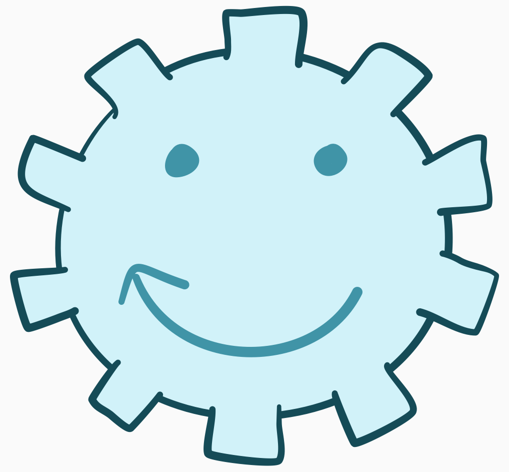
- Social network graph G
- Each person has *infection level* in $\{0, 1\}$
- *Healing:* at each node, $1 \to 0$ at rate μ
- *Infection:* at each node, $0 \to 1$ at rate $\lambda \cdot$(# of neighboring 1's)

- Track *number*, not *set*, of infections
- k = number of people infected
- *Healing:* $k \to k - 1$ at rate μk
- *Infection:* $k \to k + 1$ at rate $\lambda \eta(G)$

$$\eta(G) = \min_{\text{vertex set } A} \frac{\#\text{ edges from } A \text{ to } \overline{A}}{|A|}$$

SIS epidemic: upper bound

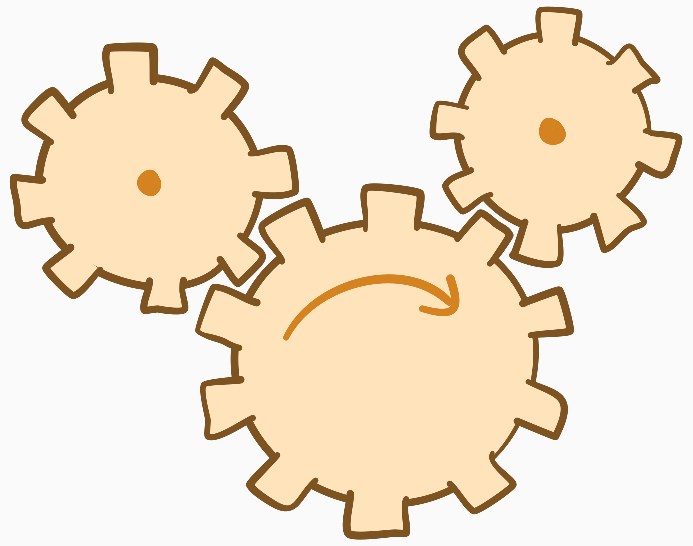




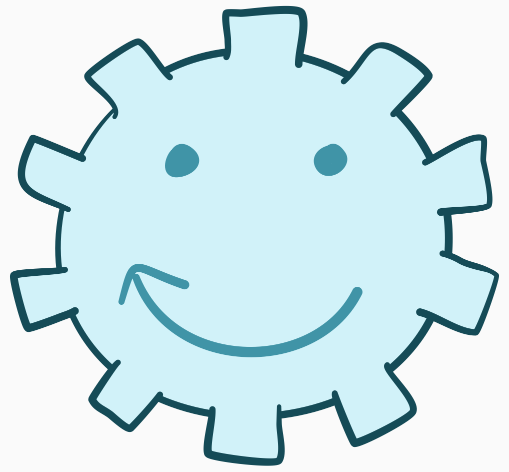
- Social network graph G
- Each person has *infection level* in $\{0, 1\}$
- *Healing:* at each node, $1 \to 0$ at rate μ
- *Infection:* at each node, $0 \to 1$ at rate $\lambda \cdot$(# of neighboring 1's)

Ganesh, Massoulié and Towsley: The effect of network topology on the spread of epidemics (2005)

SIS epidemic: upper bound



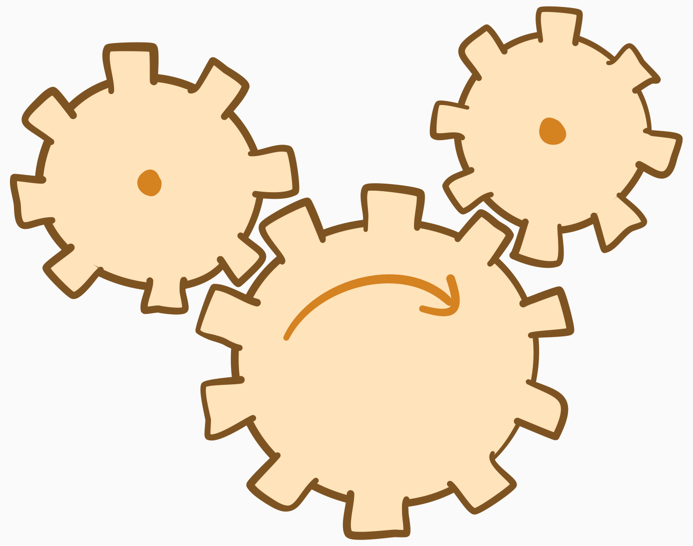
- Social network graph G
- Each person has *infection level* in $\{0, 1\}$
- *Healing:* at each node, $1 \to 0$ at rate μ
- *Infection:* at each node, $0 \to 1$ at rate $\lambda \cdot$(# of neighboring 1's)



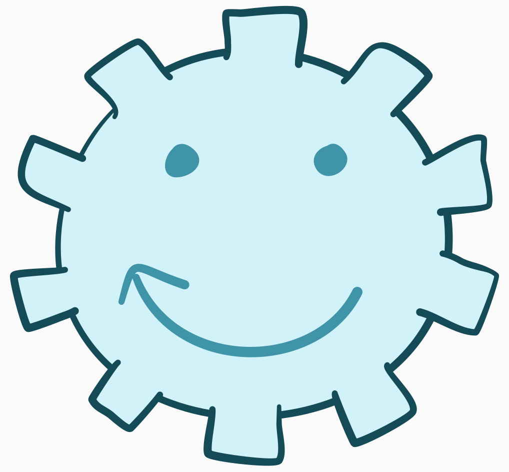
- Social network graph G

Ganesh, Massoulié and Towsley: The effect of network topology on the spread of epidemics (2005)

SIS epidemic: upper bound



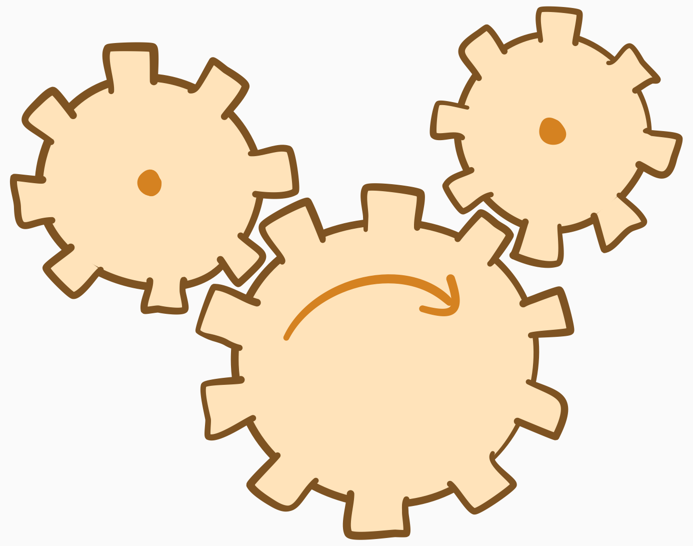
- Social network graph G
- Each person has *infection level* in $\{0, 1\}$
- *Healing:* at each node, $1 \to 0$ at rate μ
- *Infection:* at each node, $0 \to 1$ at rate $\lambda \cdot (\#$ of neighboring 1's)



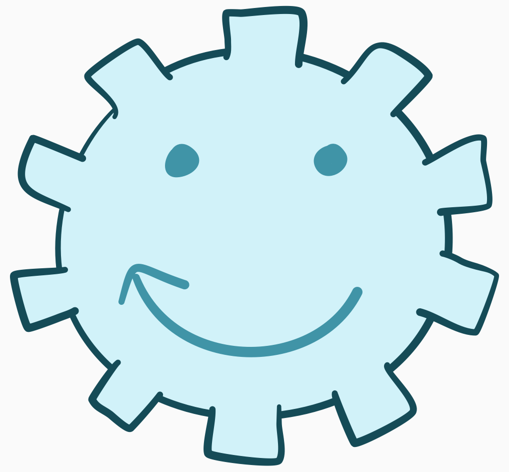
- Social network graph G
- Each person has *infection level* in $\{0, 1, 2, \ldots\}$

Ganesh, Massoulié and Towsley: The effect of network topology on the spread of epidemics (2005)

SIS epidemic: upper bound



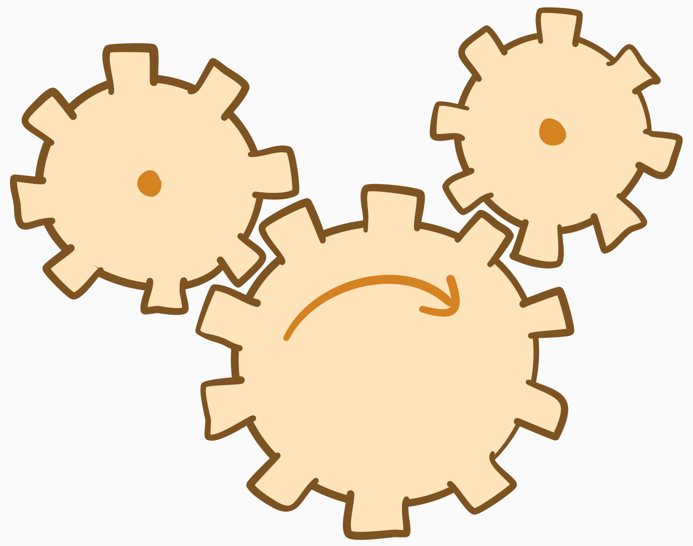
- Social network graph G
- Each person has *infection level* in $\{0, 1\}$
- *Healing:* at each node, $1 \to 0$ at rate μ
- *Infection:* at each node, $0 \to 1$ at rate $\lambda \cdot$(# of neighboring 1's)



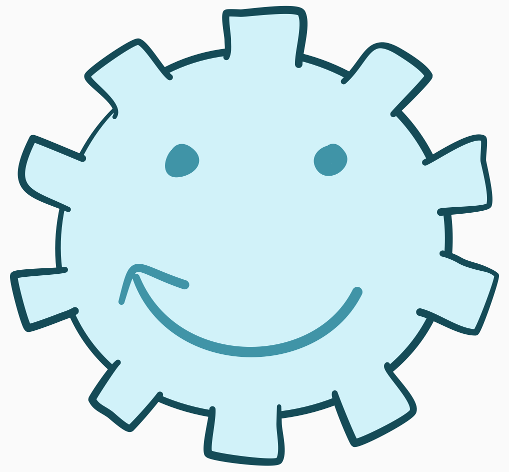
- Social network graph G
- Each person has *infection level* in $\{0, 1, 2, \ldots\}$
- *Healing:* i's become $(i - 1)$'s at rate μi

Ganesh, Massoulié and Towsley: The effect of network topology on the spread of epidemics (2005)

SIS epidemic: upper bound



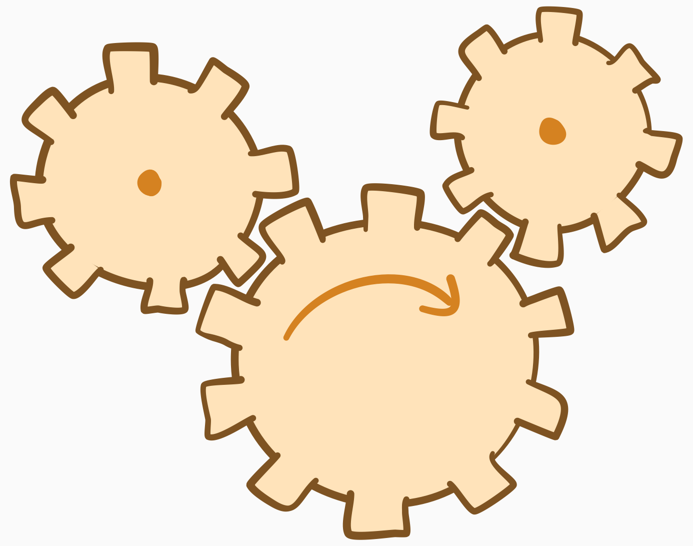
- Social network graph G
- Each person has *infection level* in $\{0, 1\}$
- *Healing:* at each node, $1 \to 0$ at rate μ
- *Infection:* at each node, $0 \to 1$ at rate $\lambda \cdot$(# of neighboring 1's)



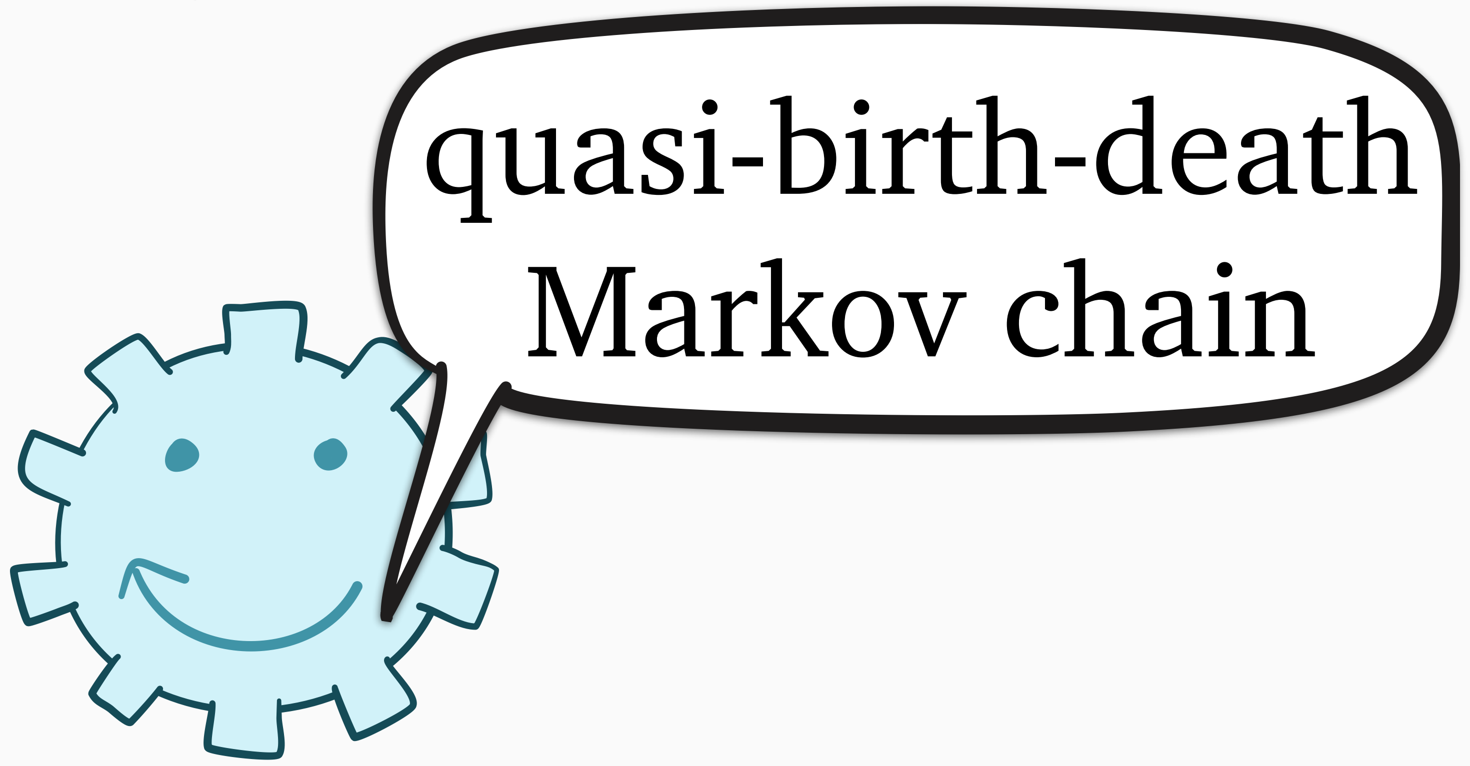
- Social network graph G
- Each person has *infection level* in $\{0, 1, 2, \ldots\}$
- *Healing:* i's become $(i - 1)$'s at rate μi
- *Infection:* at each node, $i \to i + 1$ at rate $\lambda \cdot$(total infection level of neighbors)

Ganesh, Massoulié and Towsley: The effect of network topology on the spread of epidemics (2005)

SIS epidemic: upper bound



- Social network graph G
- Each person has *infection level* in {0, 1}
- *Healing:* at each node, 1 → 0 at rate μ
- *Infection:* at each node, 0 → 1 at rate $\lambda \cdot$(# of neighboring 1's)

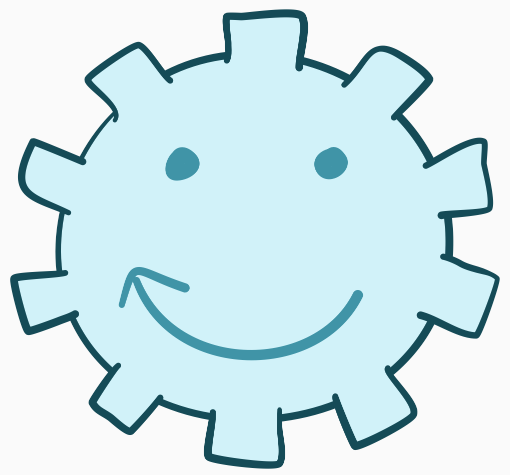


- Social network graph G
- Each person has *infection level* in {0, 1, 2, …}
- *Healing:* i's become $(i - 1)$'s at rate μi
- *Infection:* at each node, $i \to i + 1$ at rate $\lambda \cdot$(total infection level of neighbors)

Ganesh, Massoulié and Towsley: The effect of network topology on the spread of epidemics (2005)

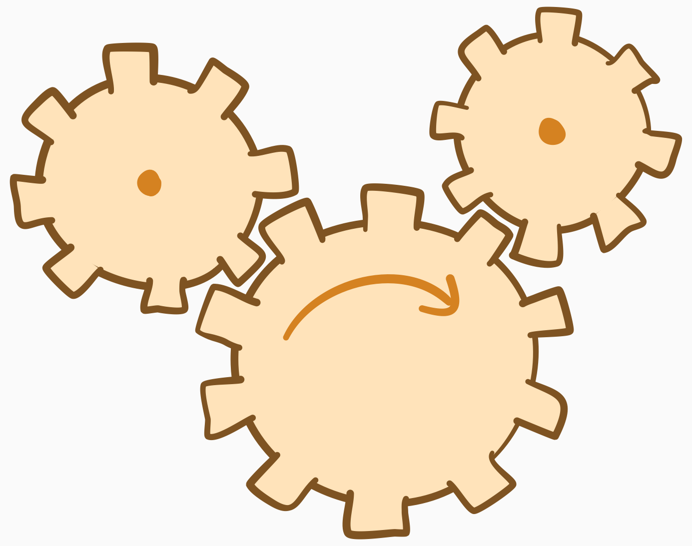
SIS epidemic: results

lower bound



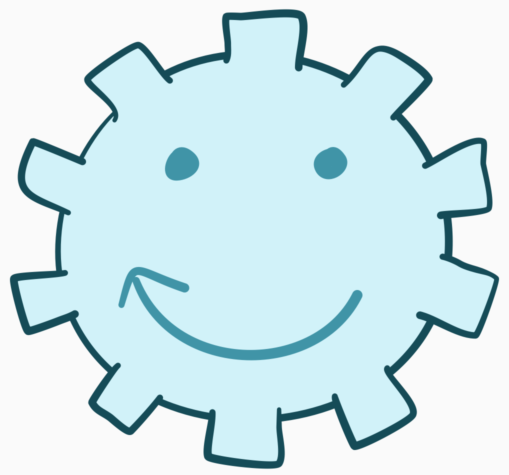
$\leq$

SIS model



$\leq$

upper bound

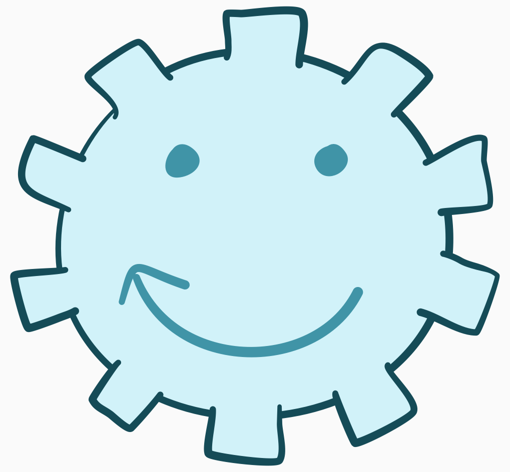
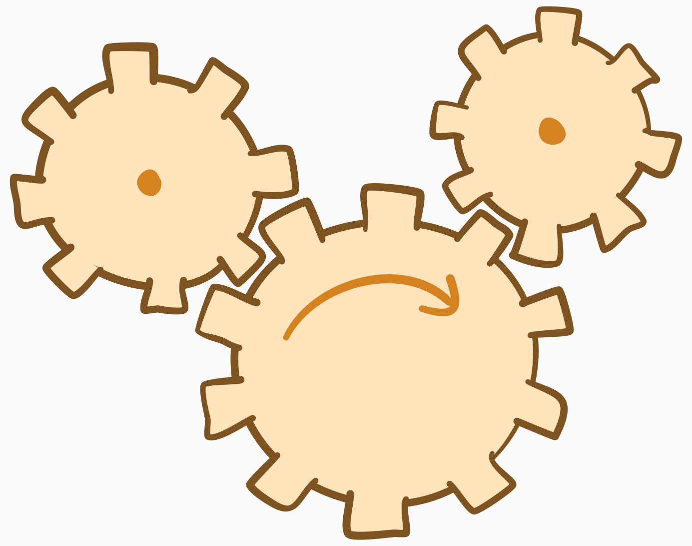
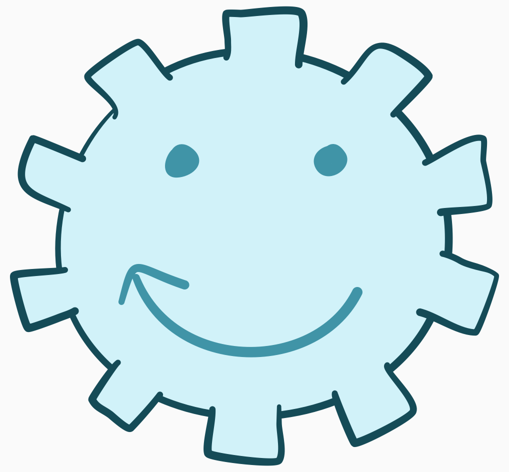


- Birth-death Markov chain
- Quasi-birth-death Markov chain

Ganesh, Massoulié and Towsley: The effect of network topology on the spread of epidemics (2005)

SIS epidemic: results

lower bound | **SIS model** | **upper bound**

 $\leq$  $\leq$ 

- Birth-death Markov chain
- Infection *decays slowly* if $\eta(G) > \mu/\lambda$

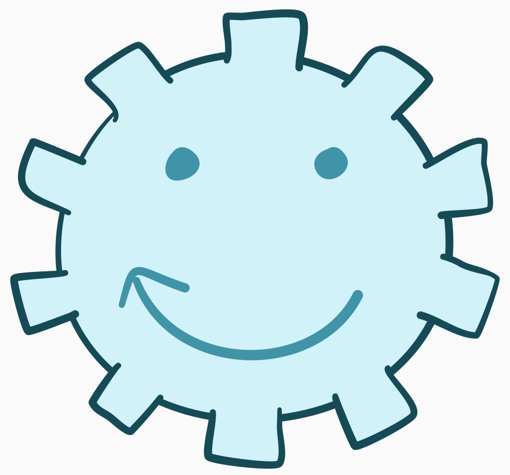
"bottleneck" ratio of G

- Quasi-birth-death Markov chain

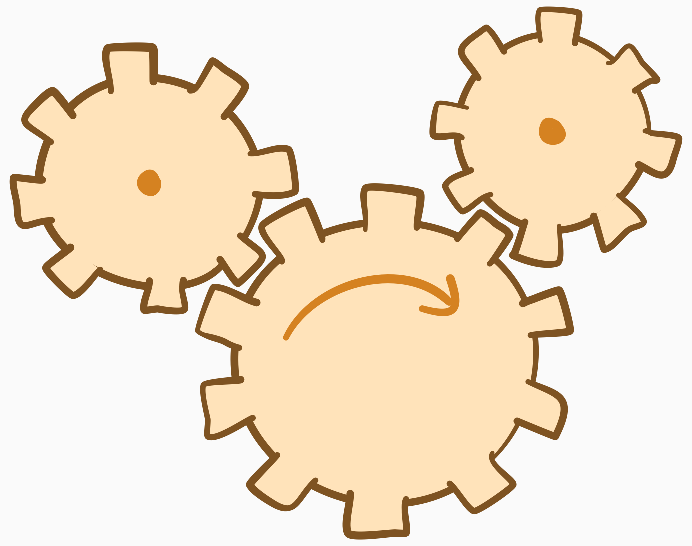
Ganesh, Massoulié and Towsley: The effect of network topology on the spread of epidemics (2005)

SIS epidemic: results

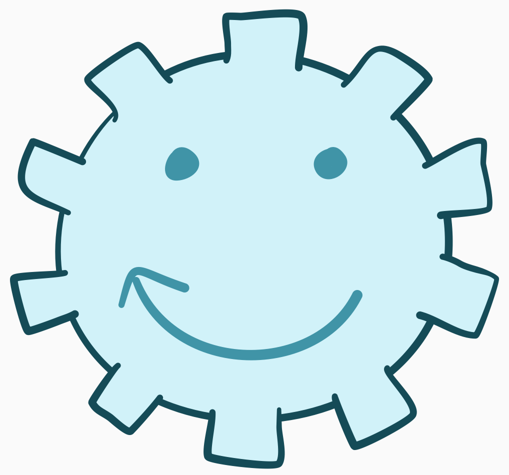
lower bound | **SIS model** | **upper bound**



$\leq$



$\leq$



- Birth-death Markov chain
- Infection *decays slowly* if $\eta(G) > \mu/\lambda$

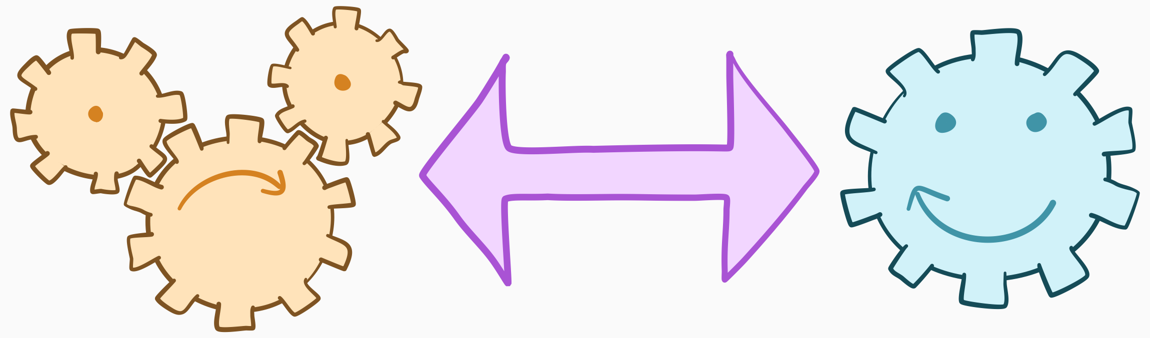
"bottleneck" ratio of G

- Quasi-birth-death Markov chain
- Infection *decays quickly* if $\rho(G) < \mu/\lambda$

largest eigenvalue of G's adjacency matrix

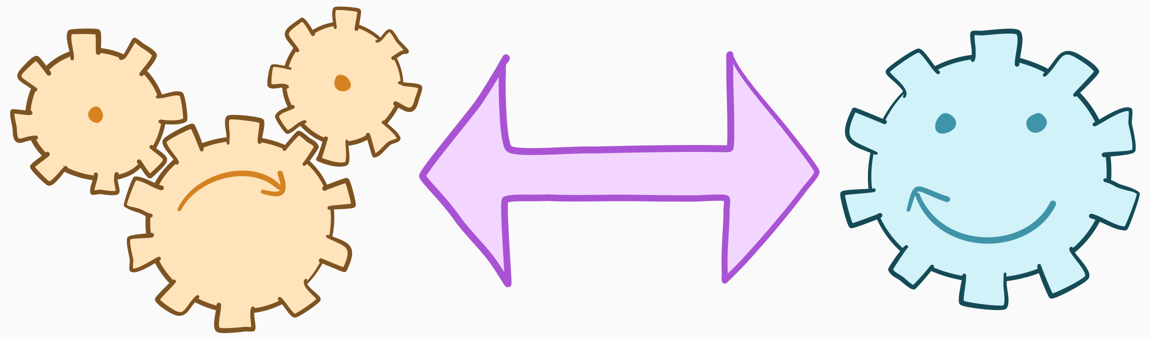
Ganesh, Massoulié and Towsley: The effect of network topology on the spread of epidemics (2005)

Classifying coupling techniques



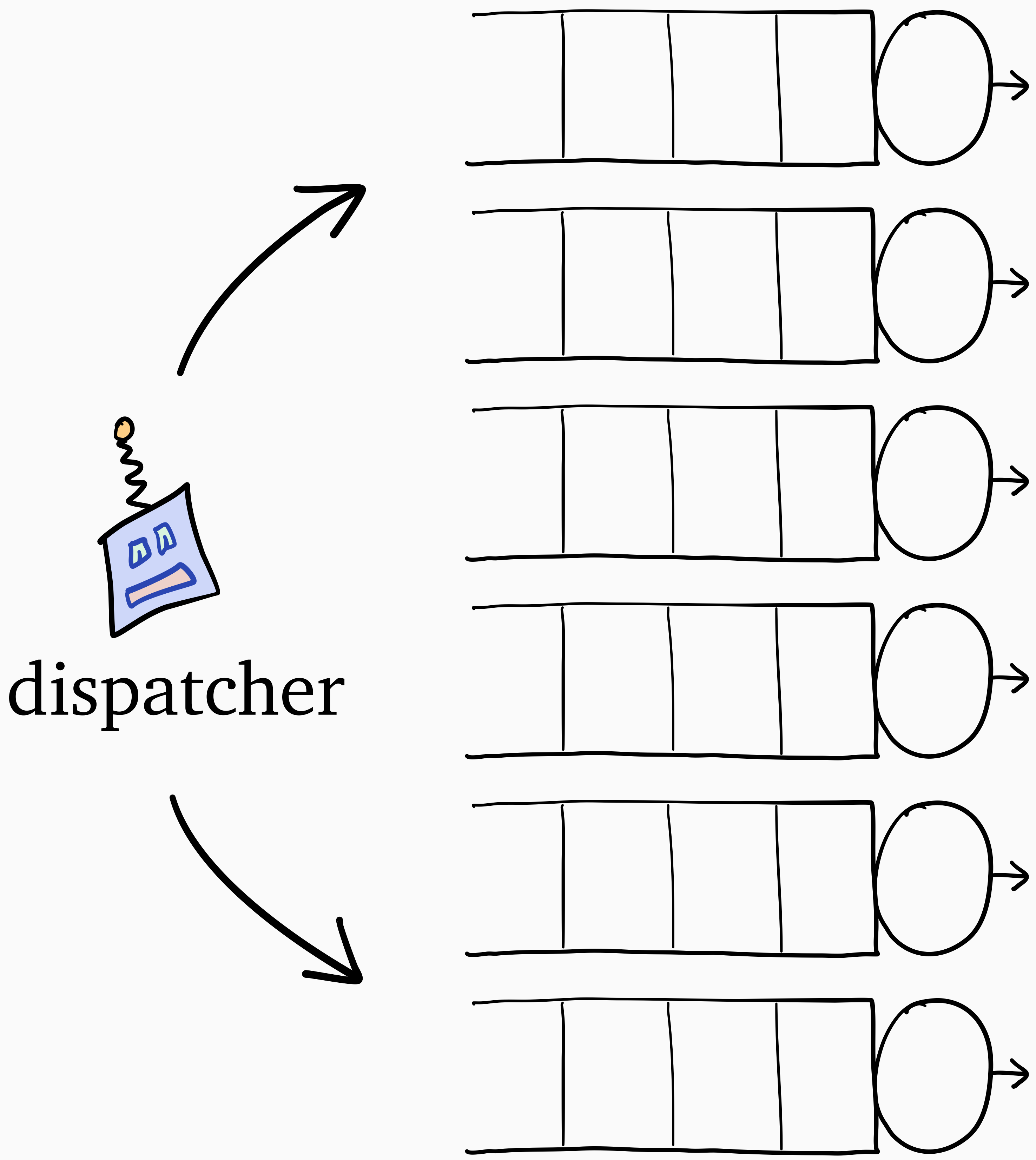
	A. Every sample path	B. Steady-state distribution
1. More information	A1	B1
2. Fewer constraints	A2 M/M/k vs. M/M/1	B2
3. Simpler dynamics	A3	B3

Classifying coupling techniques



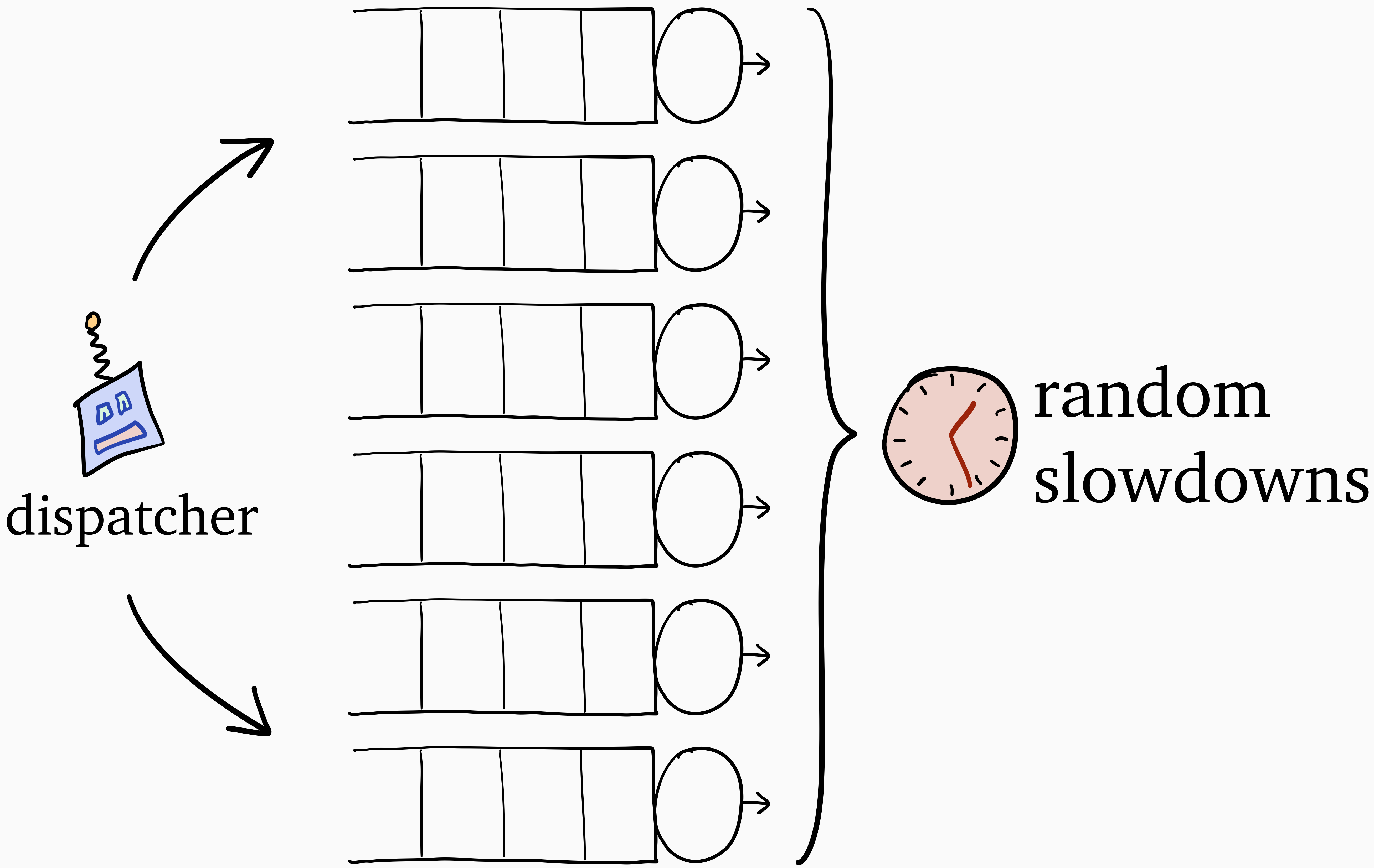
	A. Every sample path	B. Steady-state distribution
1. More information	A1	B1
2. Fewer constraints	A2 M/M/k vs. M/M/1	B2
3. Simpler dynamics	A3 SIS epidemics	B3

Queues with redundancy



Gardner, Harchol-Balter, and Scheller-Wolf: A Better Model for Job Redundancy: Decoupling Server Slowdown and Job Size (2016)

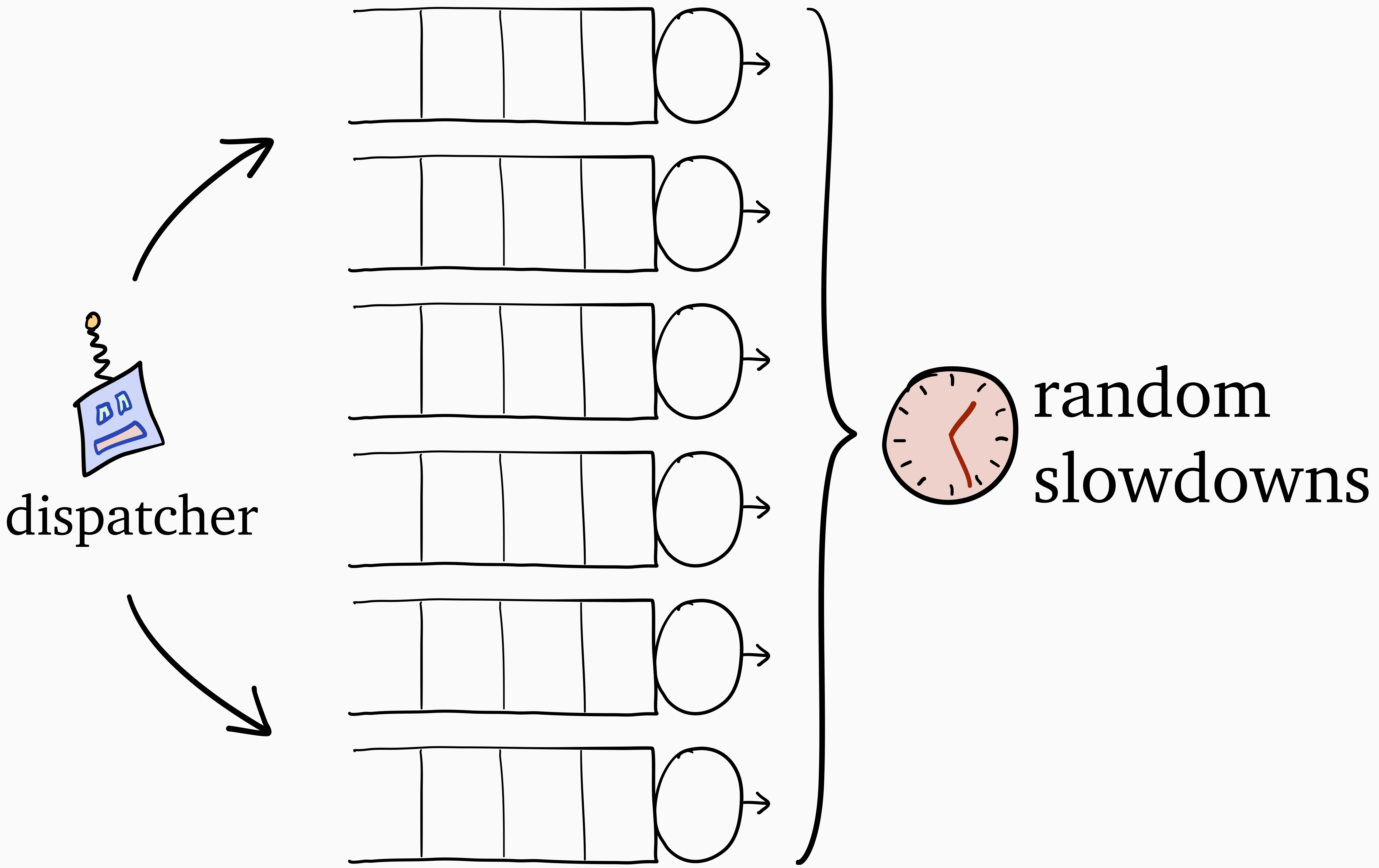
Queues with redundancy



Gardner, Harchol-Balter, and Scheller-Wolf: A Better Model for Job Redundancy: Decoupling Server Slowdown and Job Size (2016)

Queues with redundancy

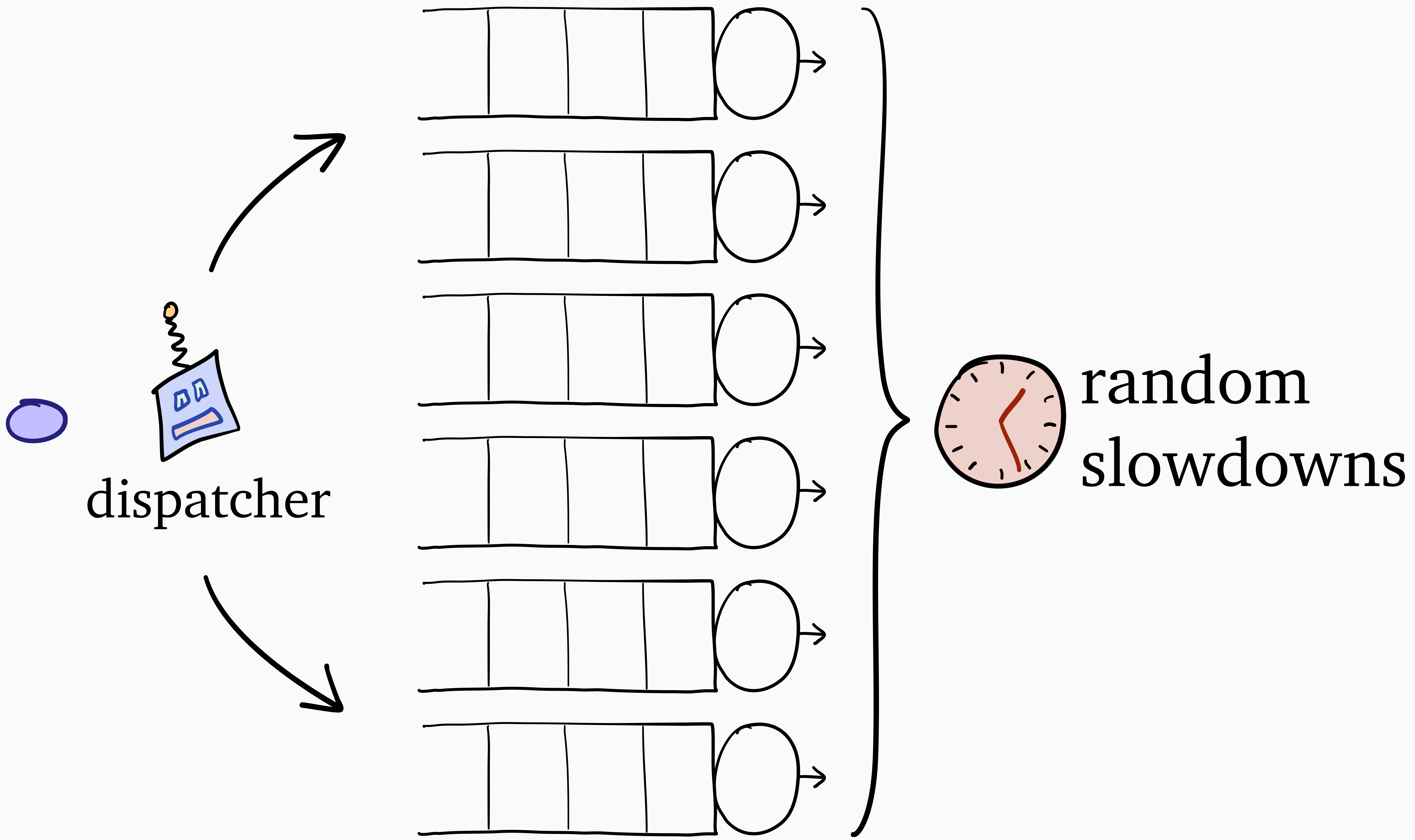
Redundancy-*d*:
make *d* copies of each job,
dispatch them randomly



Gardner, Harchol-Balter, and Scheller-Wolf: A Better Model for Job Redundancy: Decoupling Server Slowdown and Job Size (2016)

Queues with redundancy

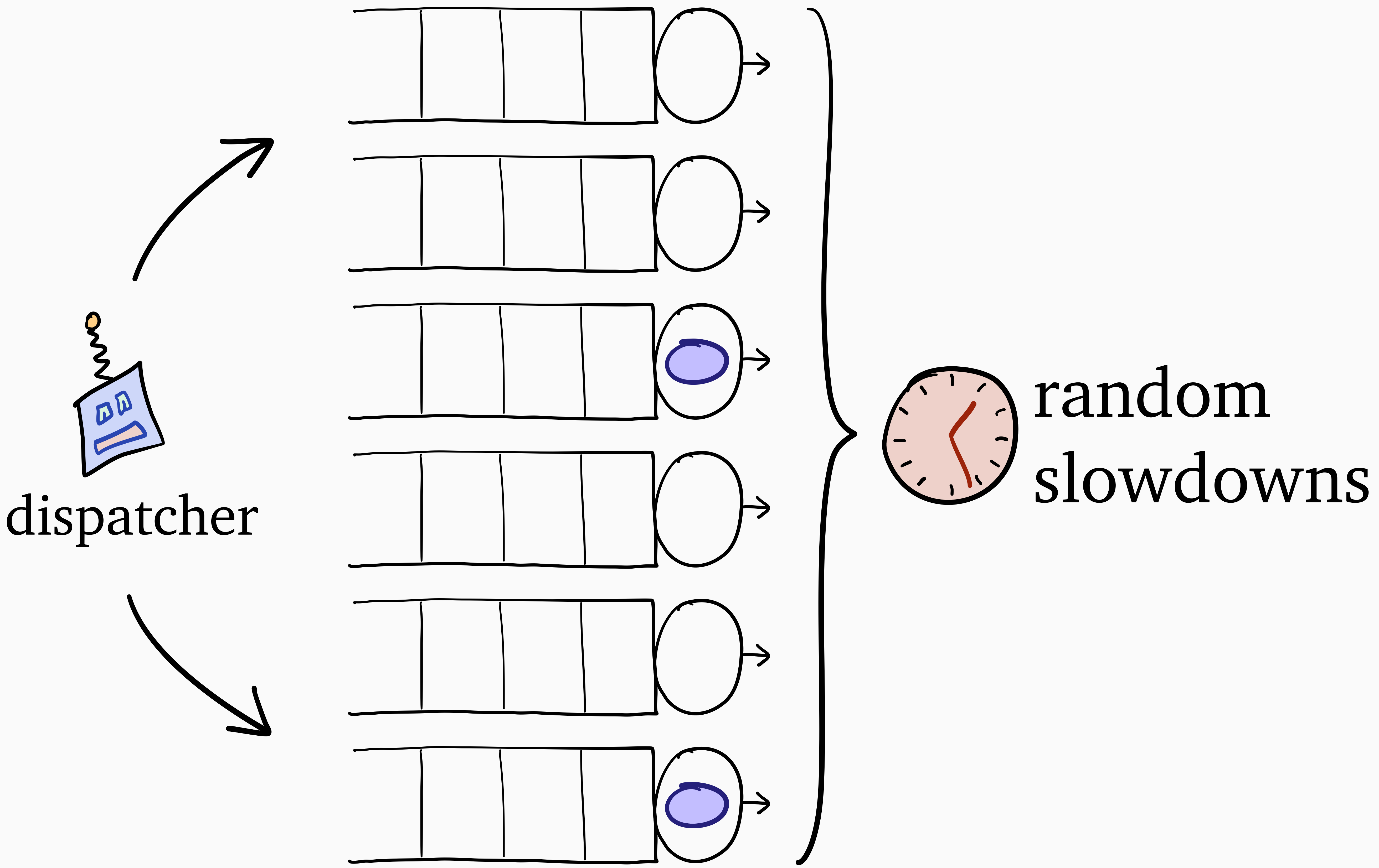
Redundancy-*d*:
make *d* copies of each job,
dispatch them randomly



Gardner, Harchol-Balter, and Scheller-Wolf: A Better Model for Job Redundancy: Decoupling Server Slowdown and Job Size (2016)

Queues with redundancy

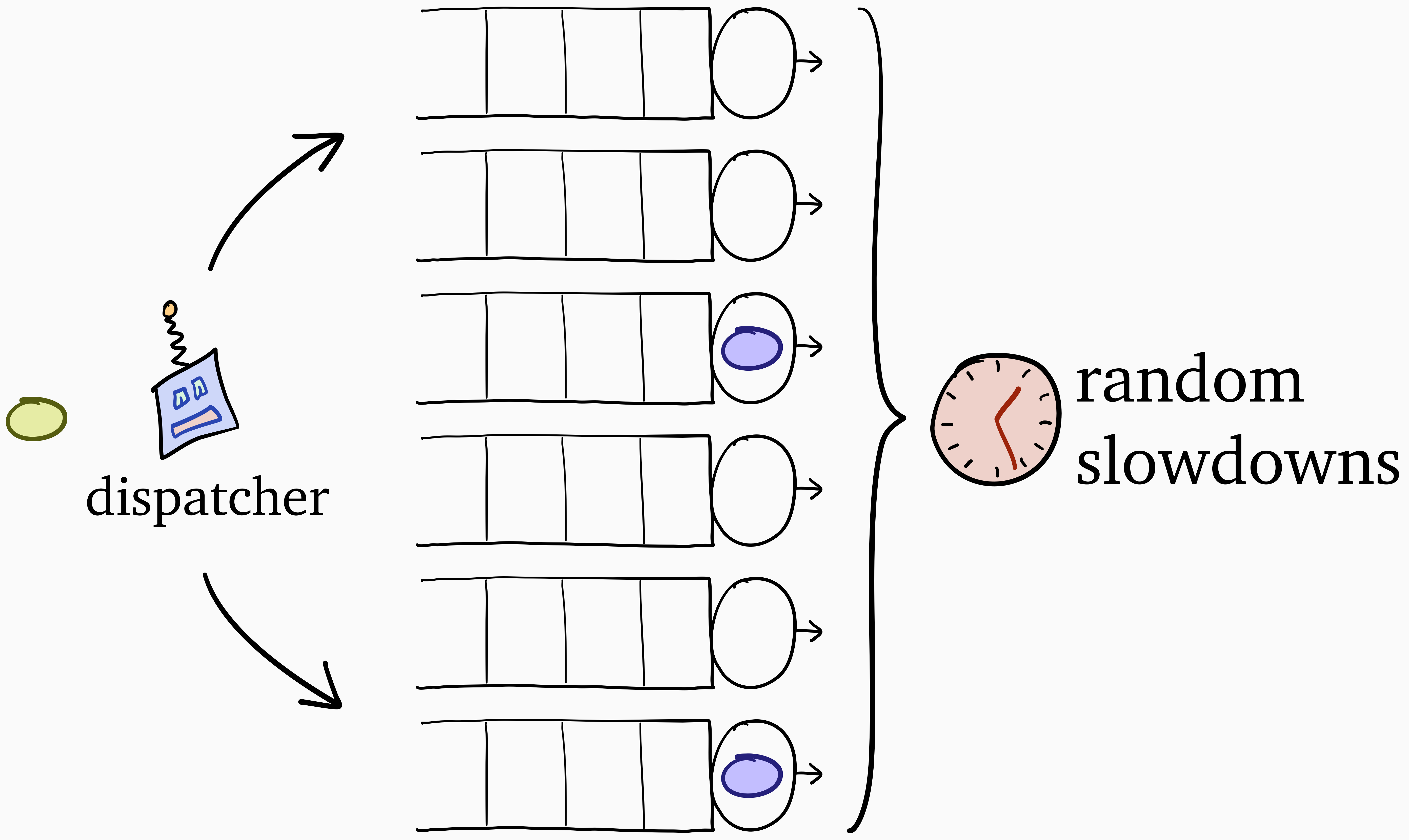
Redundancy-*d*:
make d copies of each job,
dispatch them randomly



Gardner, Harchol-Balter, and Scheller-Wolf: A Better Model for Job Redundancy: Decoupling Server Slowdown and Job Size (2016)

Queues with redundancy

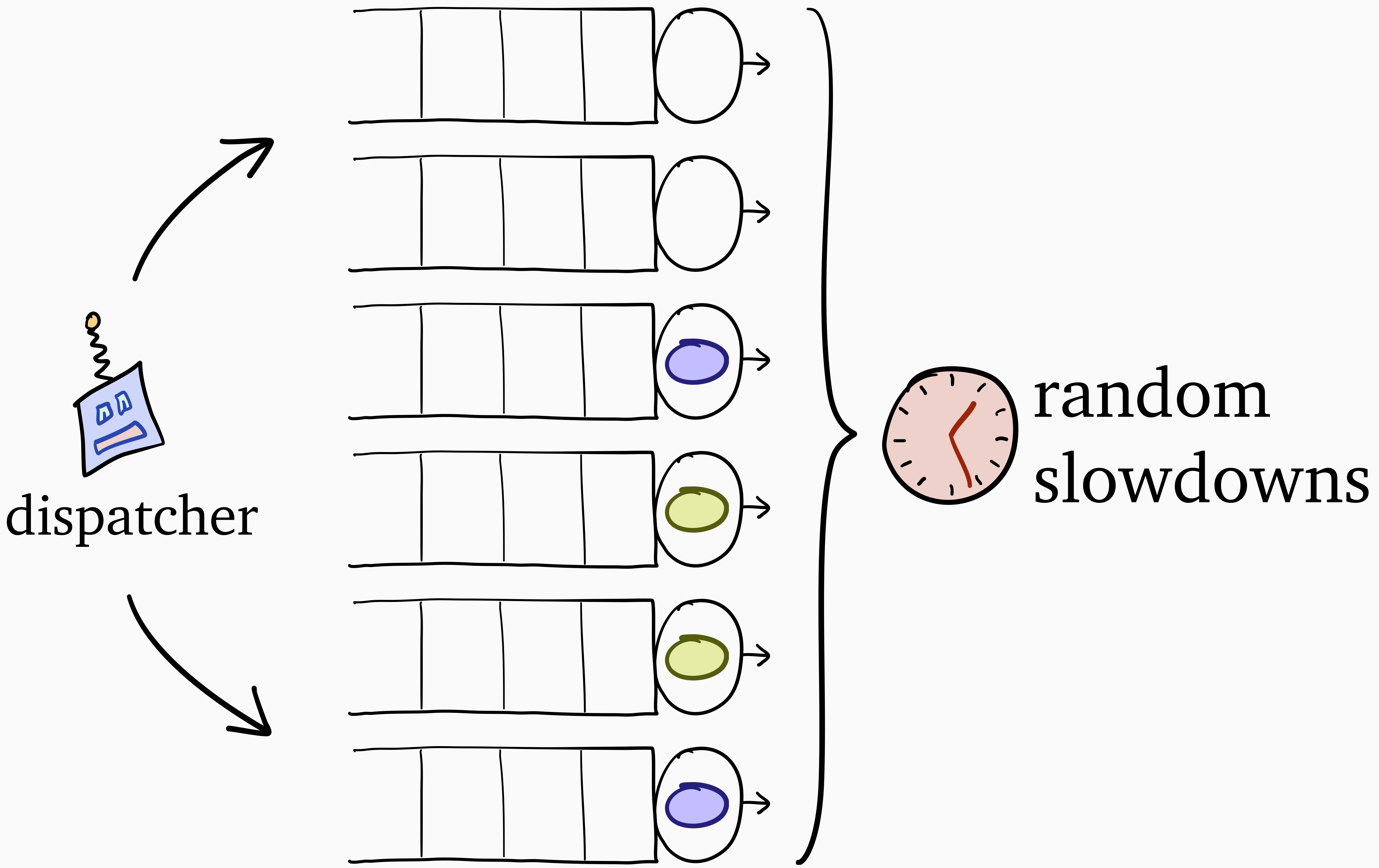
Redundancy-*d*:
make d copies of each job,
dispatch them randomly



Gardner, Harchol-Balter, and Scheller-Wolf: A Better Model for Job Redundancy: Decoupling Server Slowdown and Job Size (2016)

Queues with redundancy

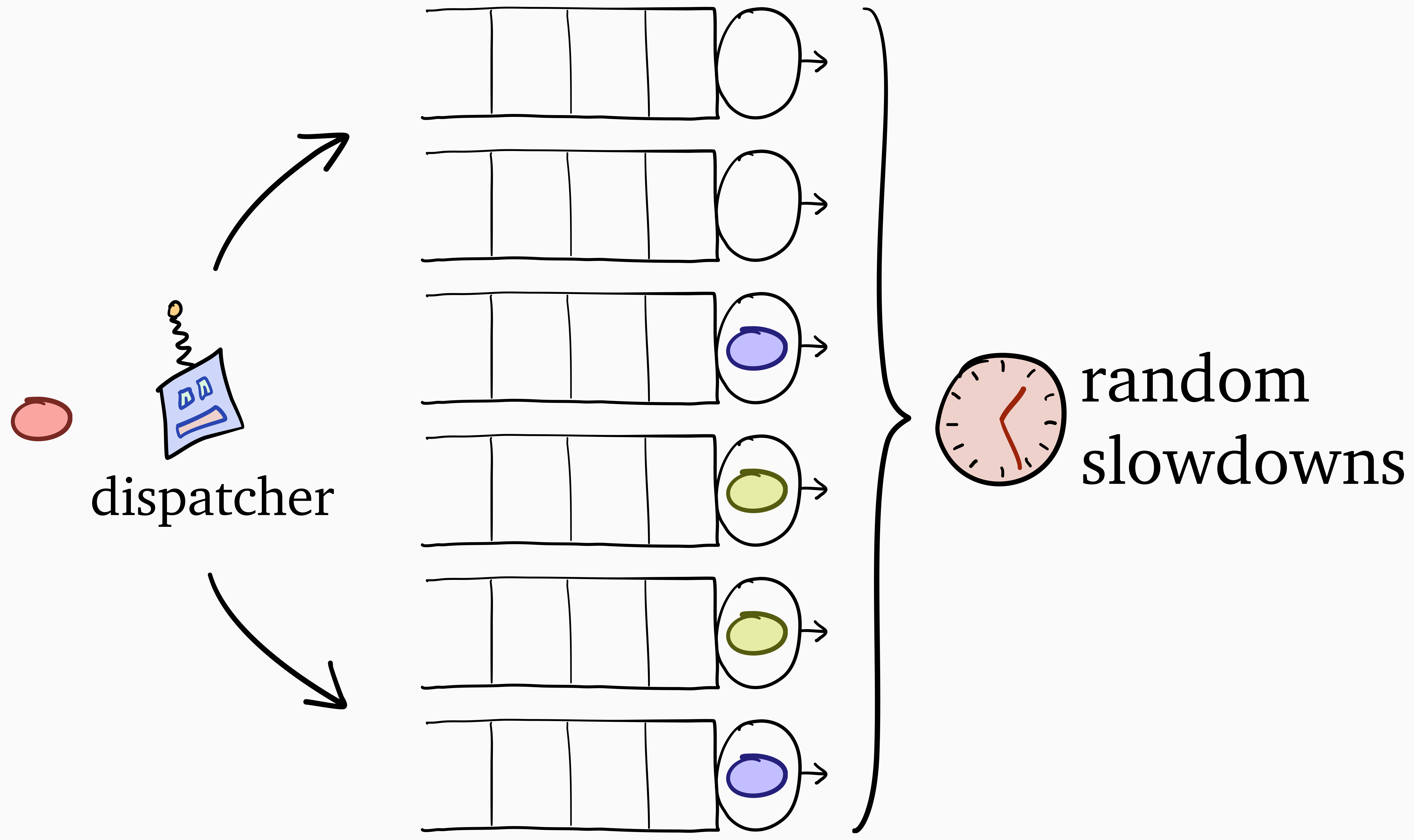
Redundancy-*d*:
make *d* copies of each job,
dispatch them randomly



Gardner, Harchol-Balter, and Scheller-Wolf: A Better Model for Job Redundancy: Decoupling Server Slowdown and Job Size (2016)

Queues with redundancy

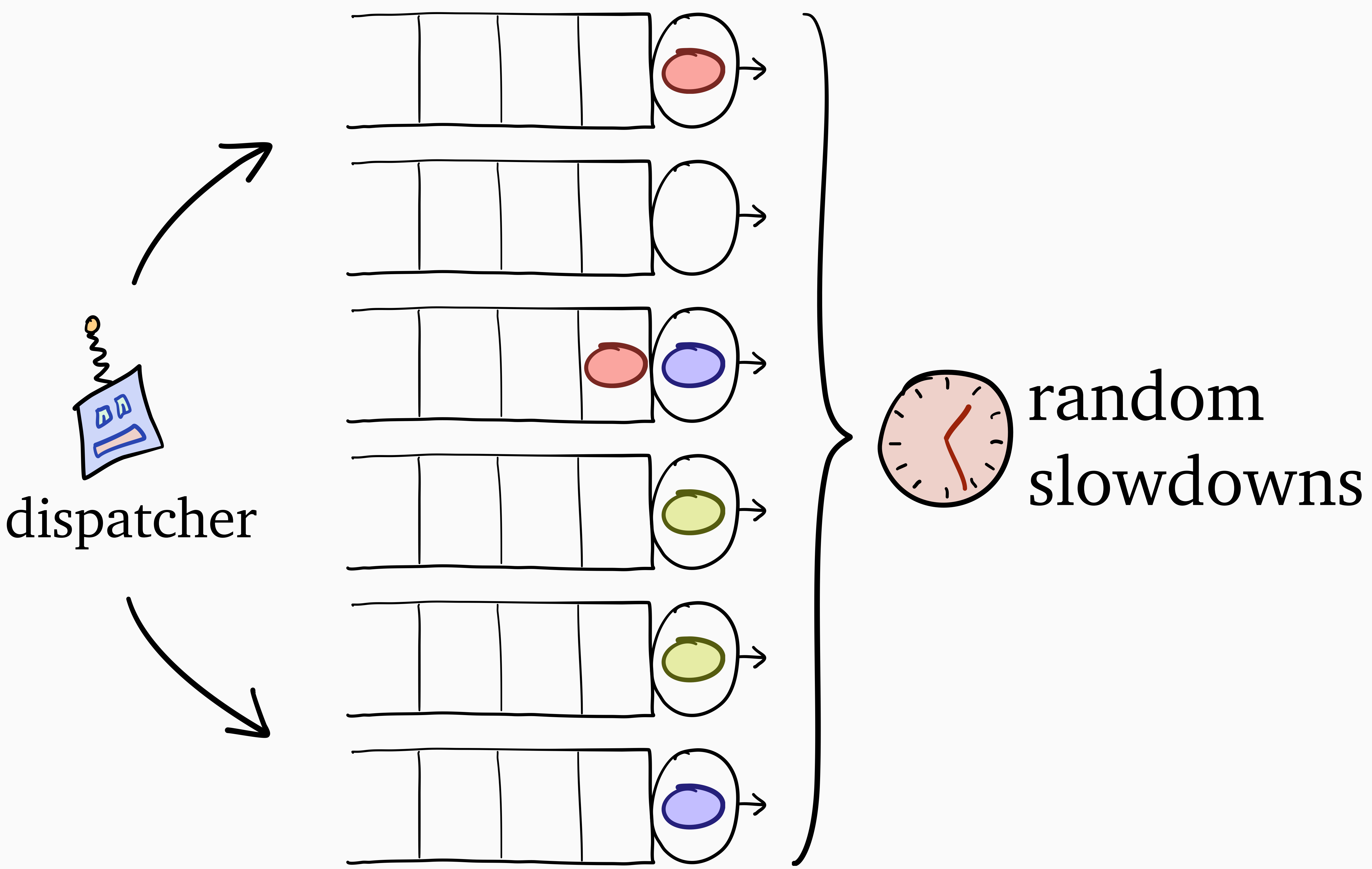
Redundancy-*d*:
make *d* copies of each job,
dispatch them randomly



Gardner, Harchol-Balter, and Scheller-Wolf: A Better Model for Job Redundancy: Decoupling Server Slowdown and Job Size (2016)

Queues with redundancy

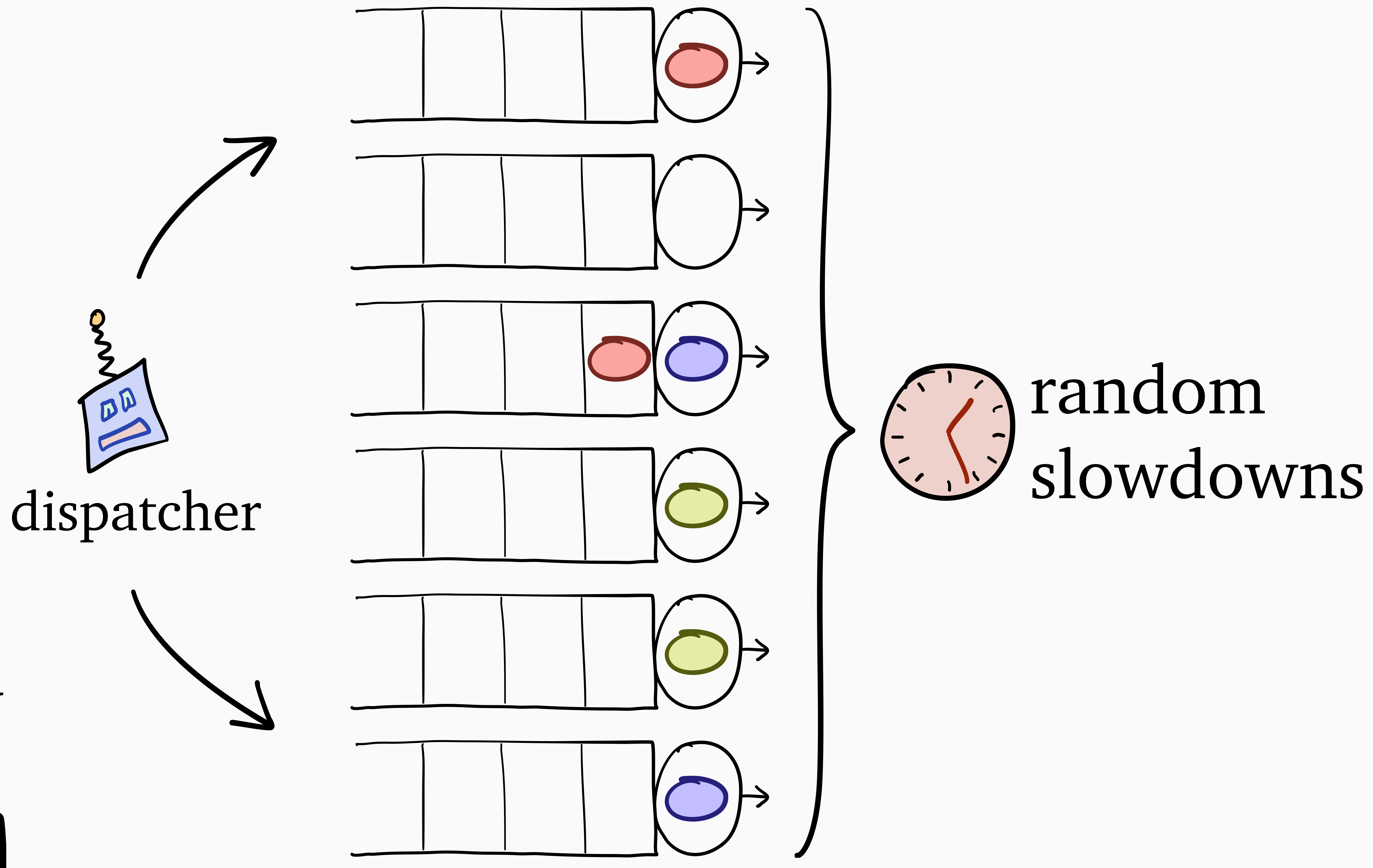
Redundancy-*d*:
make *d* copies of each job,
dispatch them randomly



Gardner, Harchol-Balter, and Scheller-Wolf: A Better Model for Job Redundancy: Decoupling Server Slowdown and Job Size (2016)

Queues with redundancy

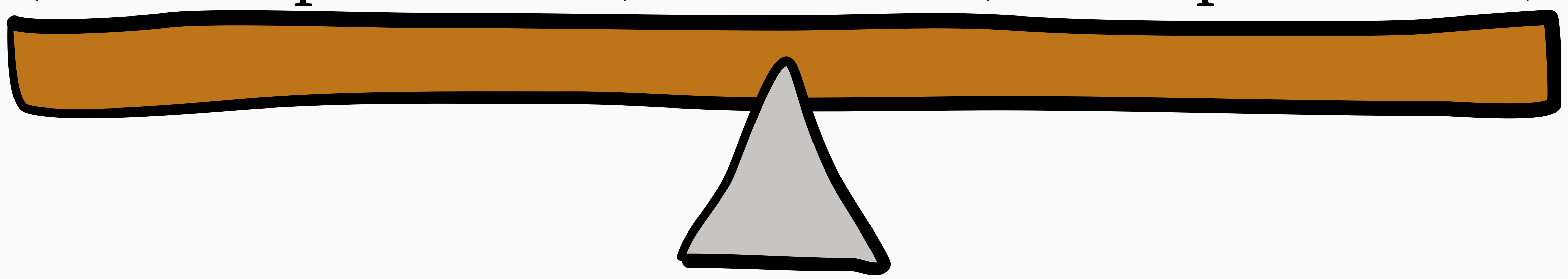
Redundancy-*d*:
make *d* copies of each job,
dispatch them randomly



Tradeoff:

mitigate slowdown
(more copies better)

ensure stability
(few copies better)



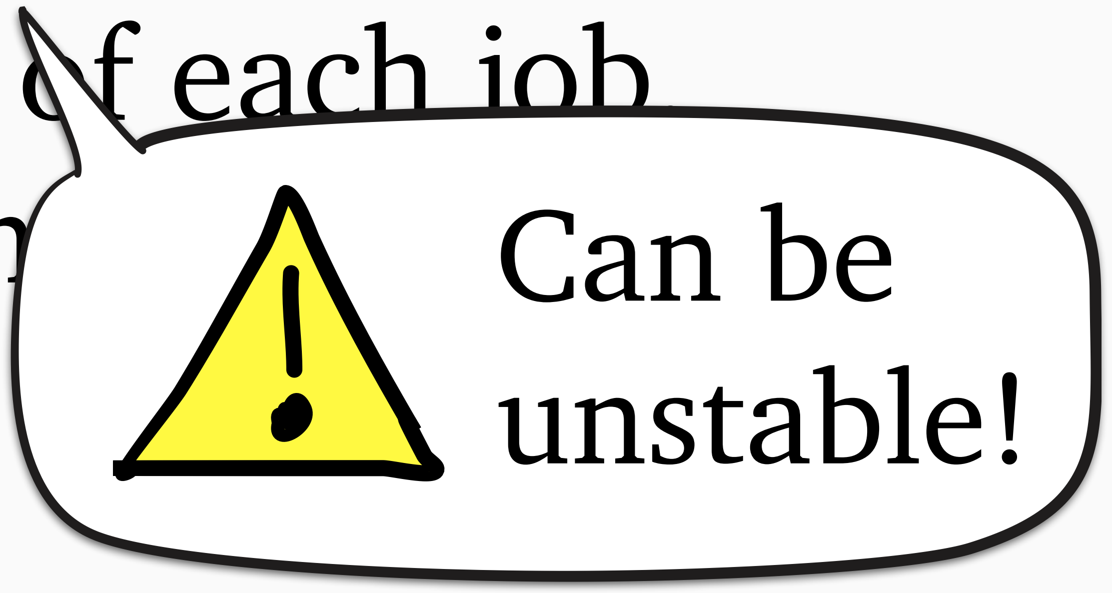
Gardner, Harchol-Balter, and Scheller-Wolf: A Better Model for Job Redundancy: Decoupling Server Slowdown and Job Size (2016)

Queues with redundancy

Redundancy-d:

make d copies of each job

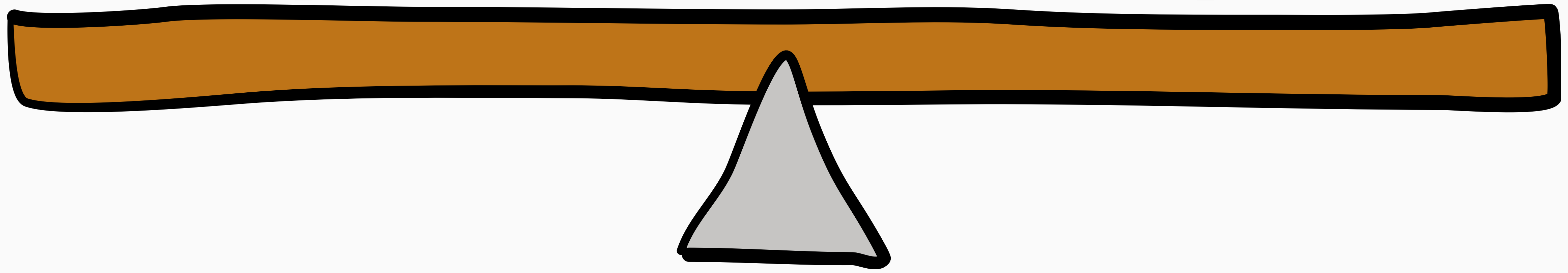
dispatch them

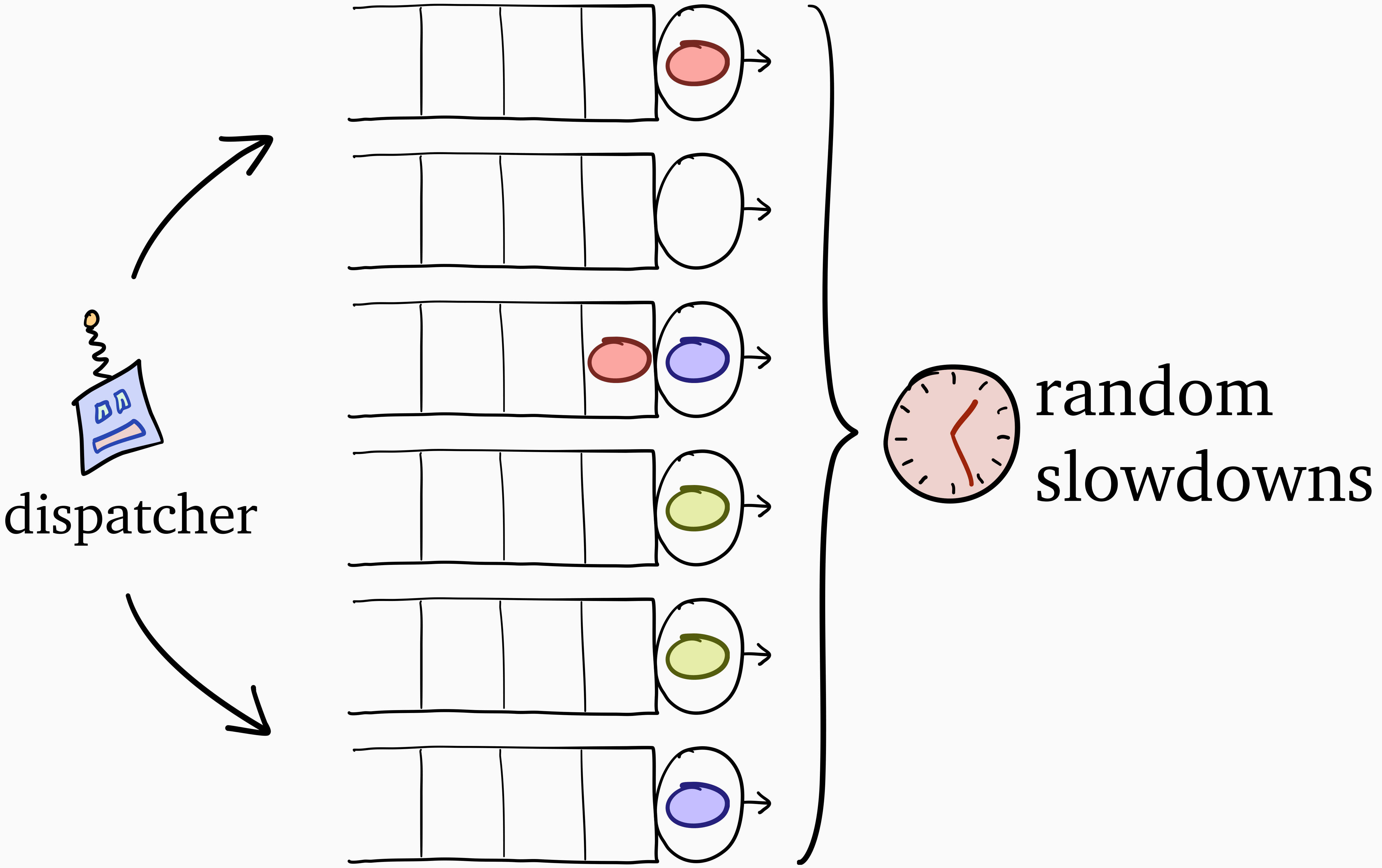


Tradeoff:

mitigate slowdown
(more copies better)

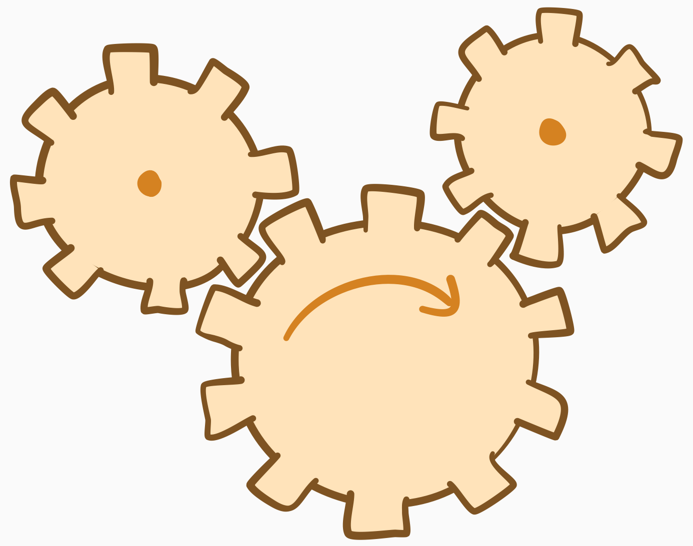
ensure stability
(few copies better)

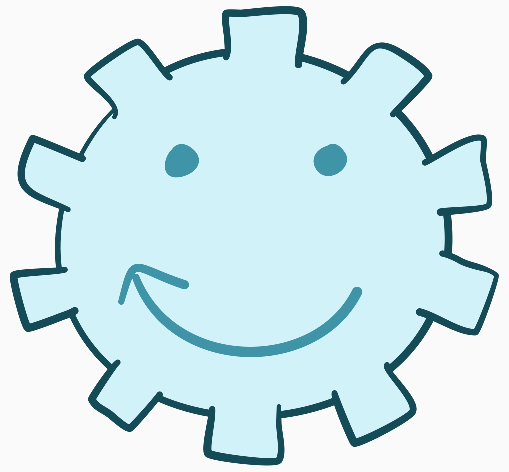




Gardner, Harchol-Balter, and Scheller-Wolf: A Better Model for Job Redundancy: Decoupling Server Slowdown and Job Size (2016)

Ensuring stability with redundancy

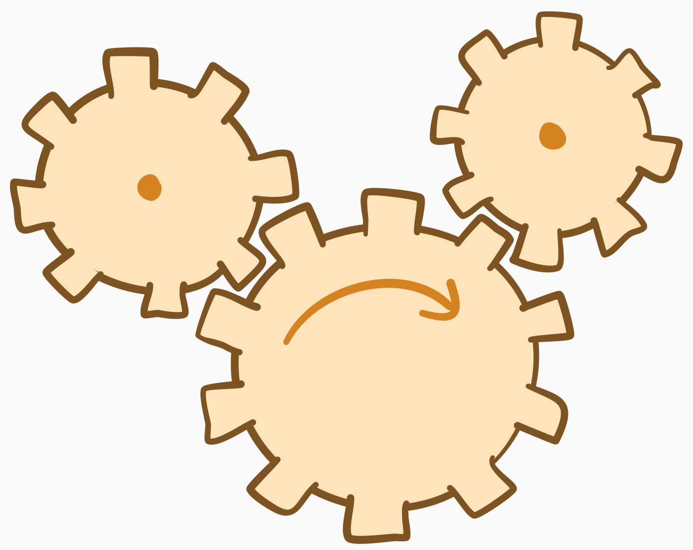


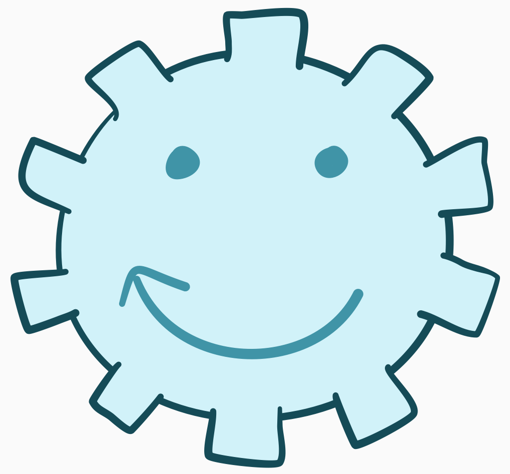


- Load-balancing queueing system with *redundancy*
- Each job has *size*, but servers have random *slowdowns*

Gardner, Harchol-Balter, and Scheller-Wolf: A Better Model for Job Redundancy: Decoupling Server Slowdown and Job Size (2016)

Ensuring stability with redundancy

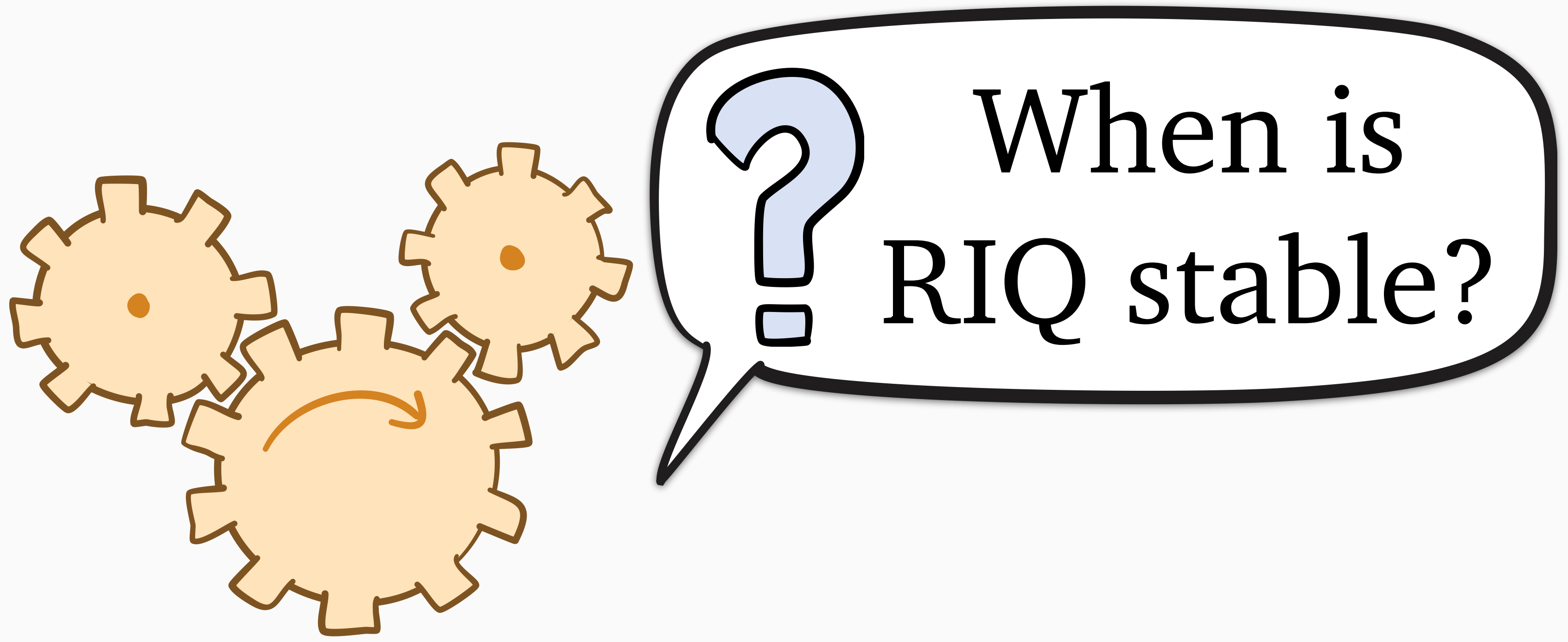


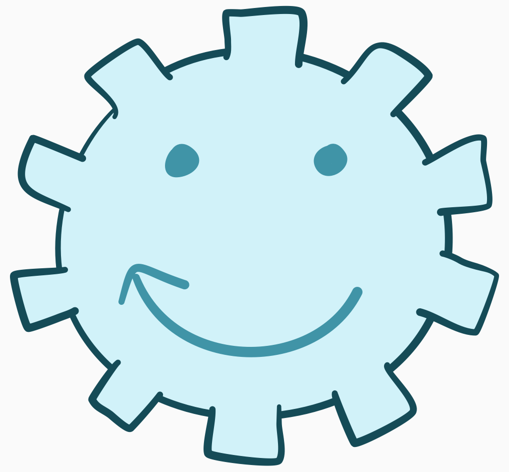


- Load-balancing queueing system with *redundancy*
- Each job has *size*, but servers have random *slowdowns*
- **Redundant-to-Idle-Queue (RIQ):** make copies of arriving jobs at *all* idle servers (and *only* idle servers)

Gardner, Harchol-Balter, and Scheller-Wolf: A Better Model for Job Redundancy: Decoupling Server Slowdown and Job Size (2016)

Ensuring stability with redundancy

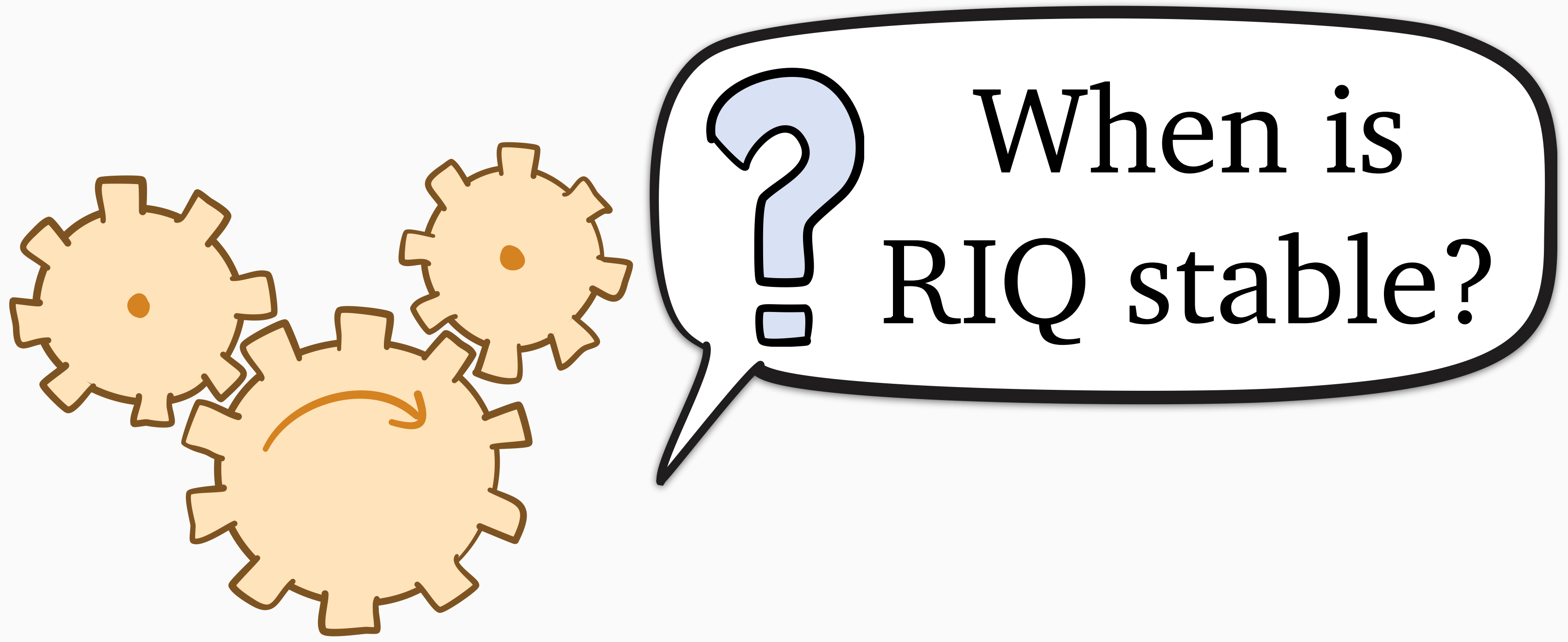


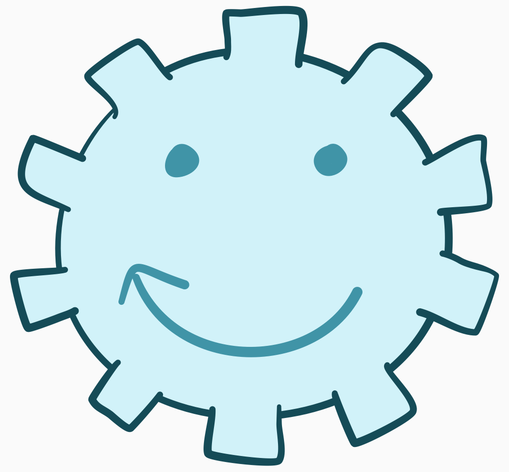


- Load-balancing queueing system with *redundancy*
- Each job has *size*, but servers have random *slowdowns*
- **Redundant-to-Idle-Queue (RIQ):** make copies of arriving jobs at *all* idle servers (and *only* idle servers)

Gardner, Harchol-Balter, and Scheller-Wolf: A Better Model for Job Redundancy: Decoupling Server Slowdown and Job Size (2016)

Ensuring stability with redundancy



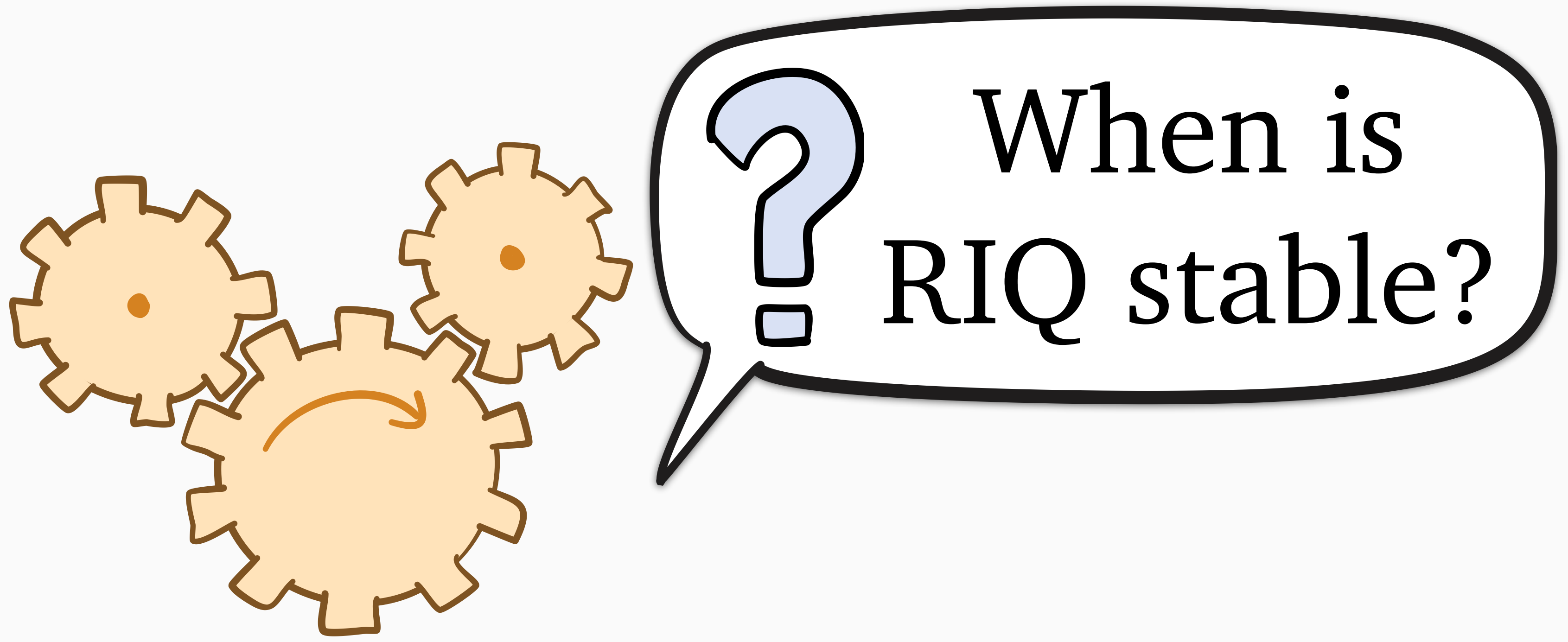


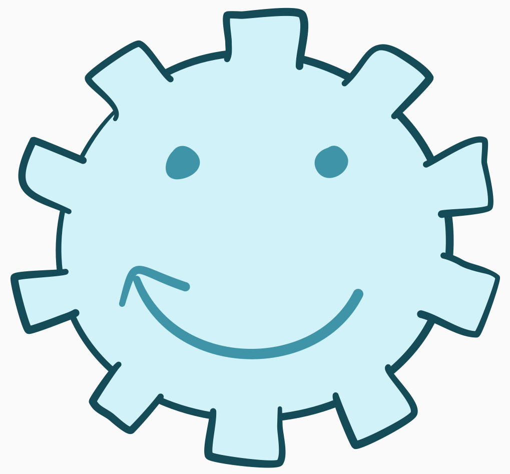
- Load-balancing queueing system with *redundancy*
- Each job has *size*, but servers have random *slowdowns*
- **Redundant-to-Idle-Queue (RIQ):** make copies of arriving jobs at *all* idle servers (and *only* idle servers)

- M/G/1 queue with *vacations*

Gardner, Harchol-Balter, and Scheller-Wolf: A Better Model for Job Redundancy: Decoupling Server Slowdown and Job Size (2016)

Ensuring stability with redundancy



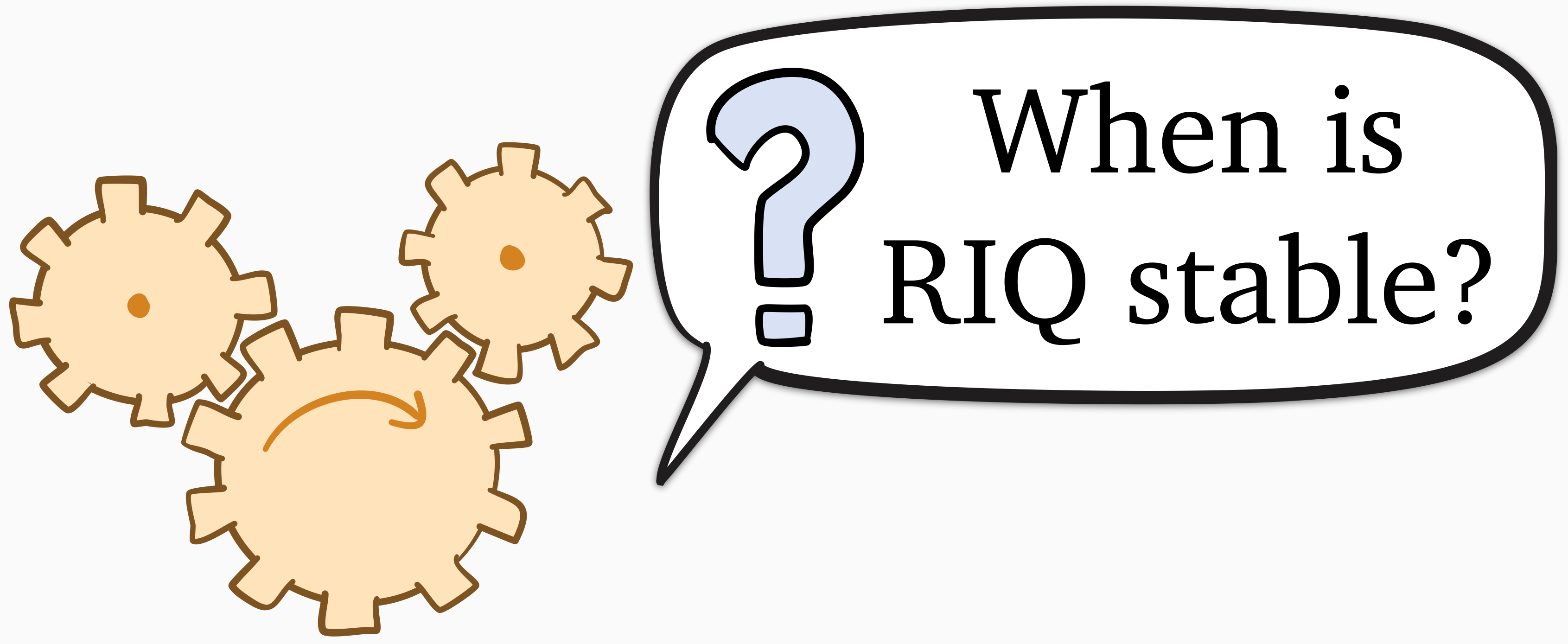


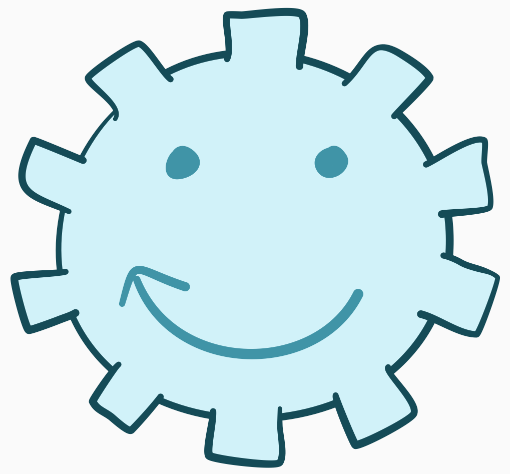
- Load-balancing queueing system with *redundancy*
- Each job has *size*, but servers have random *slowdowns*
- **Redundant-to-Idle-Queue (RIQ):** make copies of arriving jobs at *all* idle servers (and *only* idle servers)

- M/G/1 queue with *vacations*
- Whenever a server goes idle, assume a job copy is immediately created (i.e. starts a vacation)

Gardner, Harchol-Balter, and Scheller-Wolf: A Better Model for Job Redundancy: Decoupling Server Slowdown and Job Size (2016)

Ensuring stability with redundancy



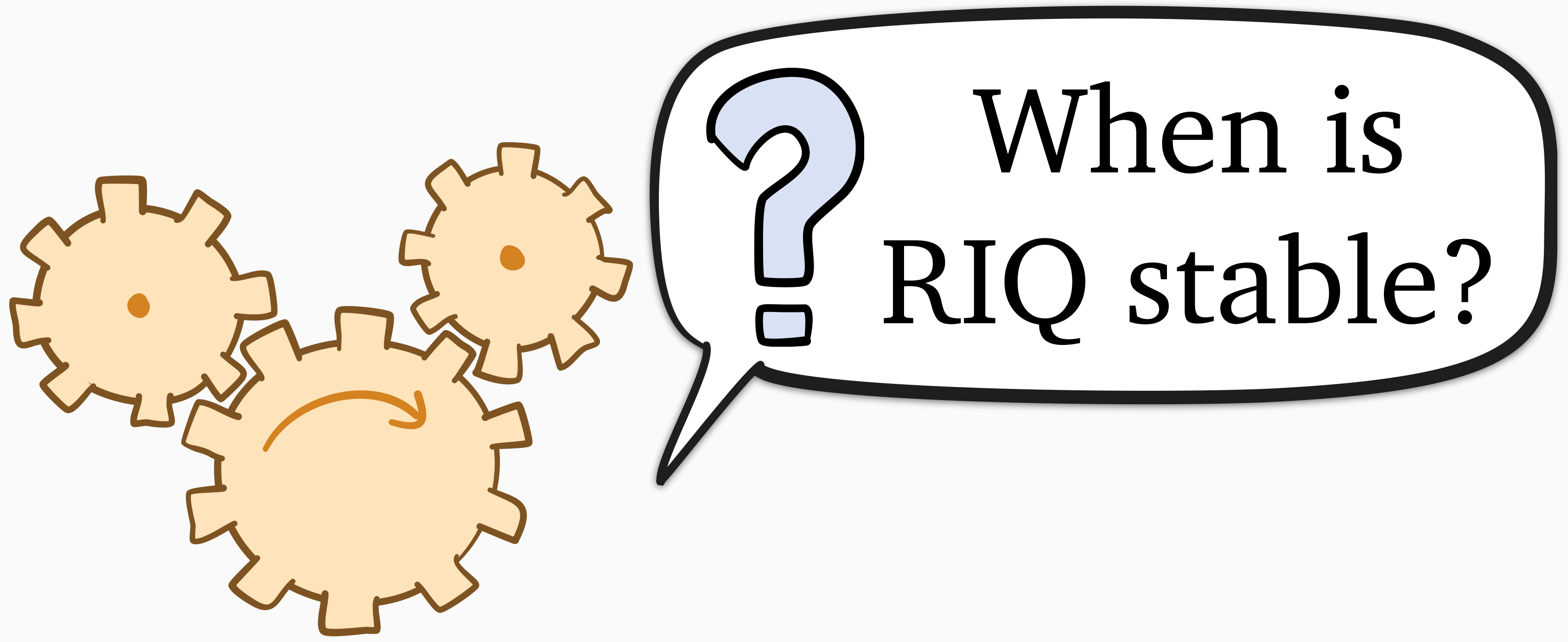


- Load-balancing queueing system with *redundancy*
- Each job has *size*, but servers have random *slowdowns*
- **Redundant-to-Idle-Queue (RIQ):** make copies of arriving jobs at *all* idle servers (and *only* idle servers)

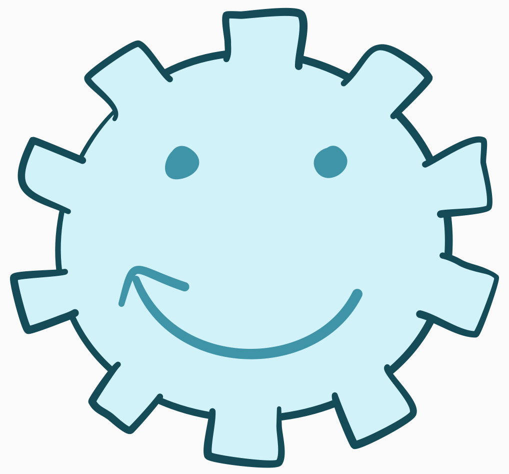
- M/G/1 queue with *vacations*
- Whenever a server goes idle, assume a job copy is immediately created (i.e. starts a vacation)
- Upper bound on RIQ system

Gardner, Harchol-Balter, and Scheller-Wolf: A Better Model for Job Redundancy: Decoupling Server Slowdown and Job Size (2016)

Ensuring stability with redundancy



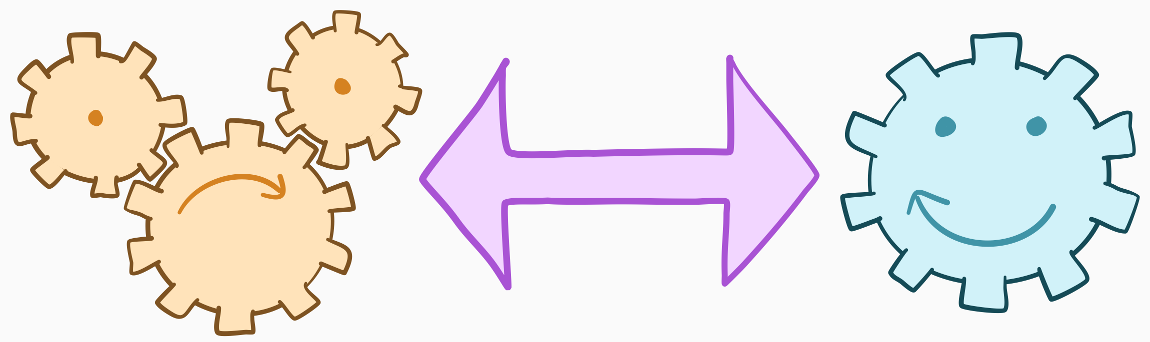
- Load-balancing queueing system with *redundancy*
- Each job has *size*, but servers have random *slowdowns*
- **Redundant-to-Idle-Queue (RIQ):** make copies of arriving jobs at *all* idle servers (and *only* idle servers)



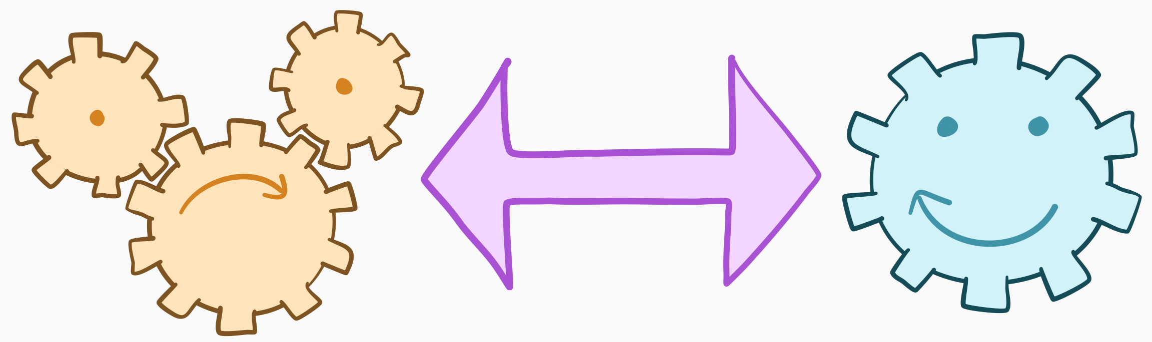
- M/G/1 queue with *vacations*
- Whenever a server goes idle, assume a job copy is immediately created (i.e. starts a vacation)
- Upper bound on RIQ system
- Implies RIQ is maximally stable

Gardner, Harchol-Balter, and Scheller-Wolf: A Better Model for Job Redundancy: Decoupling Server Slowdown and Job Size (2016)

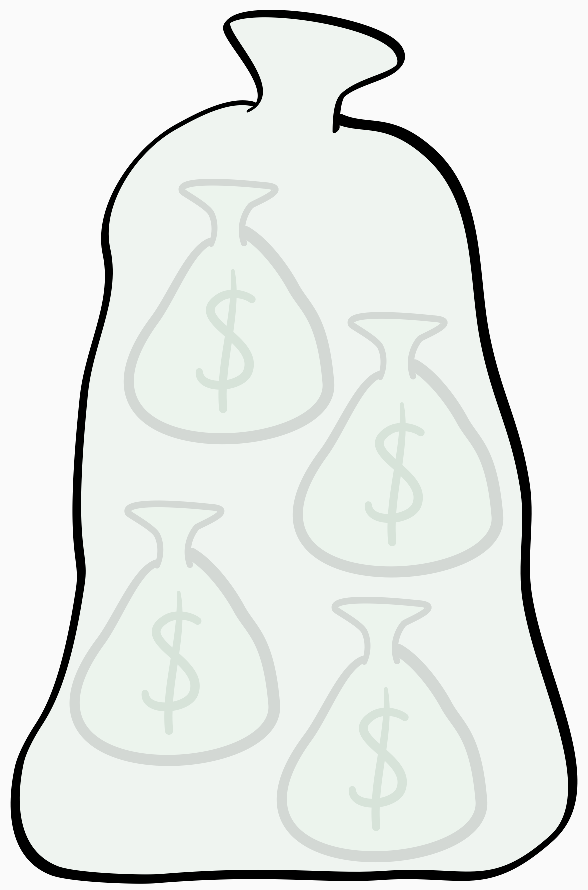
Classifying coupling techniques

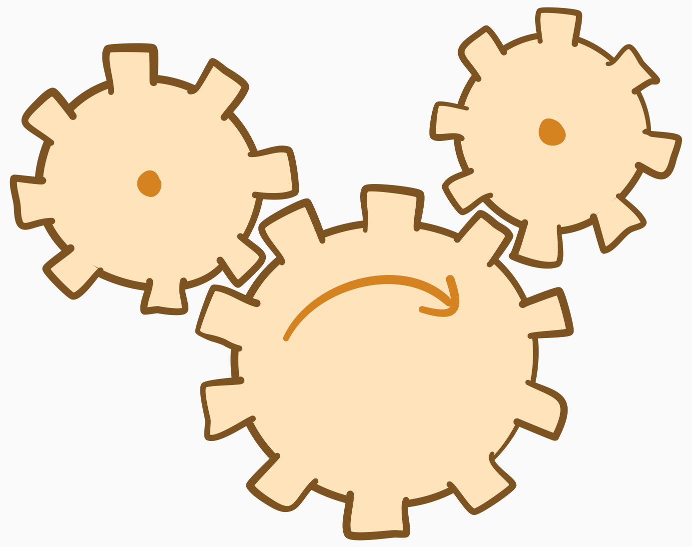
	A. Every sample path	B. Steady-state distribution
1. More information	A1	B1
2. Fewer constraints	A2 M/M/k vs. M/M/1	B2
3. Simpler dynamics	A3 SIS epidemics	B3

Classifying coupling techniques

	A. Every sample path	B. Steady-state distribution
1. More information	A1	B1
2. Fewer constraints	A2 M/M/k vs. M/M/1	B2
3. Simpler dynamics	A3 SIS epidemics Queues with redundancy	B3

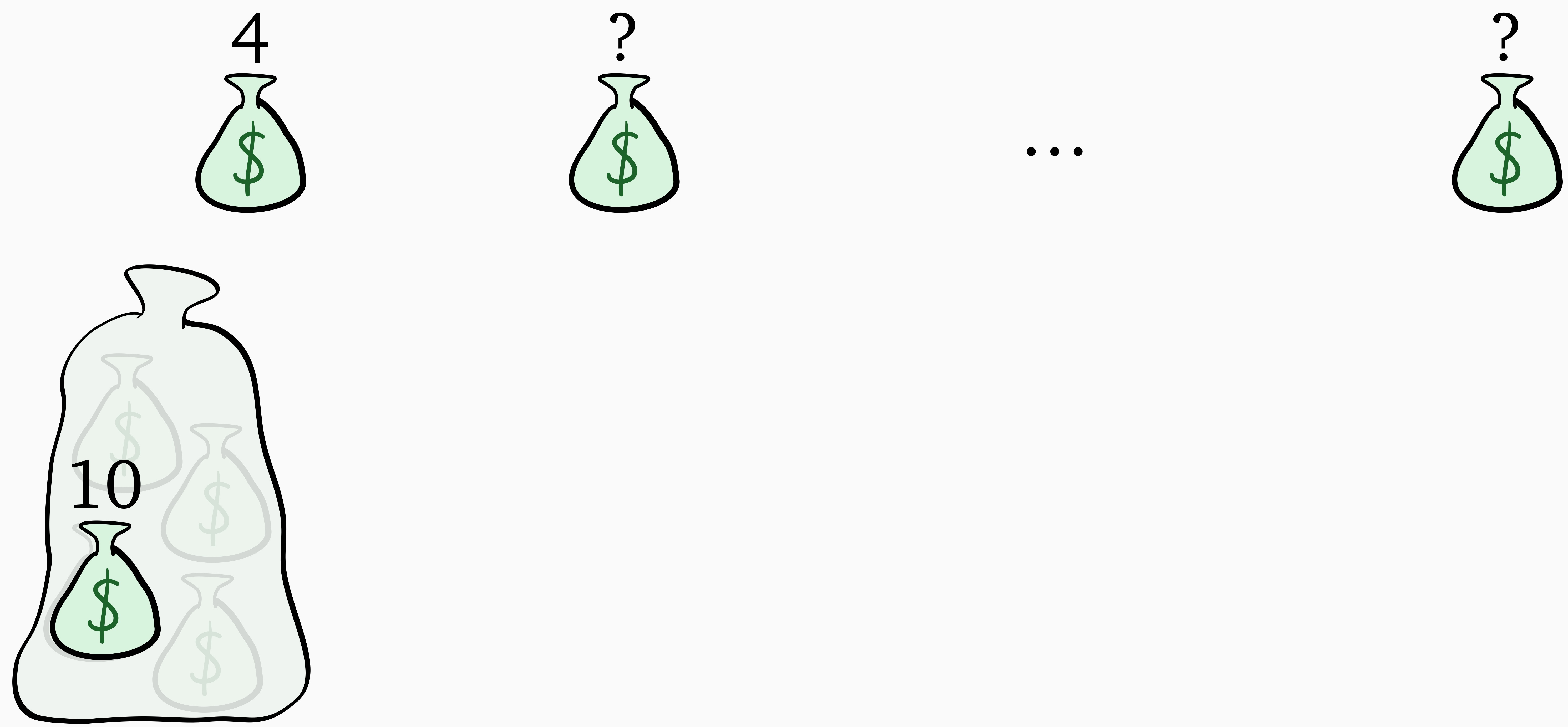
Stochastic online knapsack

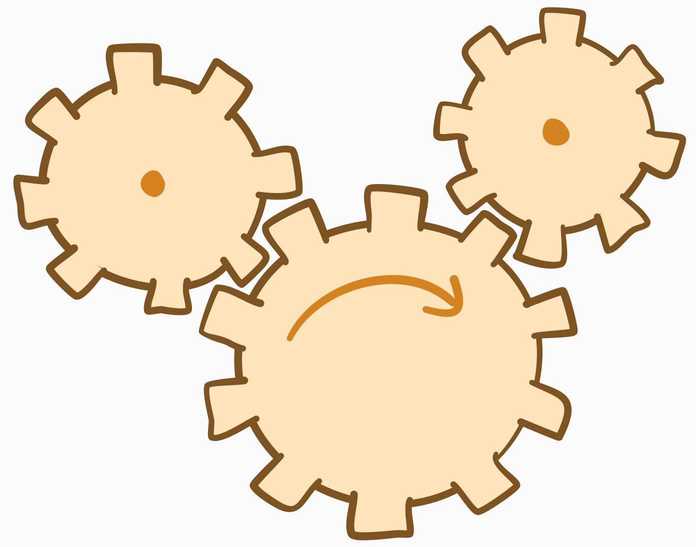




- T items arrive sequentially, with values $V_1, V_2, \ldots, V_T$
- Value distributions known to controller
- Can select up to B items

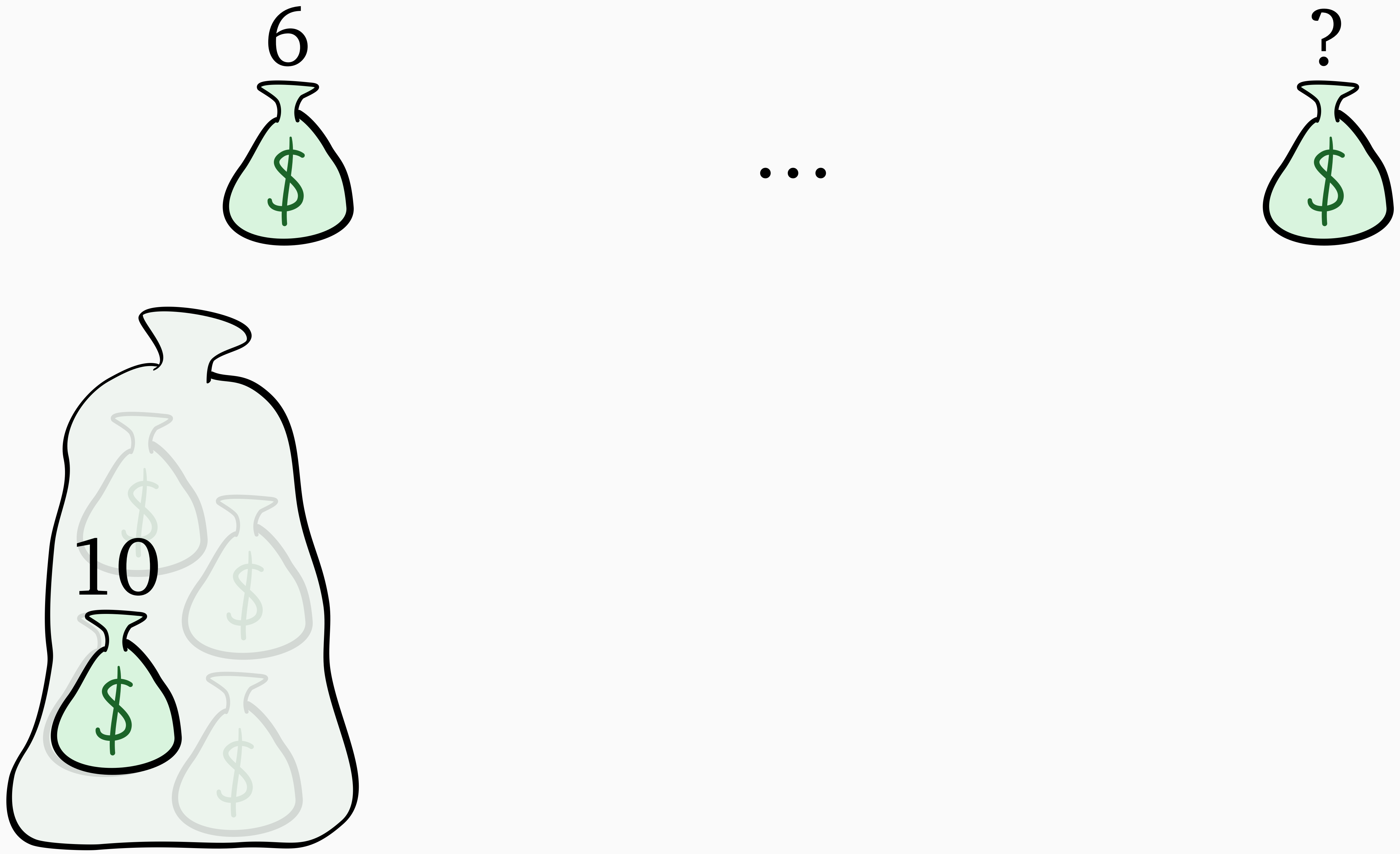
Stochastic online knapsack

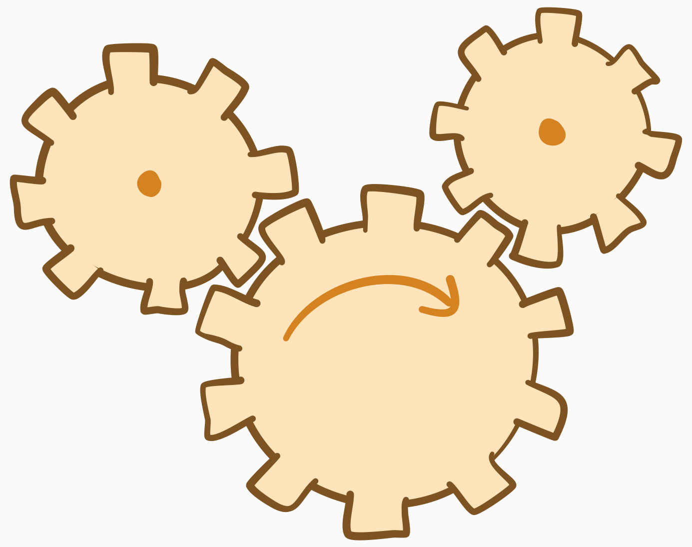




- T items arrive sequentially, with values $V_1, V_2, \ldots, V_T$
- Value distributions known to controller
- Can select up to B items

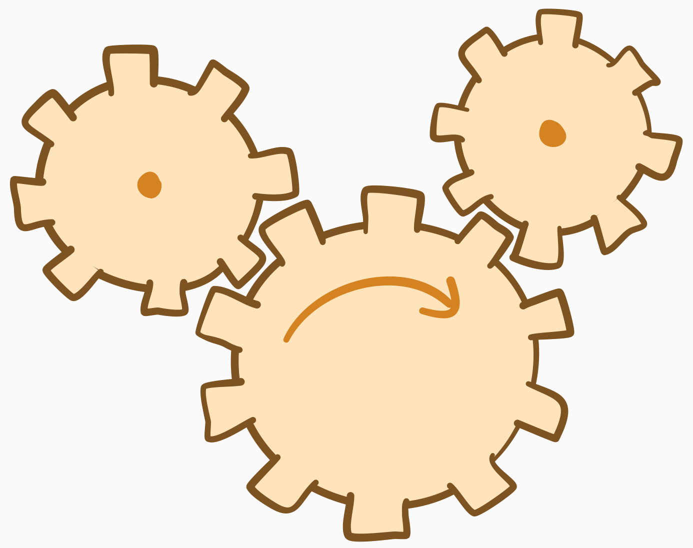
Stochastic online knapsack

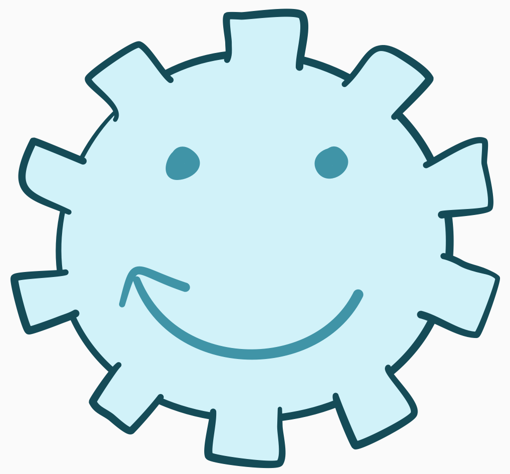




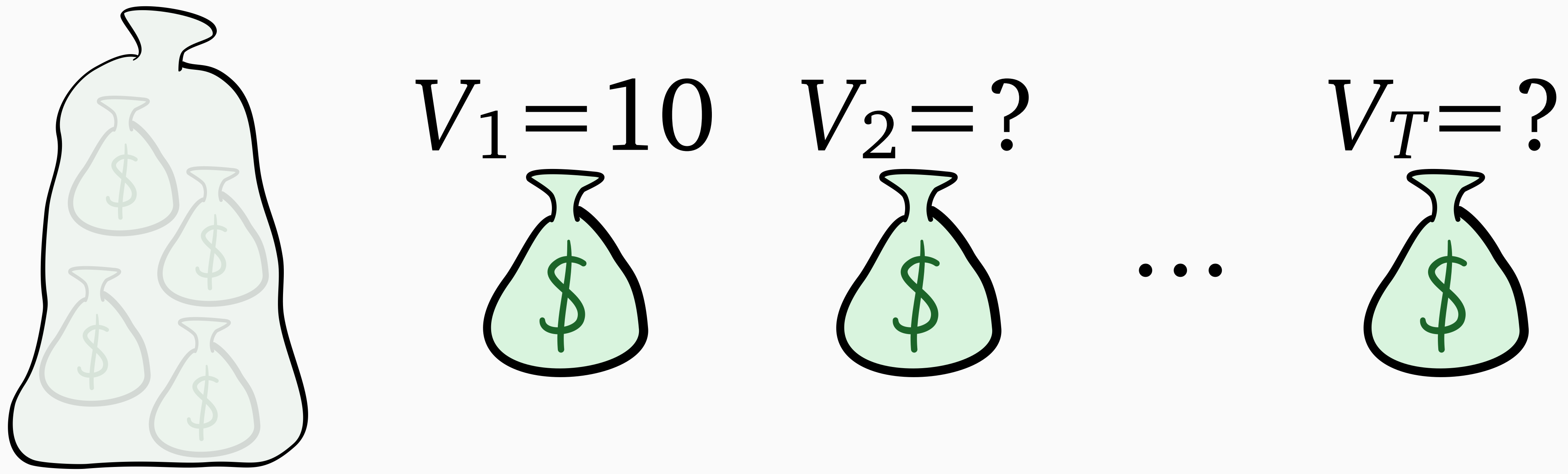
- T items arrive sequentially, with values $V_1, V_2, \ldots, V_T$
- Value distributions known to controller
- Can select up to B items

“Weakly-coupled” problems

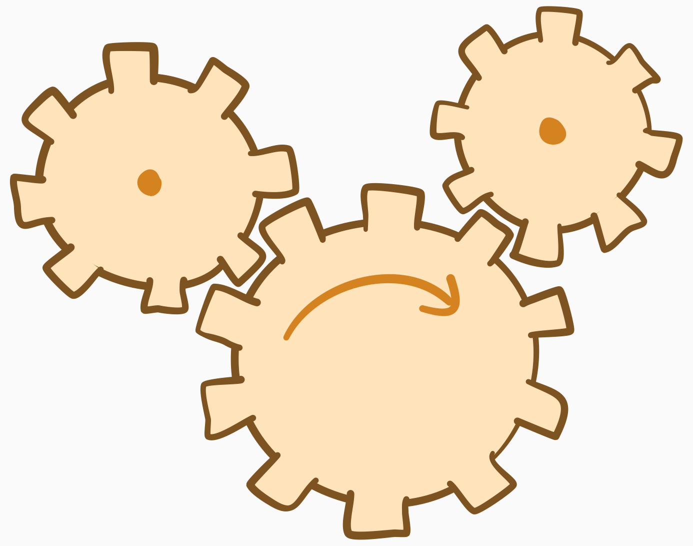


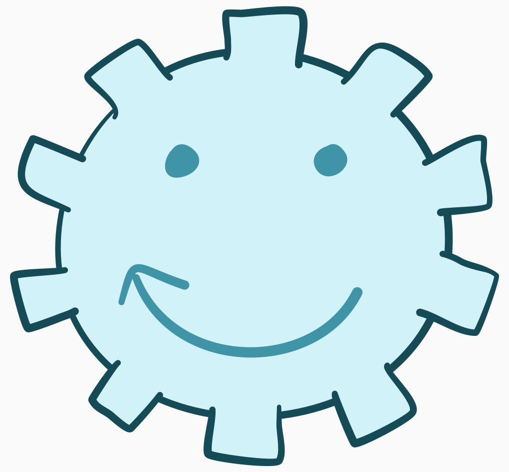


- Multiple “easy” control problems stitched together by joint constraints
- E.g. **Online Knapsack**:
  - T items arrive sequentially, with values V_1, V_2, …, V_T
  - Can select up to B items

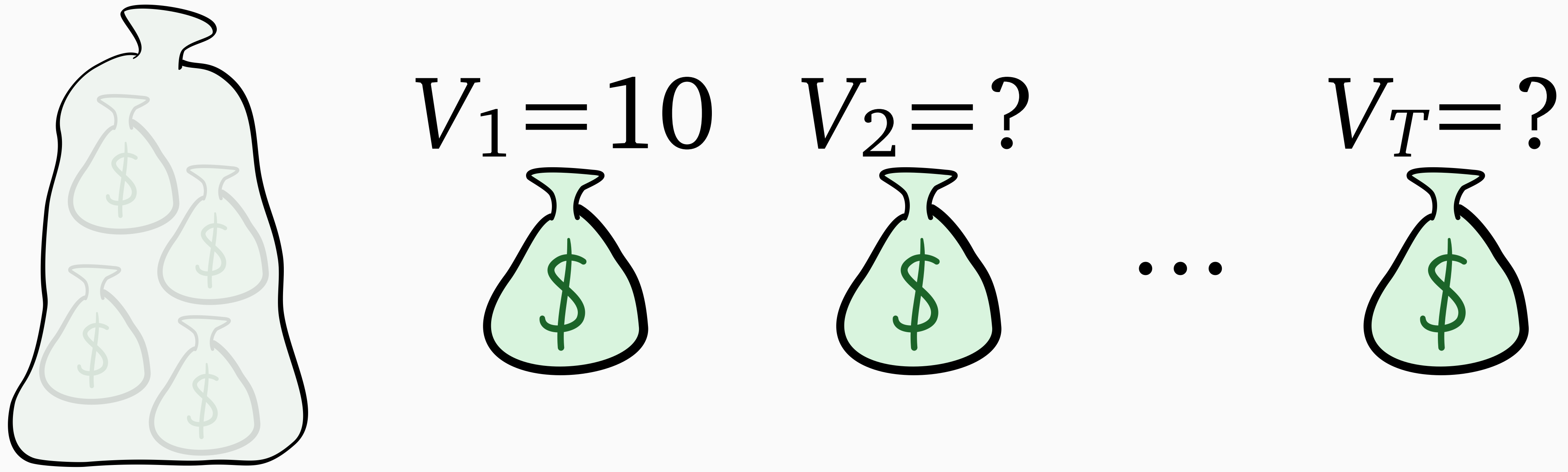


“Weakly-coupled” problems



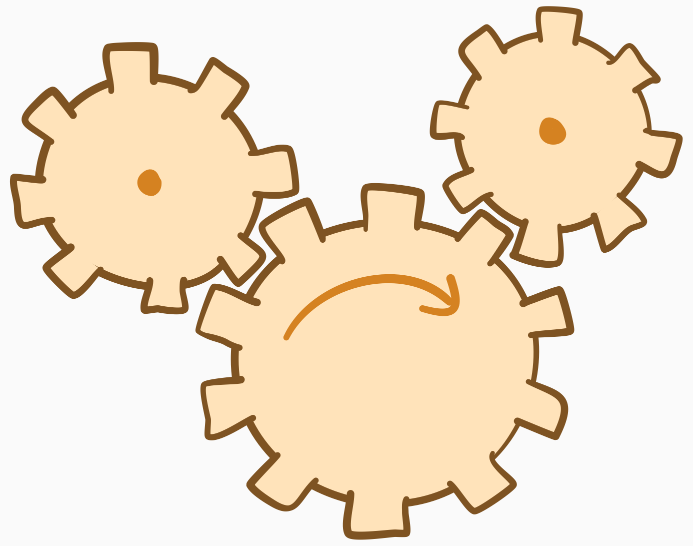


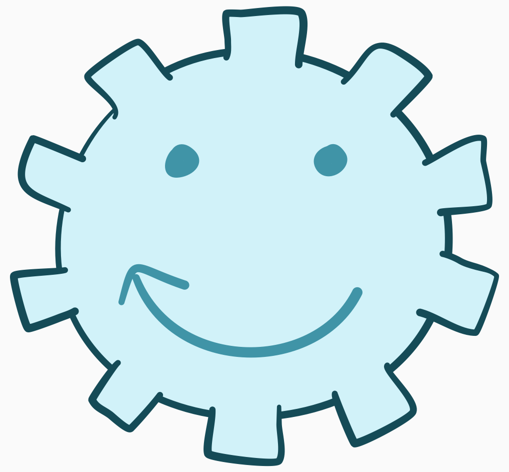
- Multiple “easy” control problems stitched together by joint constraints
- E.g. **Online Knapsack**:
  - T items arrive sequentially, with values $V_1, V_2, \dots, V_T$
  - Can select up to B items



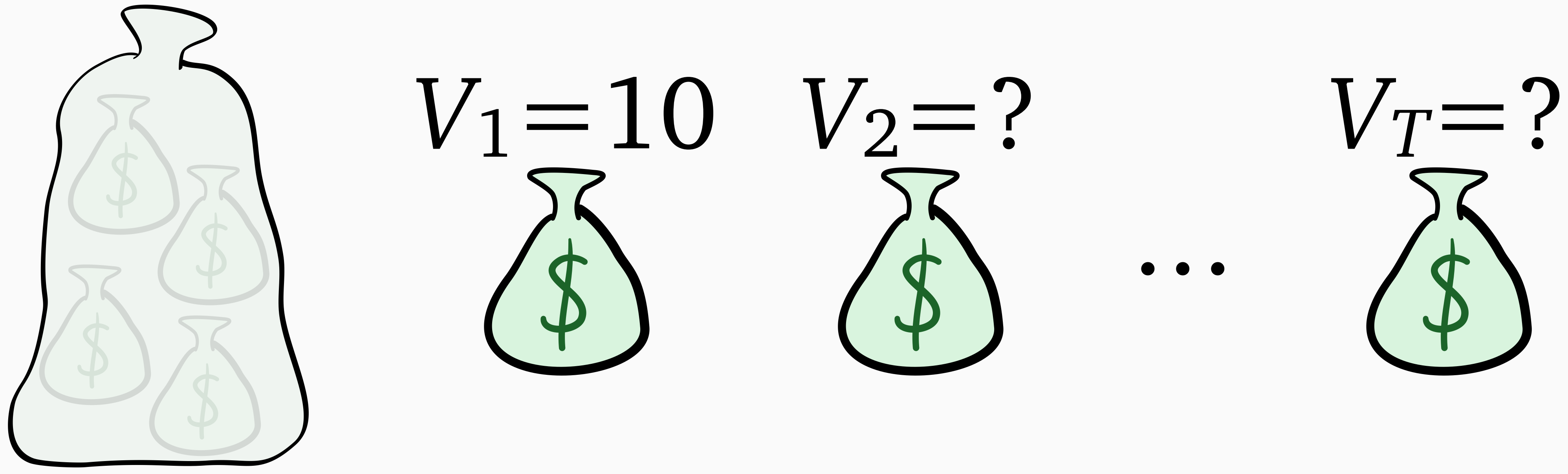
- Separate easy problems…

“Weakly-coupled” problems



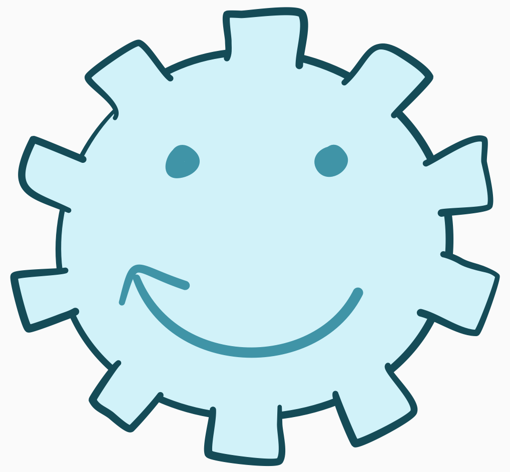
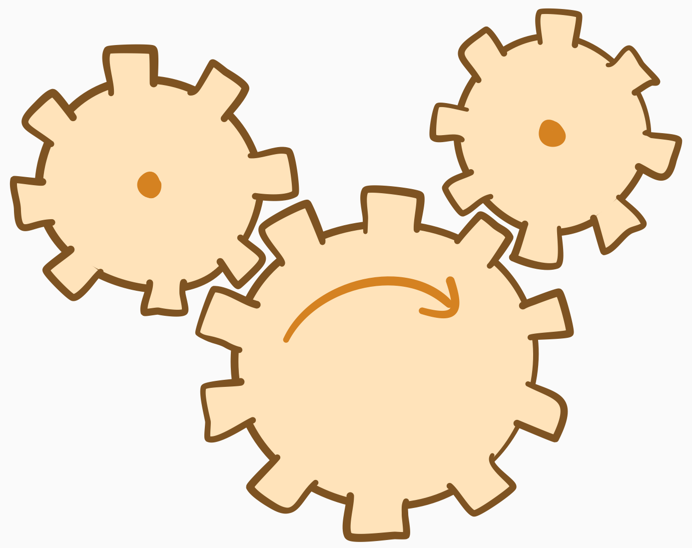
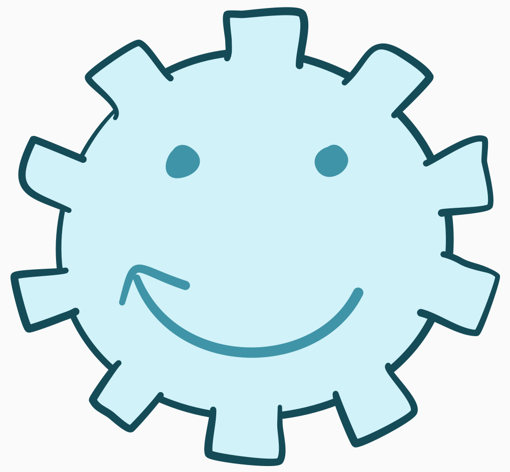


- Multiple “easy” control problems stitched together by joint constraints
- E.g. **Online Knapsack**:
  - T items arrive sequentially, with values $V_1, V_2, \ldots, V_T$
  - Can select up to B items

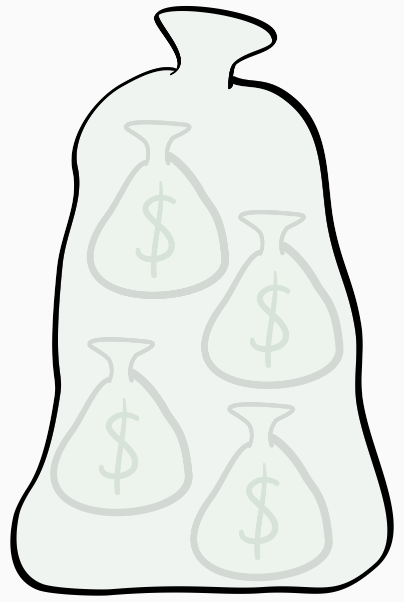


- Separate easy problems…
- E.g. **Online Purchasing**:
  Select items from incoming stream, with values *and costs*
  (no budget or limit on # items accepted)

Online knapsack: results

$\frac{1}{2}$  $\leq$  $\leq$ 

Online knapsack



$V_1 = 10$ $V_2 = ?$ $\cdots$ $V_T = ?$

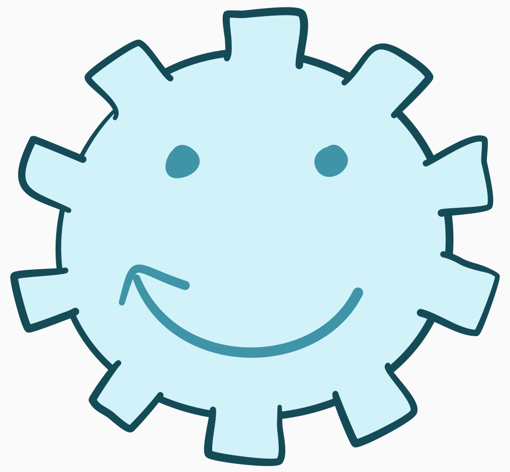
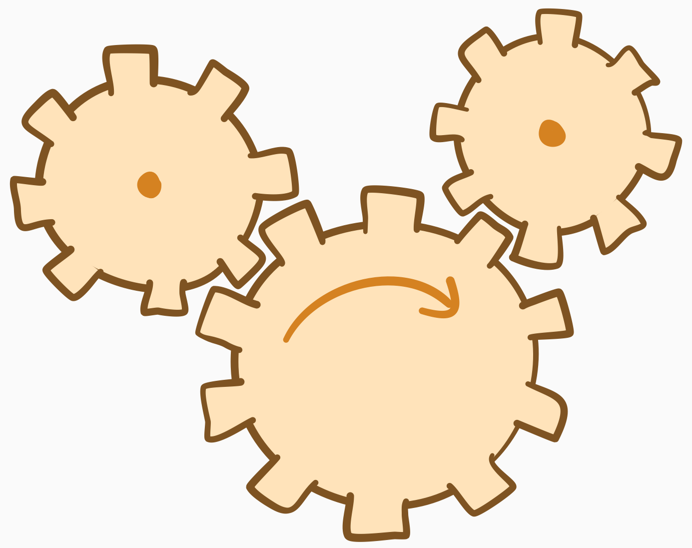
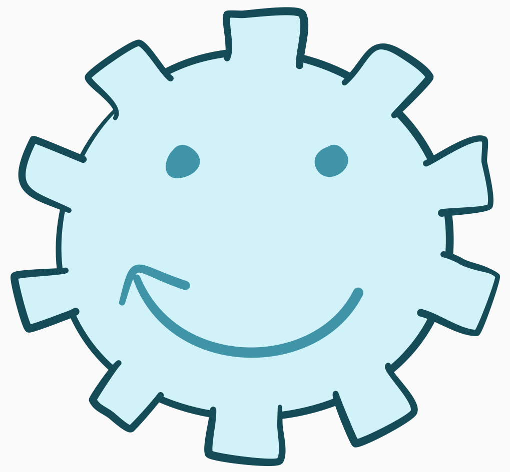
Online purchasing

$V_1 = 10$ $V_2 = ?$ $\cdots$ $V_T = ?$

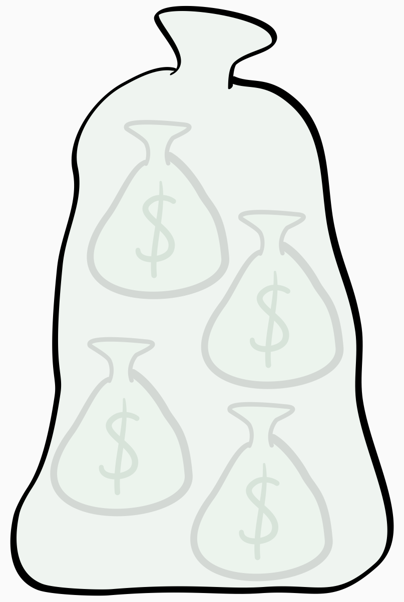
$c_1 = \lambda$ $c_2 = \lambda$ $c_T = \lambda$

Kamesh Munagala: Lecture notes on "Optimization and Decision-Making under Uncertainty" (2016)

Online knapsack: results

$\frac{1}{2}$  $\leq$  $\leq$ 

Online knapsack



$V_1 = 10$ $V_2 = ?$ $\cdots$ $V_T = ?$

Online purchasing

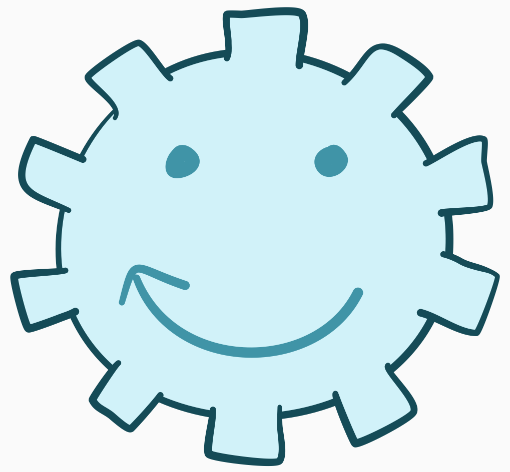
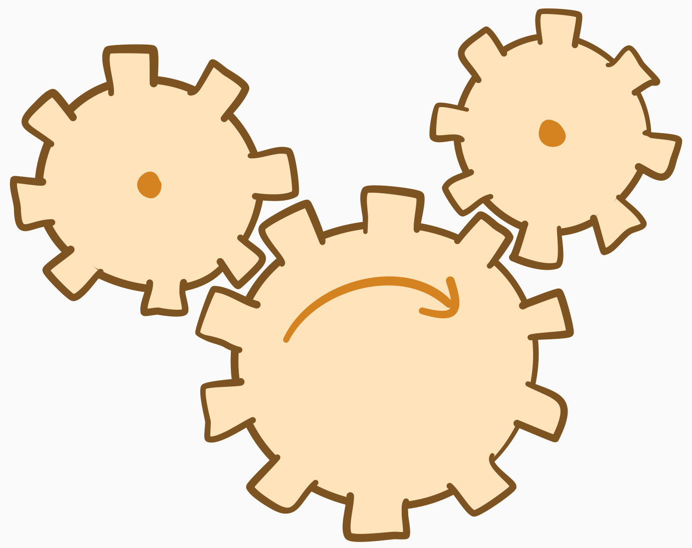
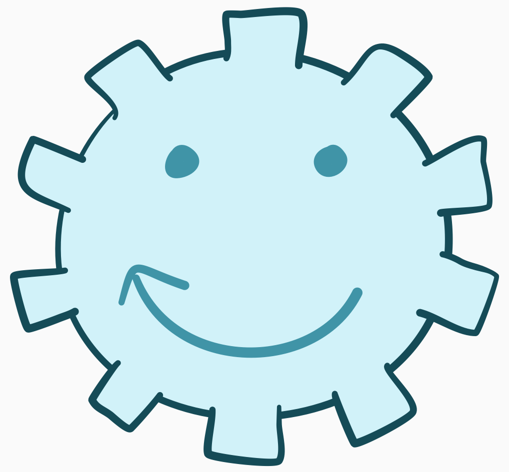
$V_1 = 10$ $V_2 = ?$ $\cdots$ $V_T = ?$

$c_1 = \lambda$ $c_2 = \lambda$ $c_T = \lambda$

Can choose a ‘cost’ λ s.t. accepting all $V_t > \lambda$ while space available
gives a 2-approximation

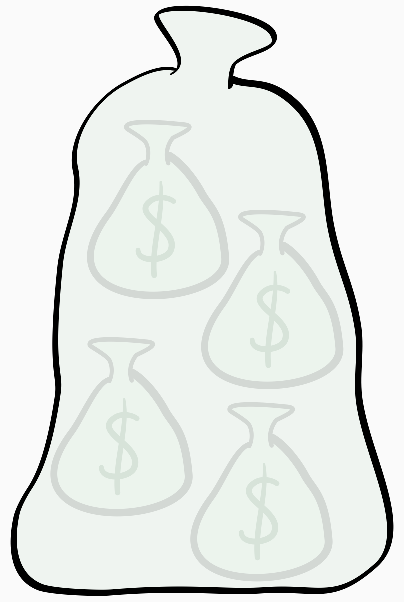
Kamesh Munagala: Lecture notes on “Optimization and Decision-Making under Uncertainty” (2016)

Online knapsack: results

$\frac{1}{2}$  $\leq$  $\leq$ 

Online knapsack

$V_1 = 10$ $V_2 = ?$ … $V_T = ?$



"balanced" threshold

Online purchasing

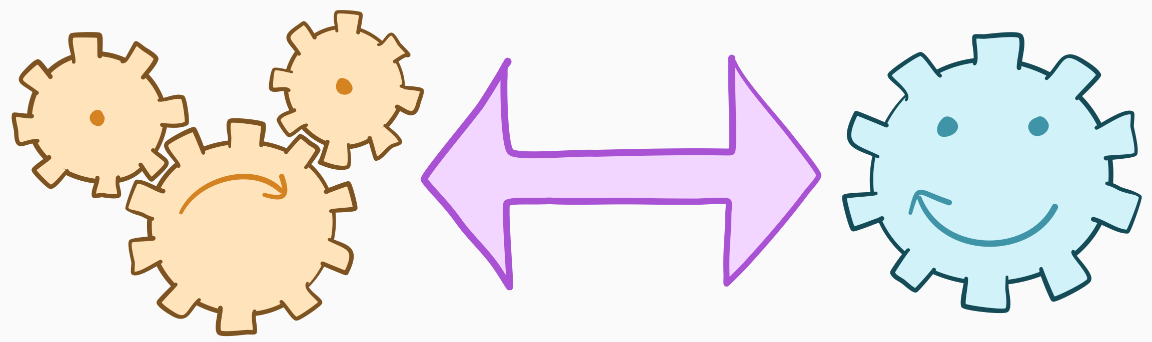
$V_1 = 10$ $V_2 = ?$ … $V_T = ?$

$c_1 = \lambda$ $c_2 = \lambda$ $c_T = \lambda$

Can choose a 'cost' λ s.t. accepting all $V_t > \lambda$ while space available
gives a 2-approximation

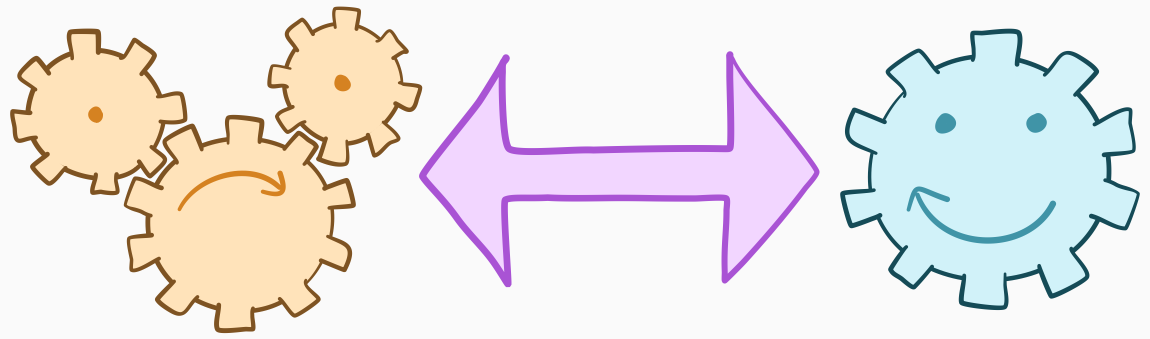
Kamesh Munagala: Lecture notes on "Optimization and Decision-Making under Uncertainty" (2016)

Classifying coupling techniques

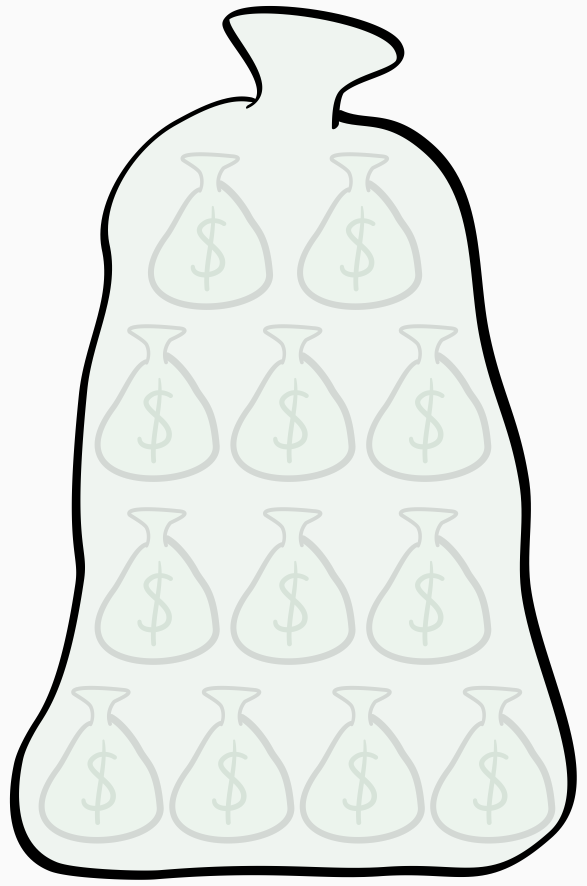


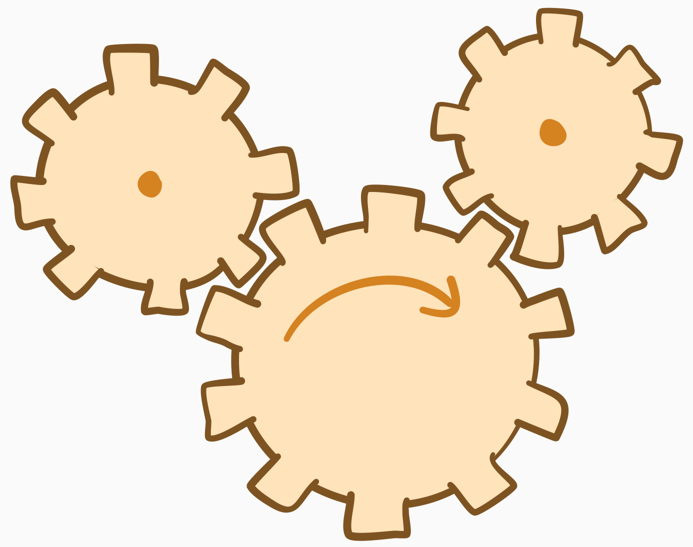
	A. Every sample path	B. Steady-state distribution
1. More information	A1	B1
2. Fewer constraints	A2 M/M/k vs. M/M/1	B2
3. Simpler dynamics	A3 Queues with redundancy SIS epidemics	B3

Classifying coupling techniques

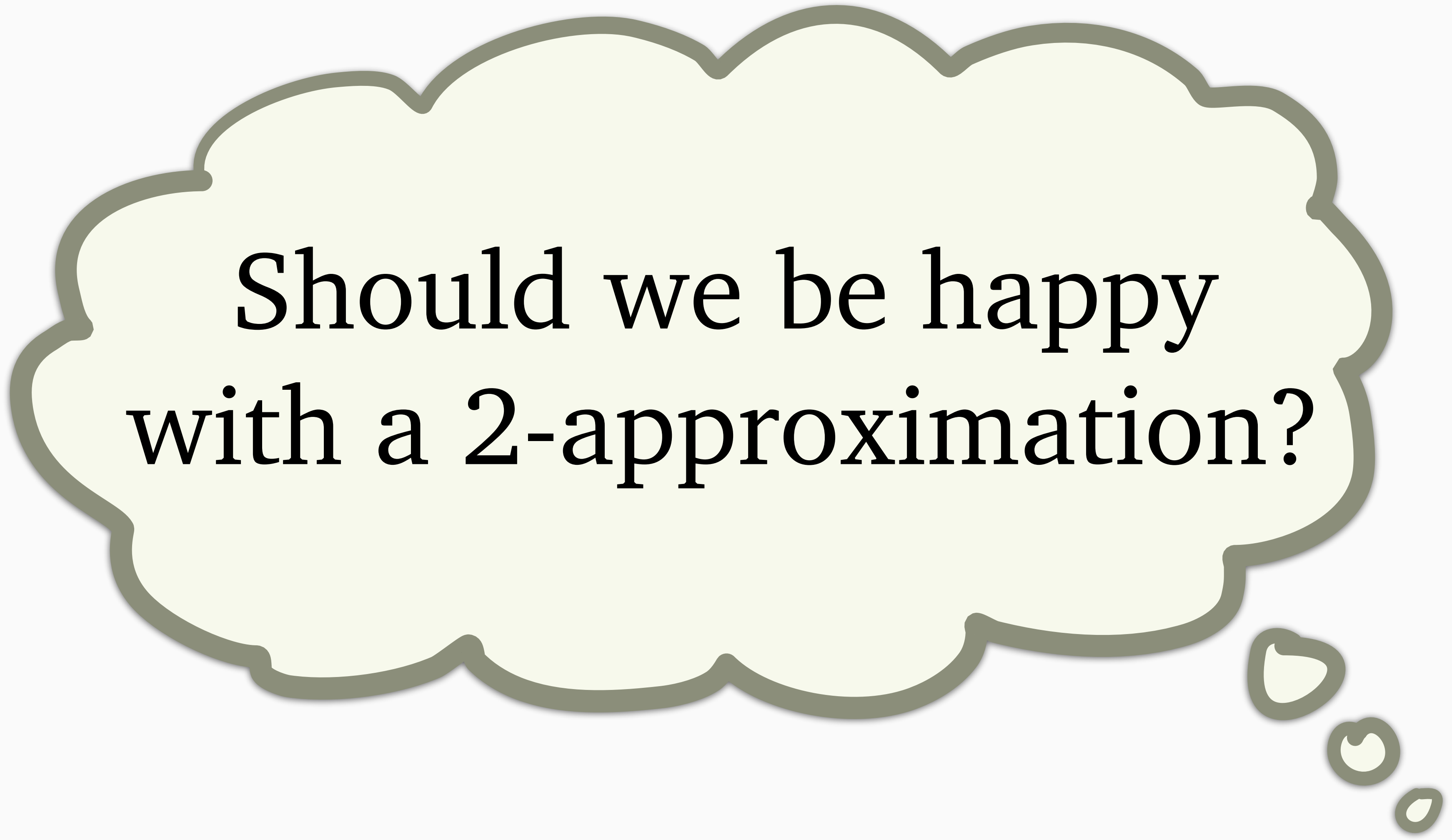
	A. Every sample path	B. Steady-state distribution
1. More information	A1	B1
2. Fewer constraints	A2 M/M/k vs. M/M/1 Online knapsack (via constraints-to-costs)	B2
3. Simpler dynamics	A3 Queues with redundancy SIS epidemics	B3

BIG online knapsacks

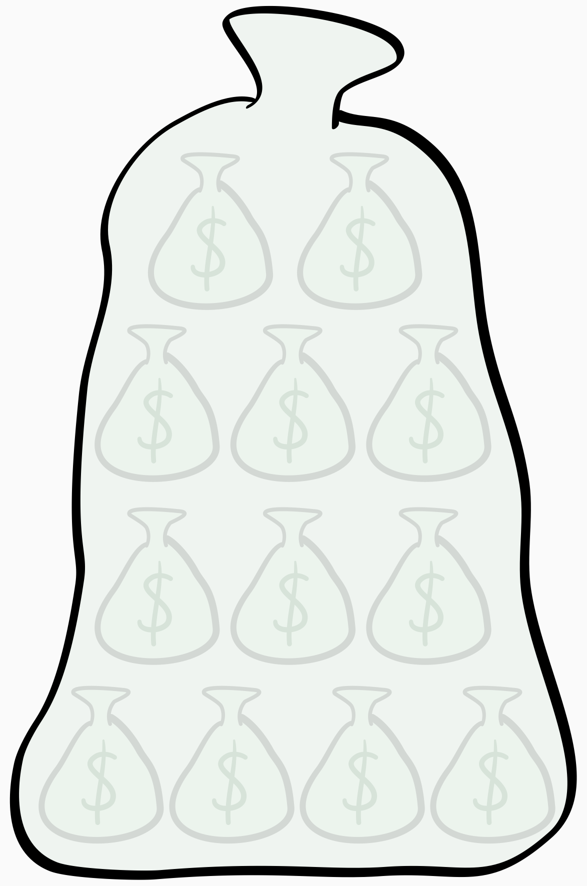


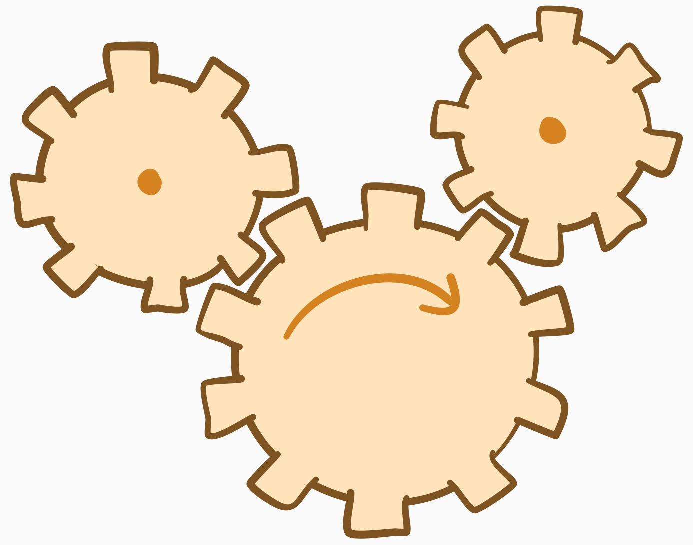


- T items arrive sequentially, with values $V_1, V_2, \ldots, V_T$
- Can select up to B items
- **Both B and T are large** (with say $B = 0.1\ T$)



BIG online knapsacks



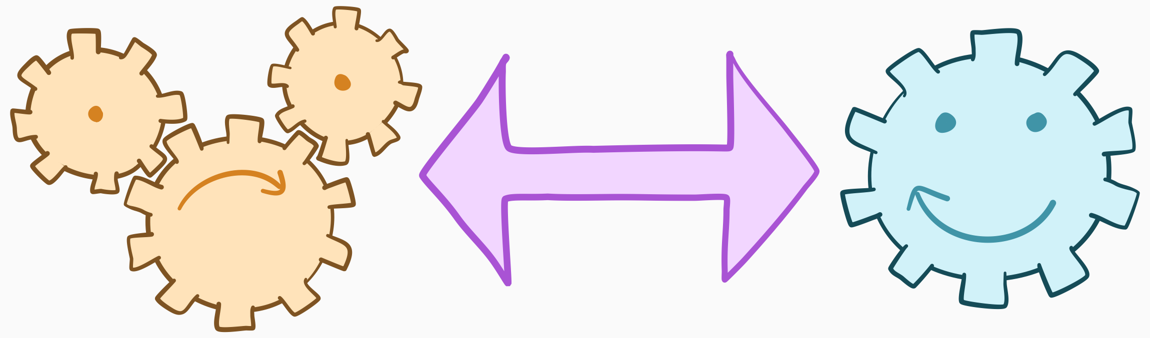


- T items arrive sequentially, with values $V_1, V_2, \ldots, V_T$
- Can select up to B items
- **Both B and T are large** (with say $B = 0.1\ T$)

Should we be happy with a 2-approximation?

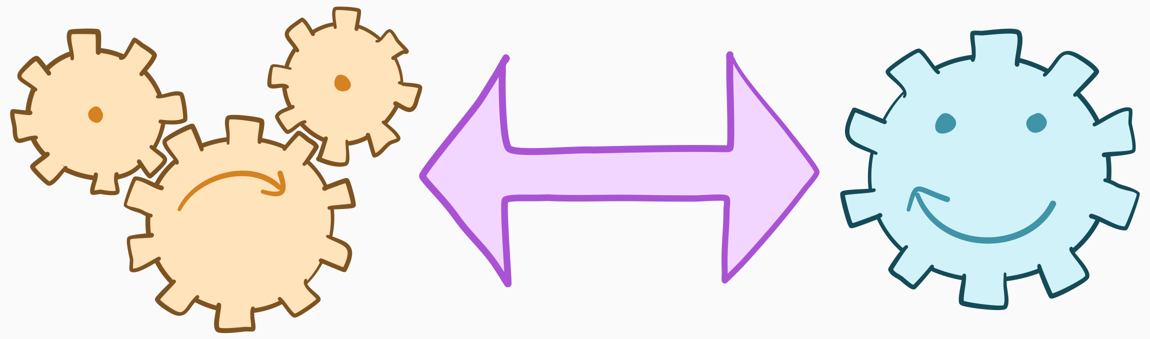
That means our **regret grows linearly with T**

Classifying coupling techniques



	A. Every sample path	B. Steady-state distribution
1. More information	A1	B1
2. Fewer constraints	A2 M/M/k vs. M/M/1 Online knapsack (via constraints-to-costs)	B2
3. Simpler dynamics	A3 SIS epidemics Queues with redundancy	B3

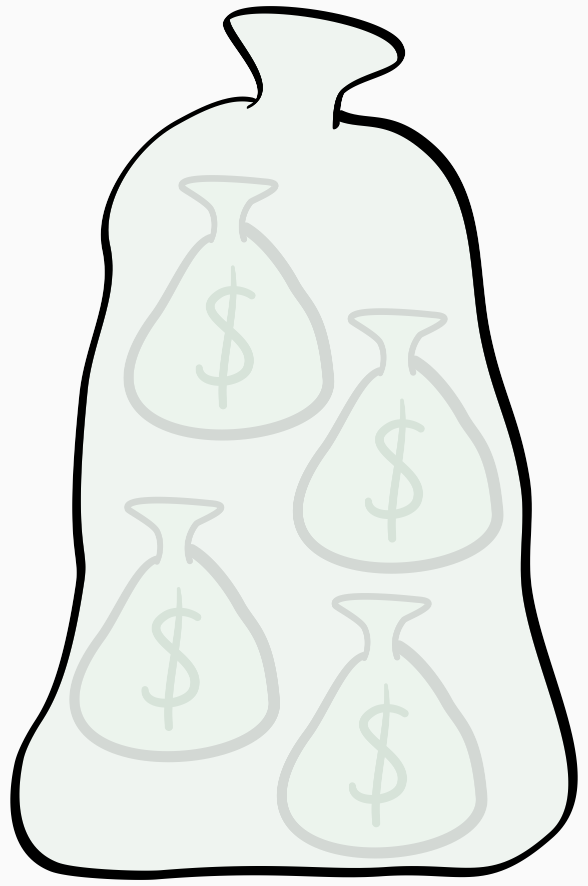
Classifying coupling techniques



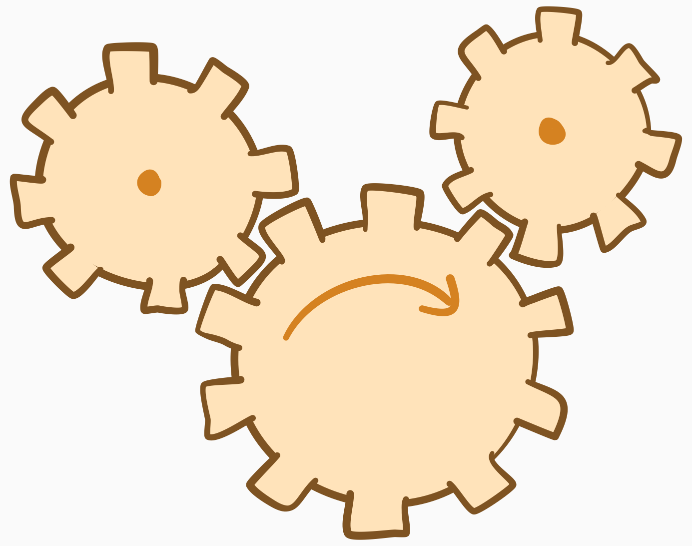
	A. Every sample path	B. Steady-state distribution
1. More information	A1 BIG online knapsack (stay tuned!)	B1
2. Fewer constraints	A2 M/M/k vs. M/M/1 Online knapsack (via constraints-to-costs)	B2
3. Simpler dynamics	A3 SIS epidemics Queues with redundancy	B3

In-Depth Study 1:
Online Resource Allocation

The (stochastic) online knapsack



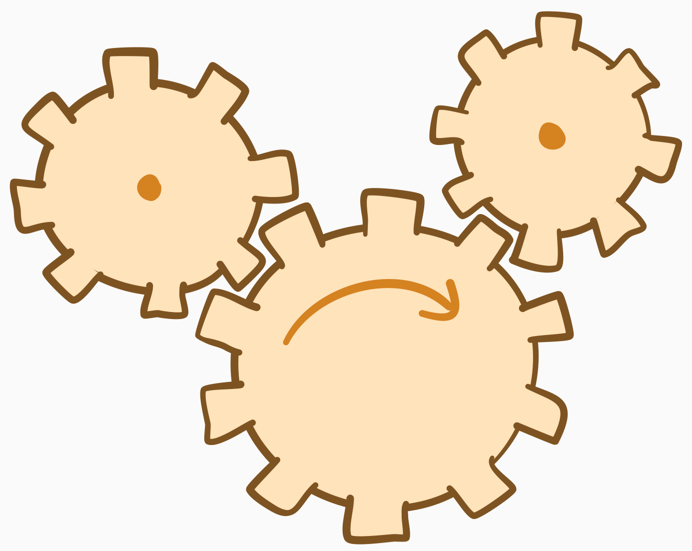
- T items arrive sequentially, with values $V_1, V_2, \ldots, V_T$
- Can select up to B items
- Values are i.i.d from distribution



$$V_t = \begin{cases} 5 & \text{with probability } p_5 \\ 10 & \text{with probability } p_{10} \\ 20 & \text{with probability } p_{20} \end{cases}$$

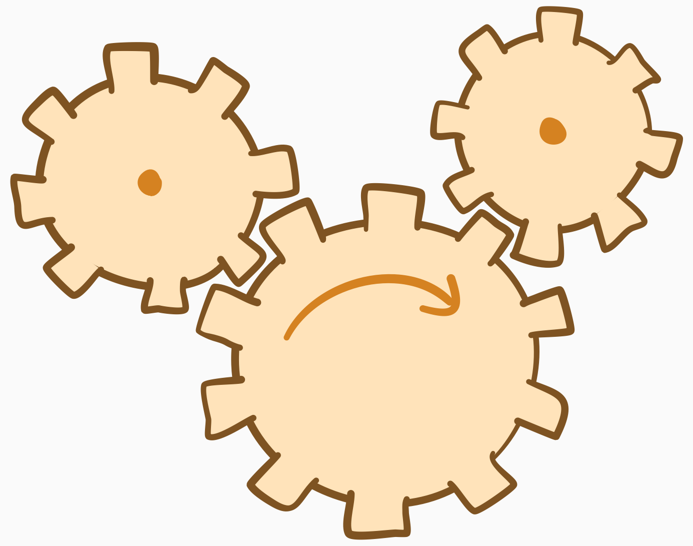
The (stochastic) online knapsack

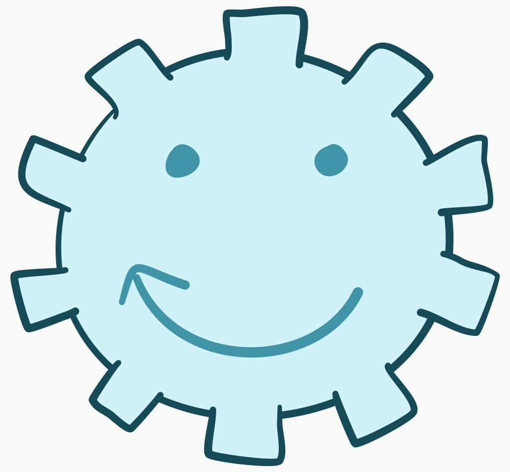
- T items arrive sequentially, with values $V_1, V_2, \ldots, V_T$
- Can select up to B items
- Values are i.i.d from distribution



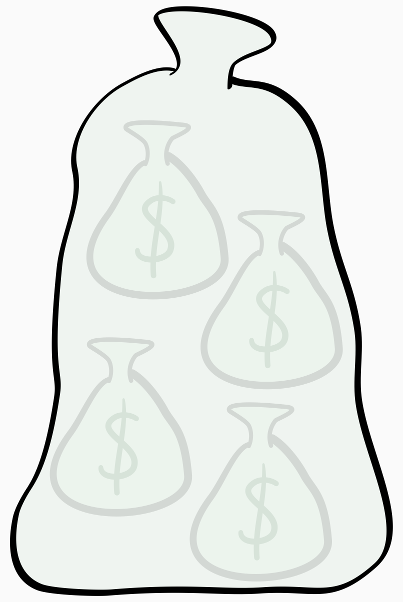
$$V_t = \begin{cases} 5 & \text{with probability } p_5 \\ 10 & \text{with probability } p_{10} \\ 20 & \text{with probability } p_{20} \end{cases}$$

Prophet benchmarks



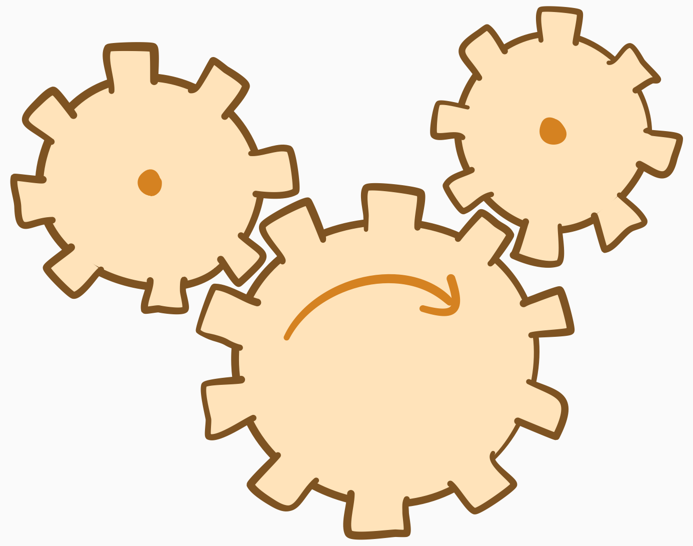


- Uncertainty about the future
- **Online Knapsack**

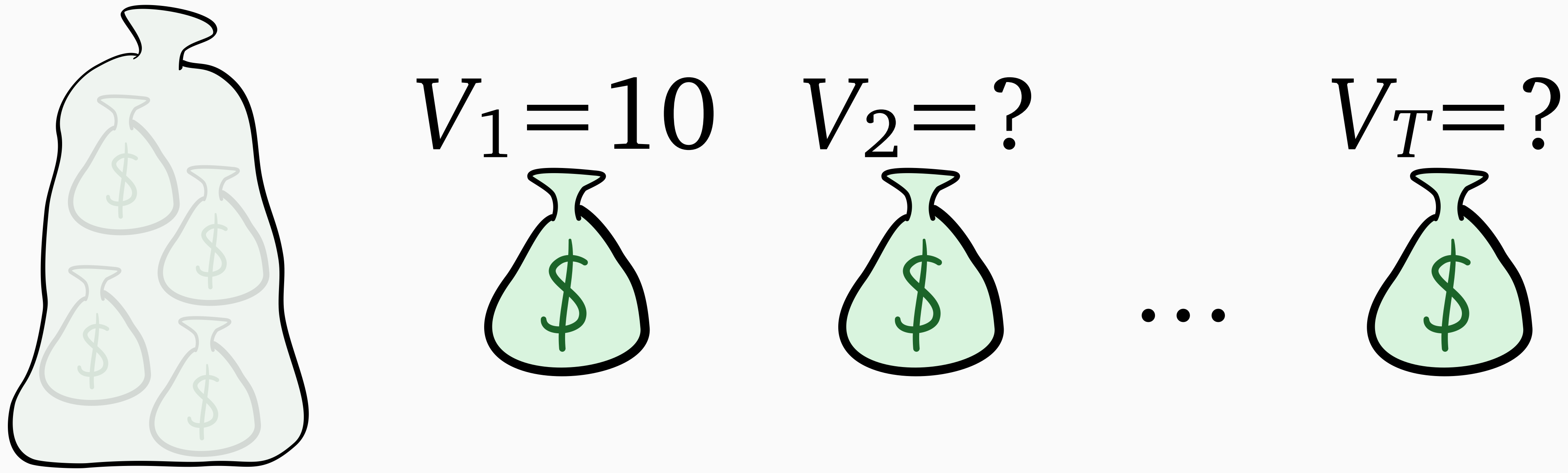


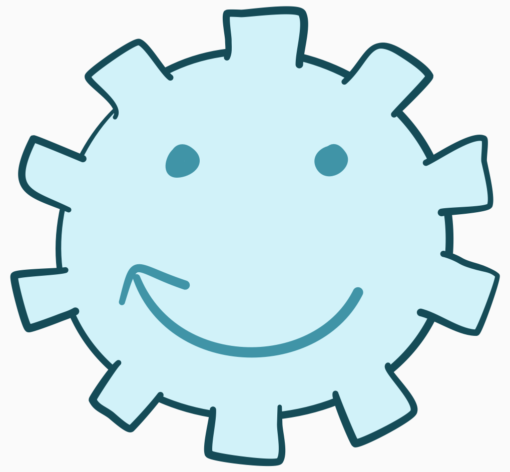
$V_1 = 10$ $V_2 = ?$ $V_T = ?$

Prophet benchmarks



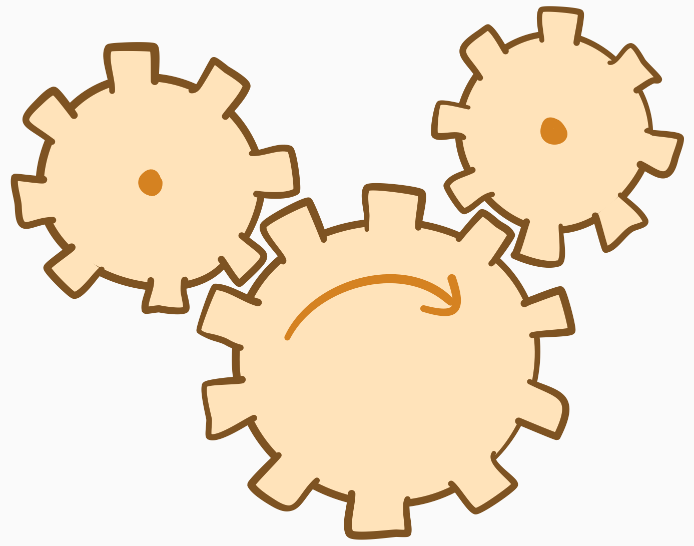
- Uncertainty about the future
- **Online Knapsack**

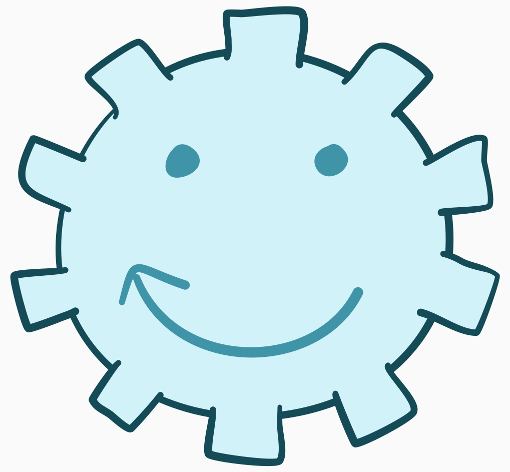




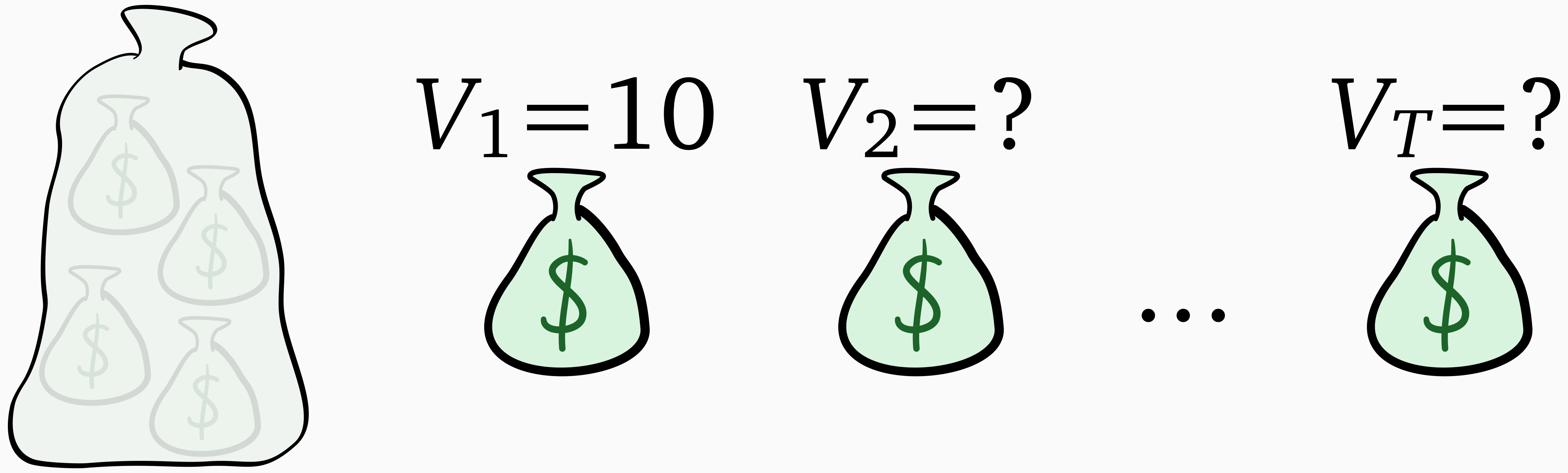
- Knows the future!

Prophet benchmarks

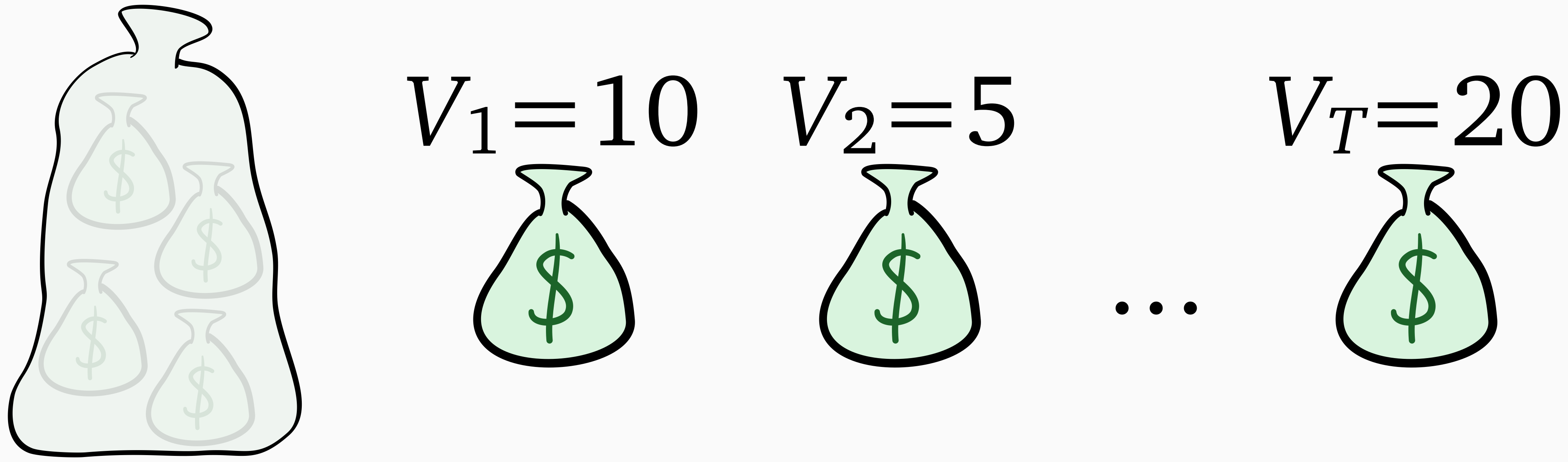




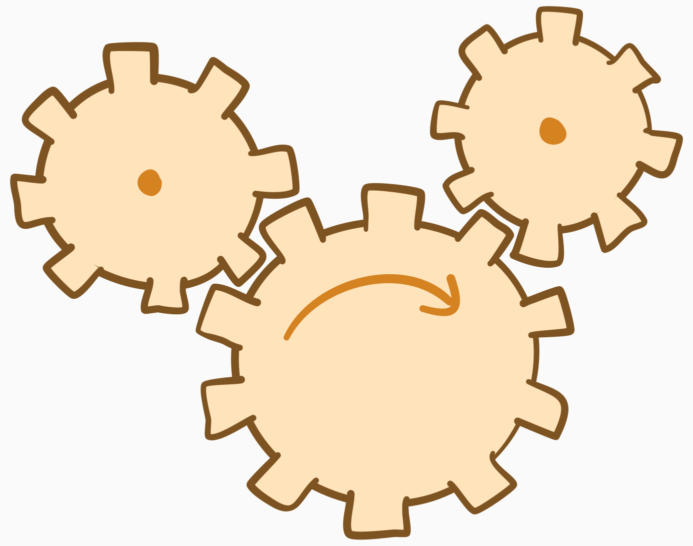
- Uncertainty about the future
- **Online Knapsack**

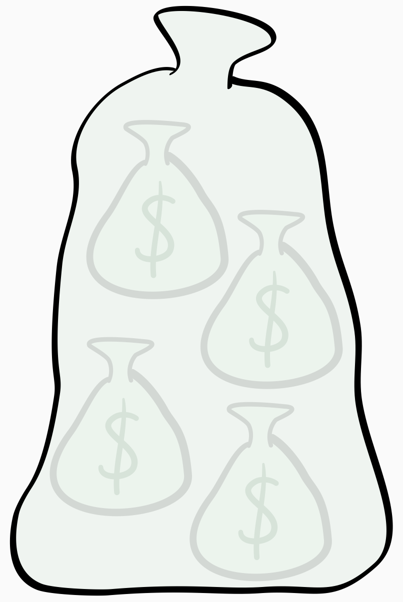


- Knows the future!
- **Offline Knapsack**

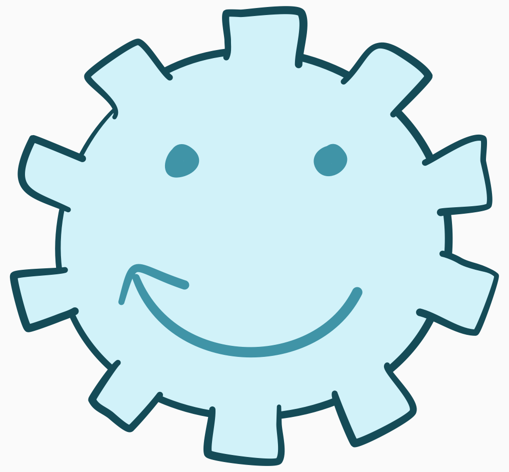


Coupling to **OFFLINE**

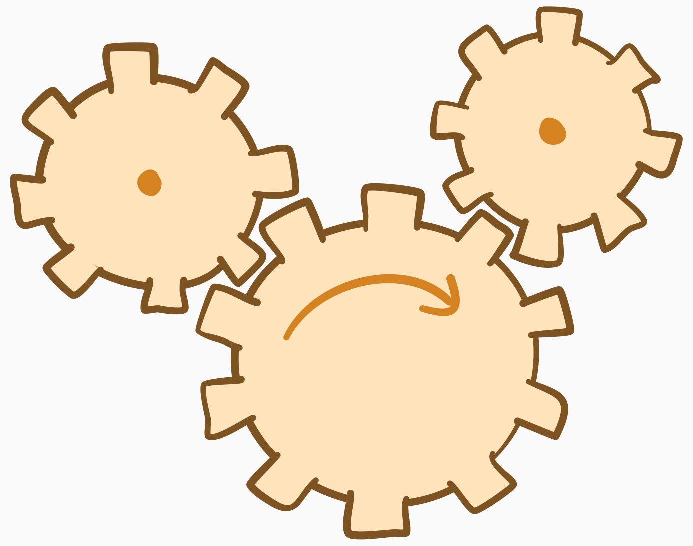
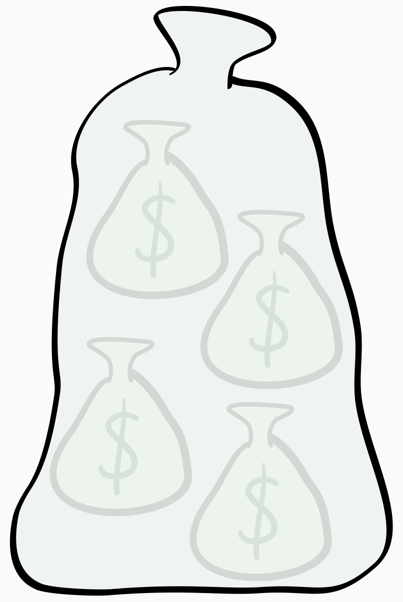




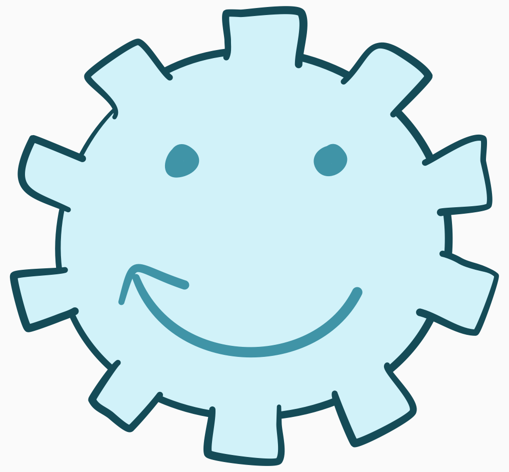
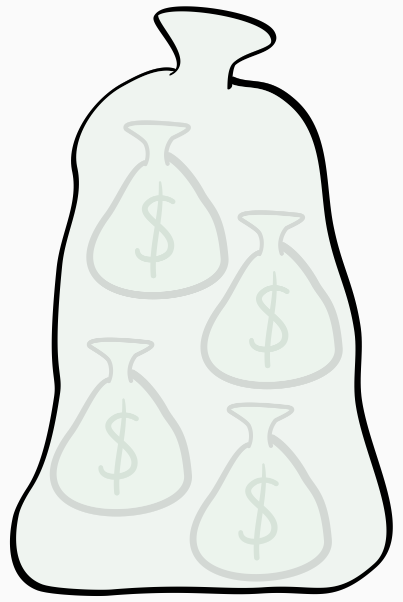
$V_1 = 10$ $V_2 = ?$ $V_3 = ?$ $V_4 = ?$ $V_5 = ?$ $V_6 = ?$ $V_7 = ?$ $V_8 = ?$



Coupling to **OFFLINE**

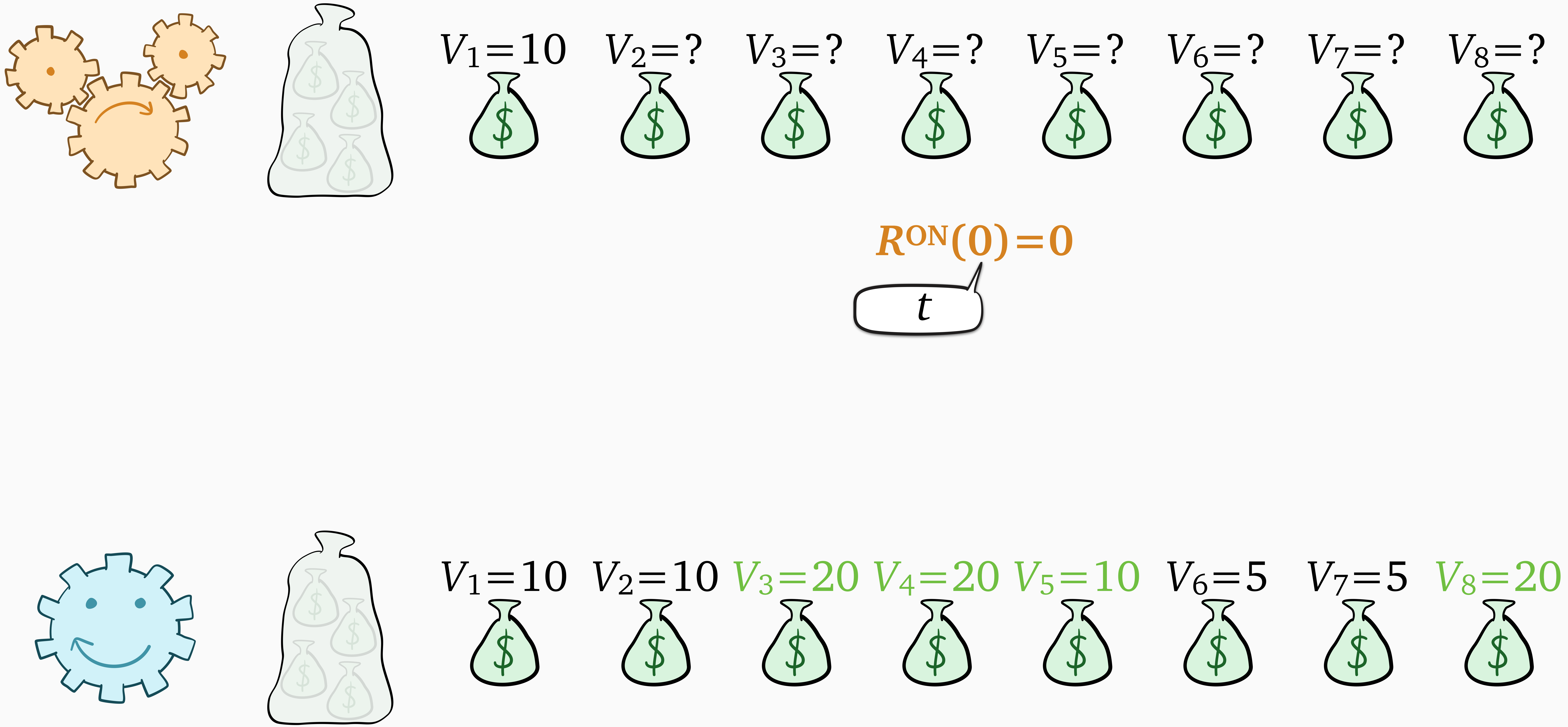



$V_1 = 10$ $V_2 = ?$ $V_3 = ?$ $V_4 = ?$ $V_5 = ?$ $V_6 = ?$ $V_7 = ?$ $V_8 = ?$

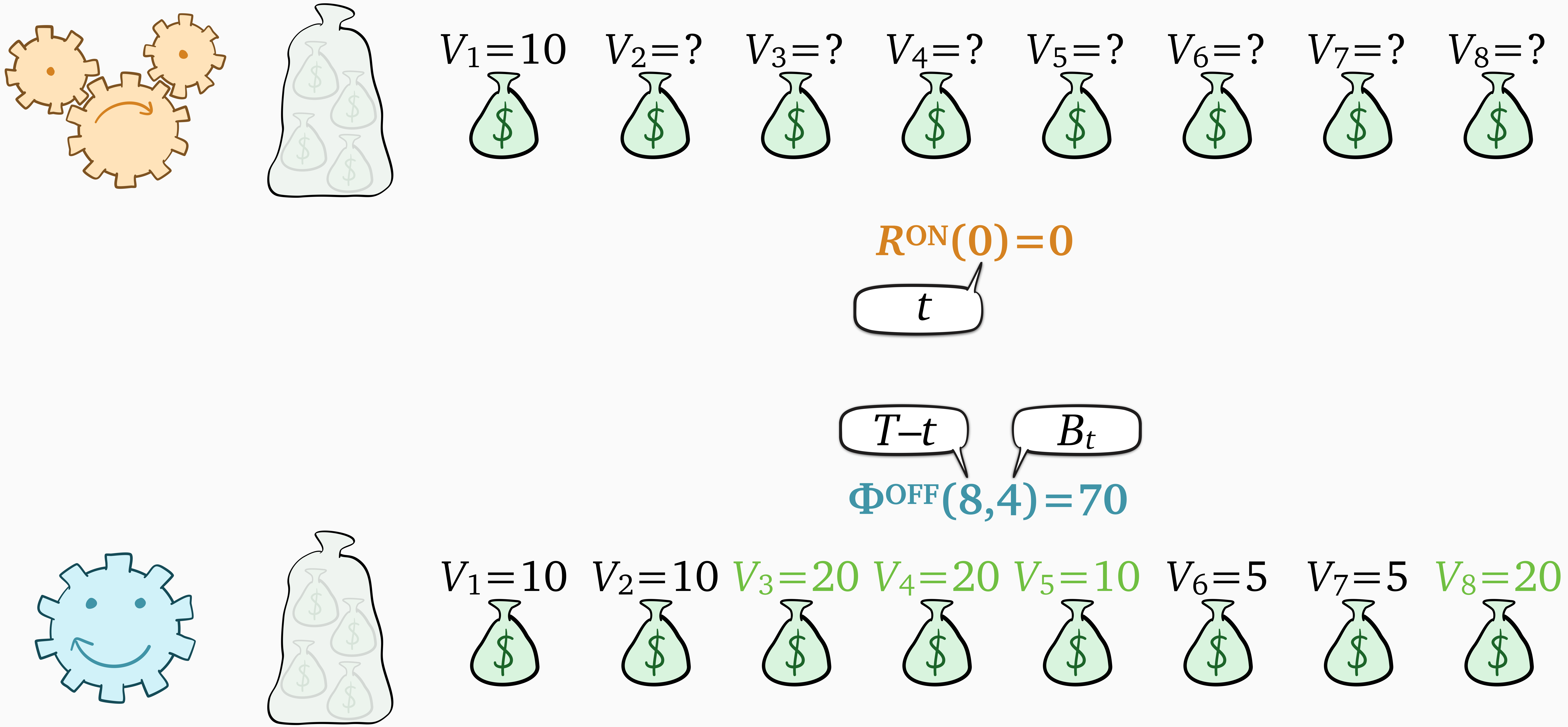



$V_1 = 10$ $V_2 = 10$ $V_3 = 20$ $V_4 = 20$ $V_5 = 10$ $V_6 = 5$ $V_7 = 5$ $V_8 = 20$

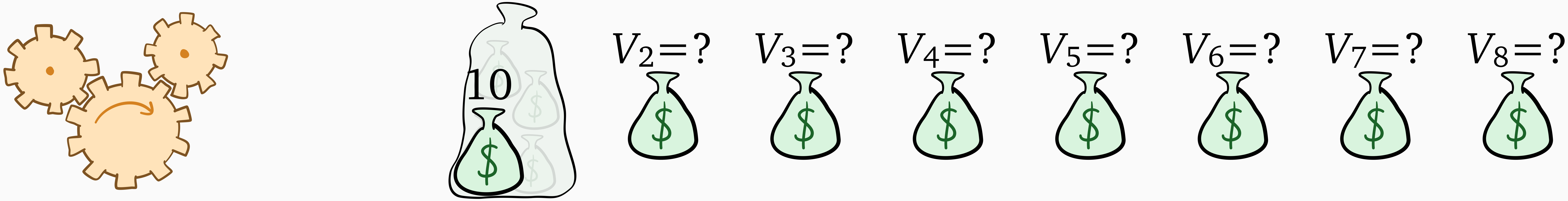
Coupling to **OFFLINE**



Coupling to OFFLINE

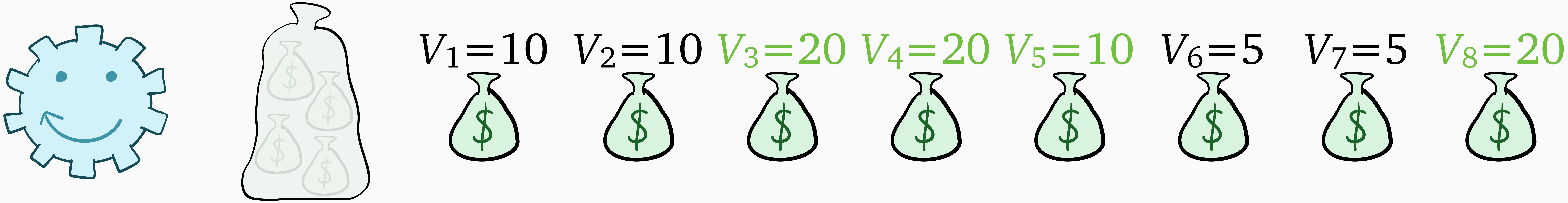


Coupling to OFFLINE

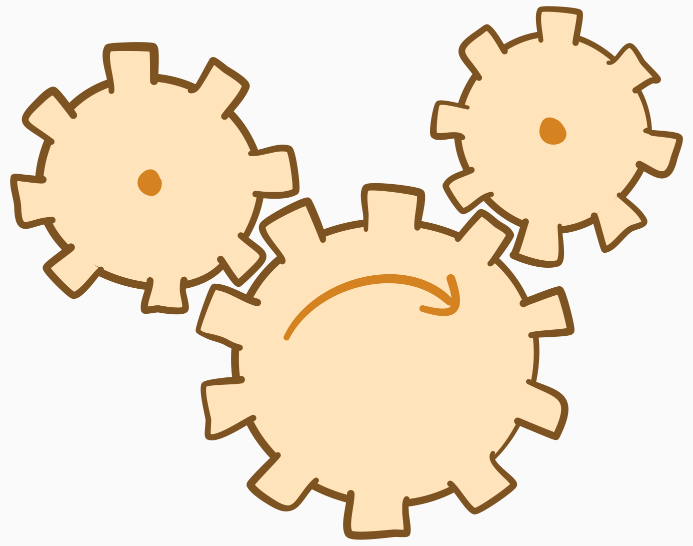


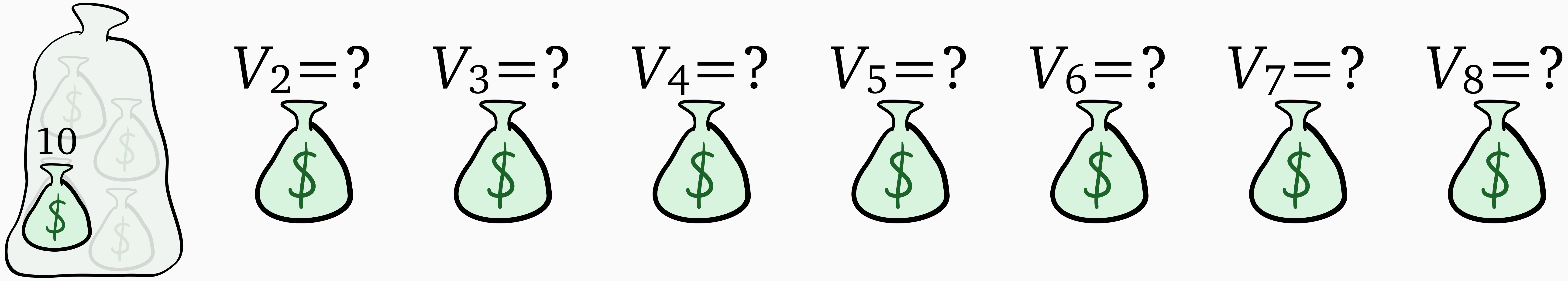
$R^{\mathrm{ON}}(1) = 10$

$\Phi^{\mathrm{OFF}}(8,4) = 70$



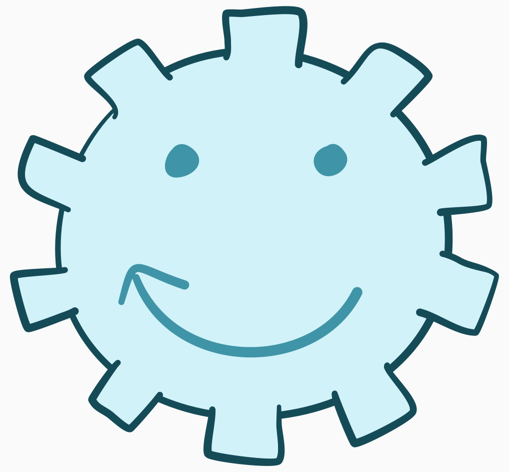
Coupling to **OFFLINE**

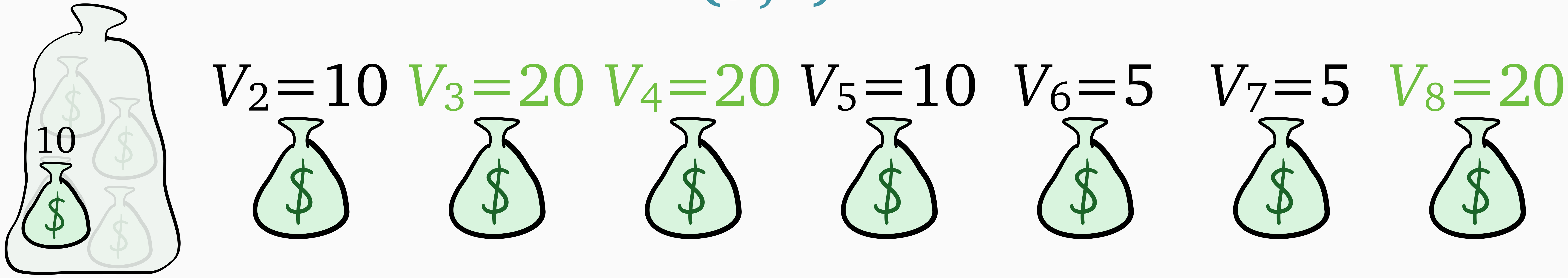




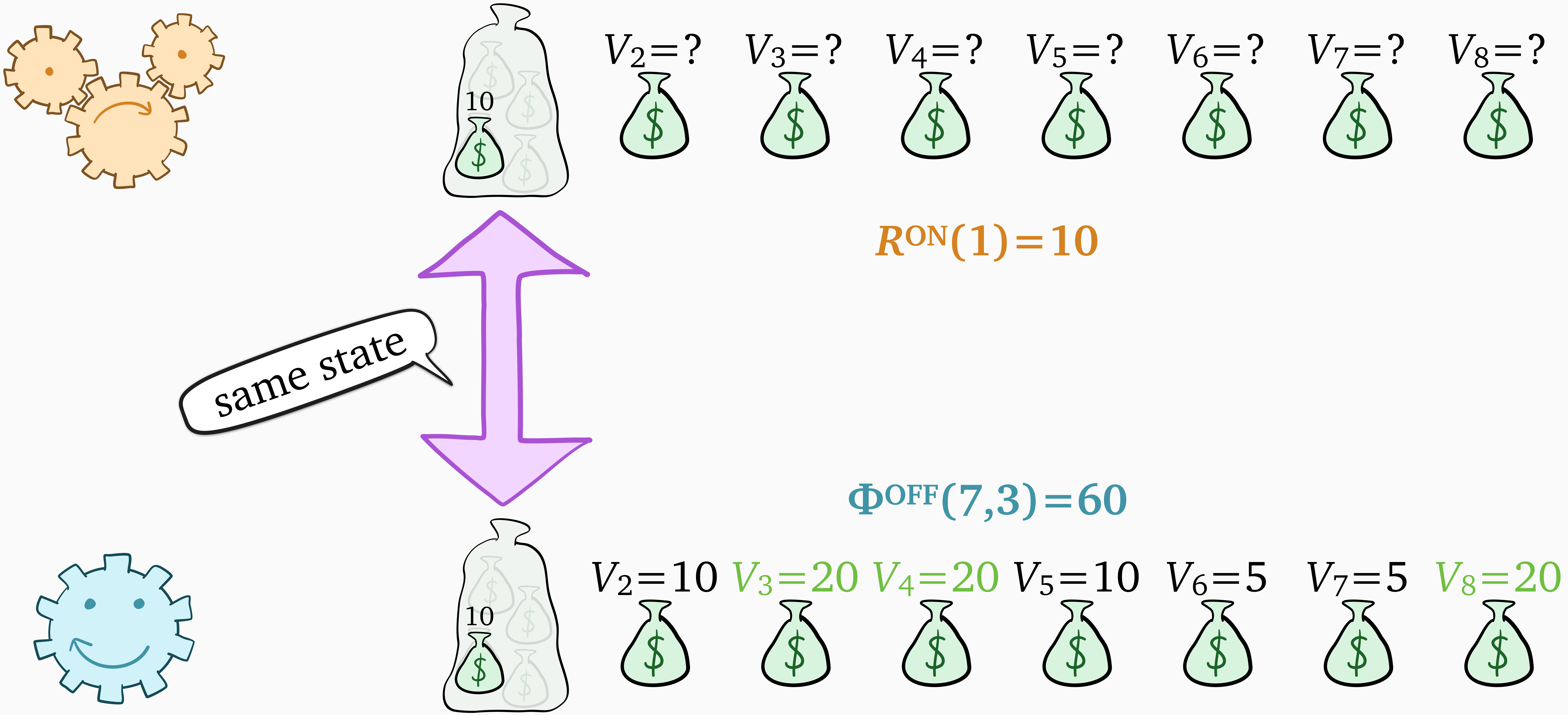
$R^{ON}(1) = 10$

$\Phi^{OFF}(7,3) = 60$

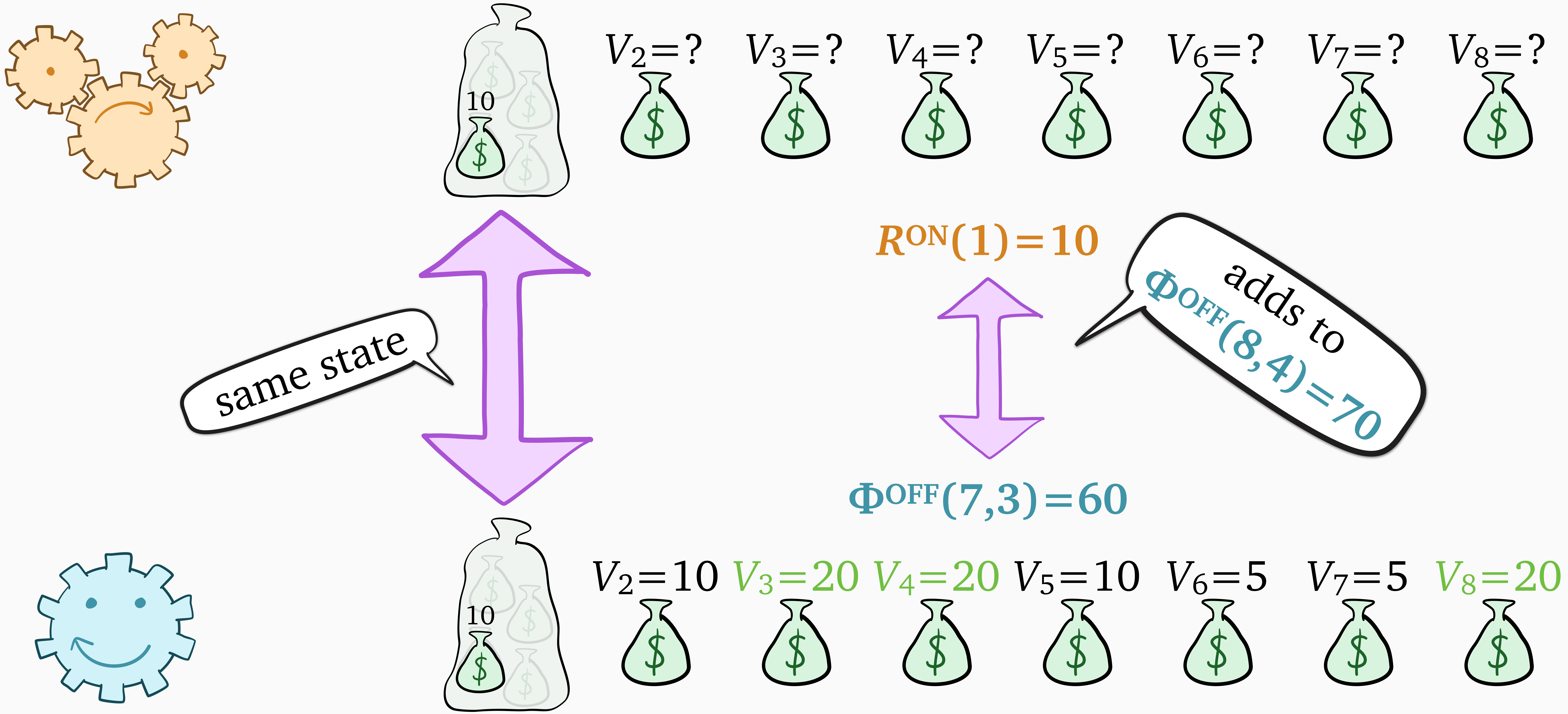




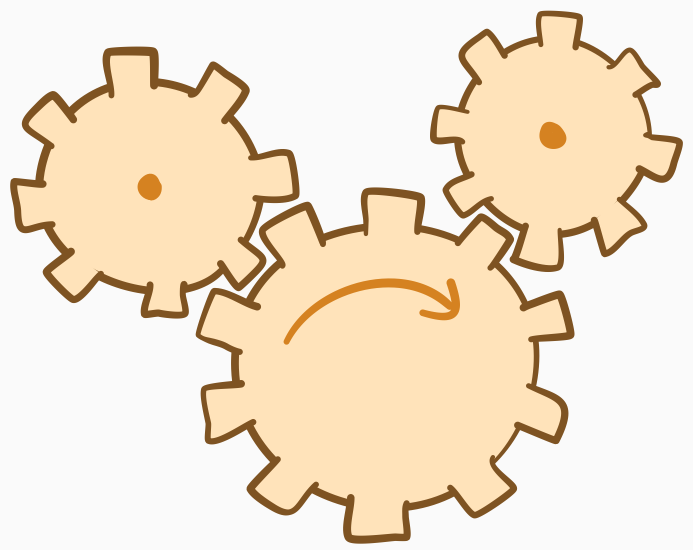
Coupling to **OFFLINE**

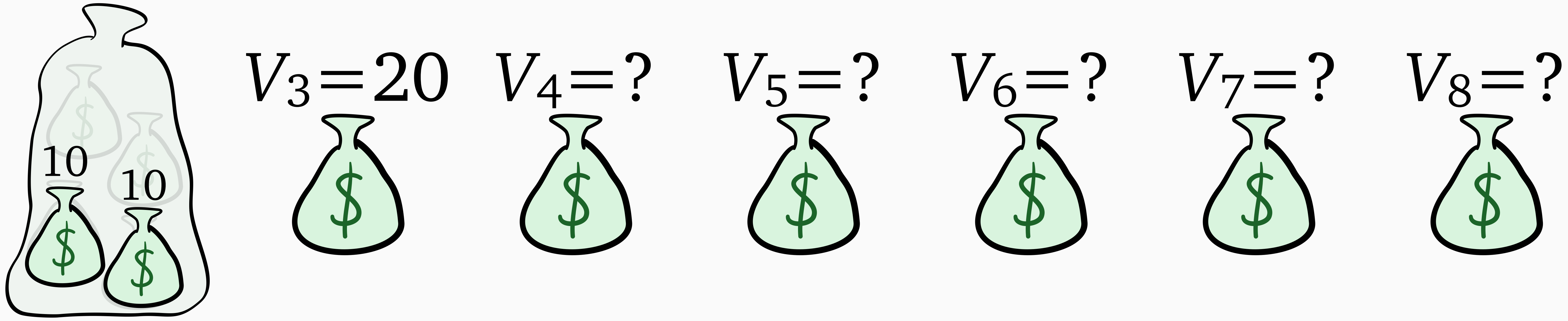


Coupling to **OFFLINE**



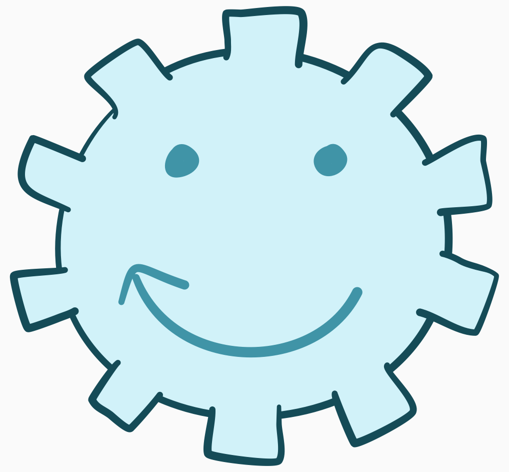
Coupling to OFFLINE with compensations

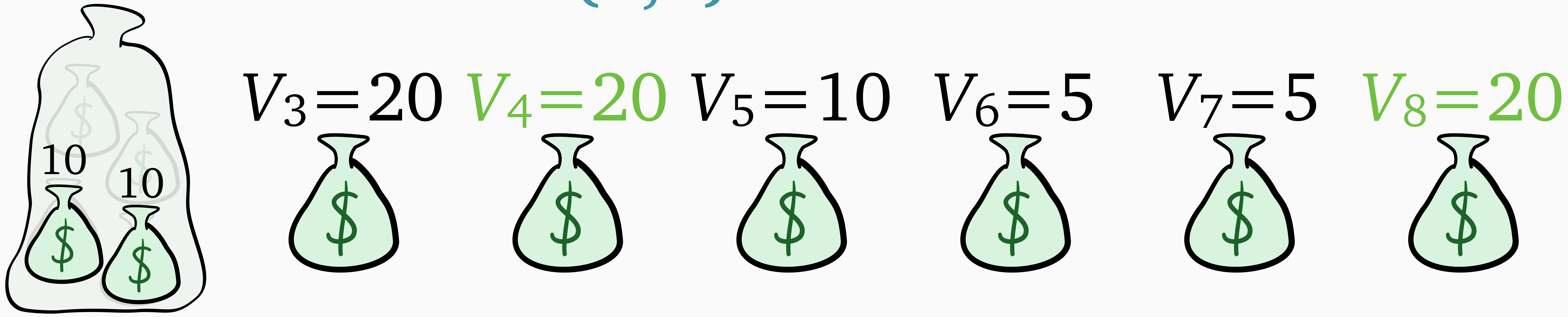




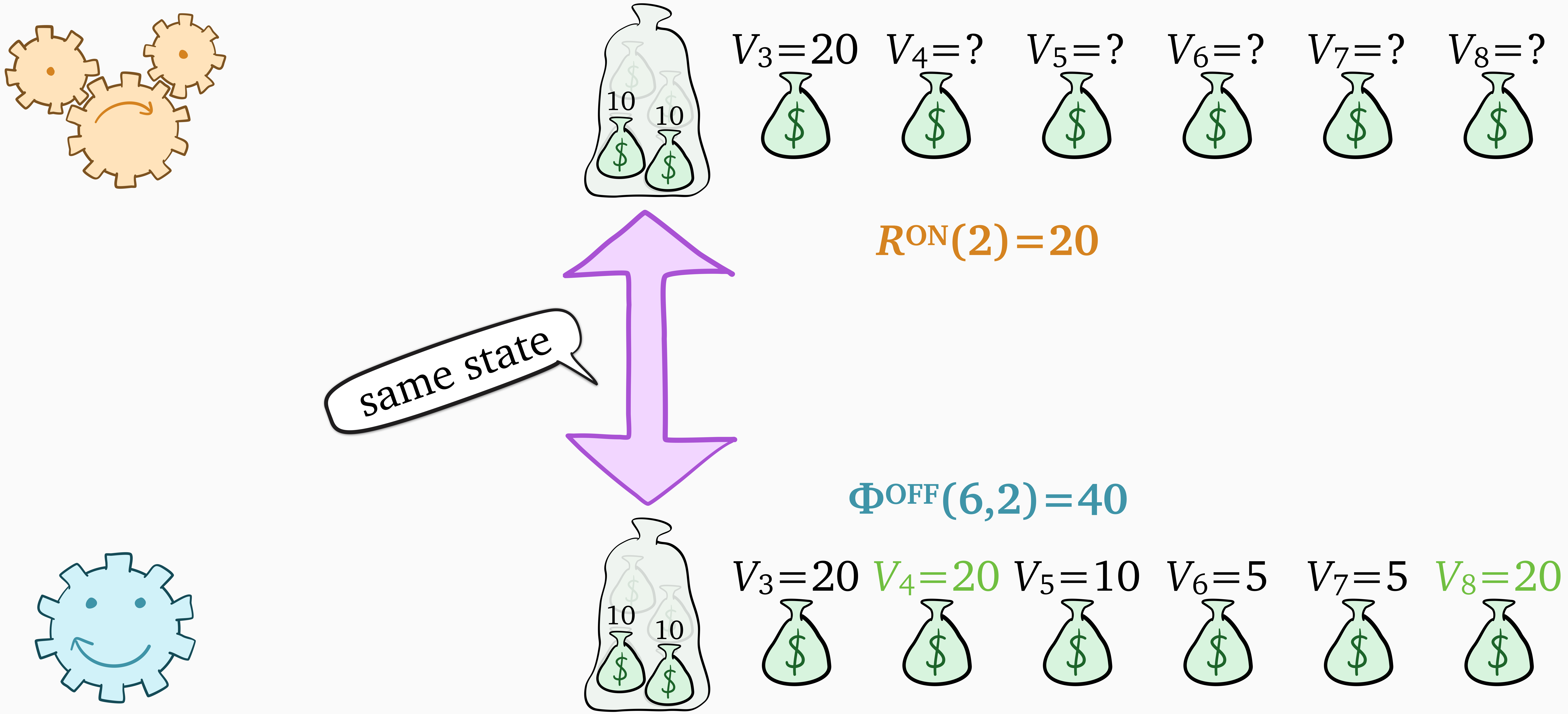
$R^{ON}(2) = 20$

$\Phi^{OFF}(6,2) = 40$

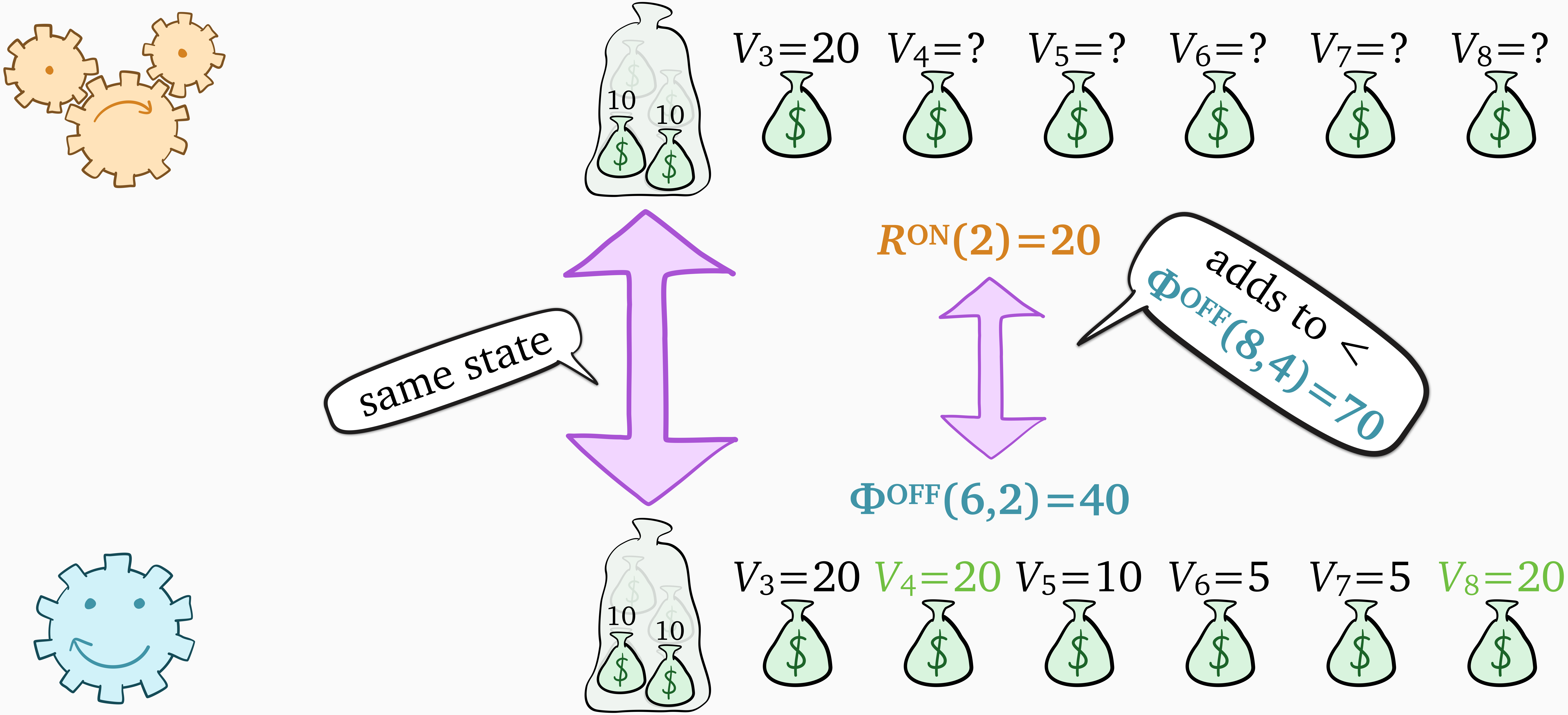




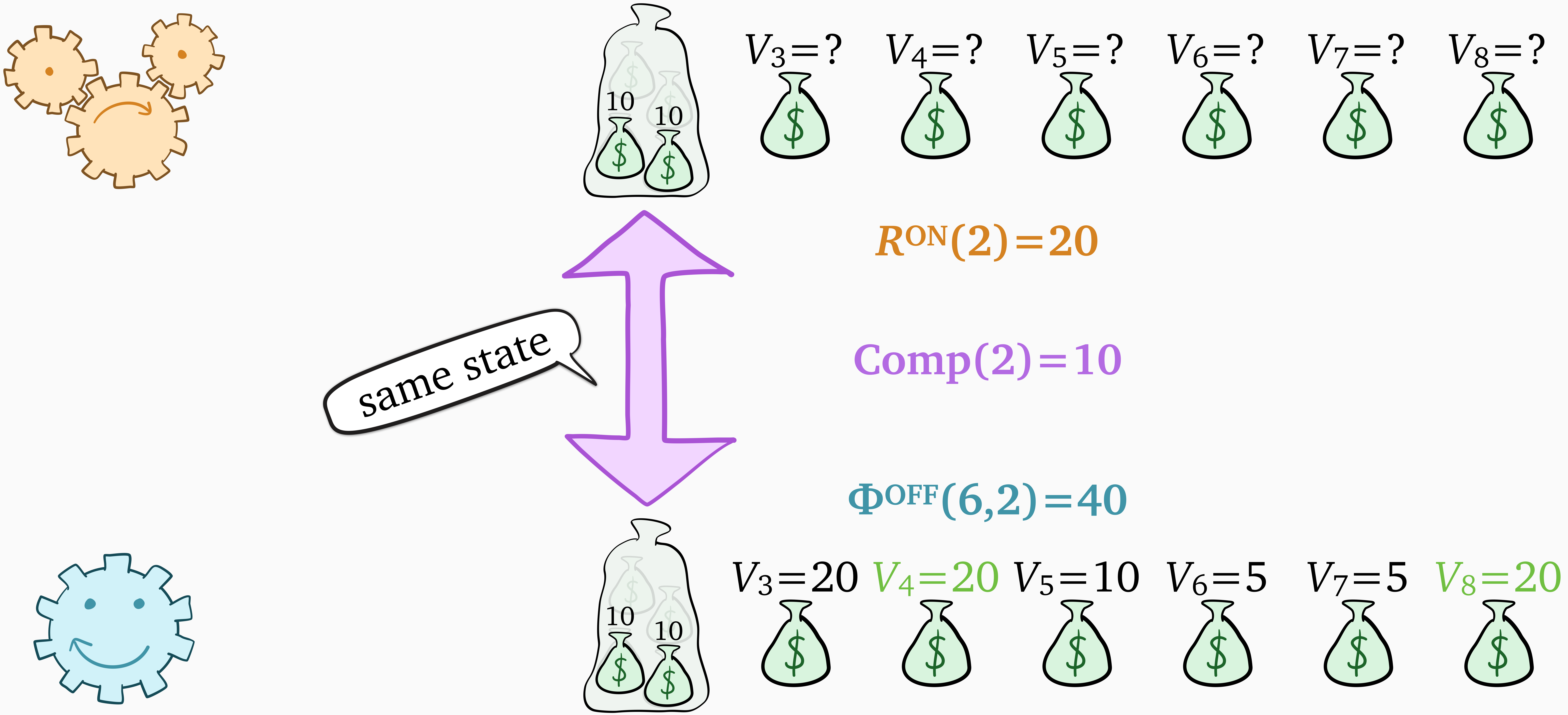
Coupling to OFFLINE with compensations



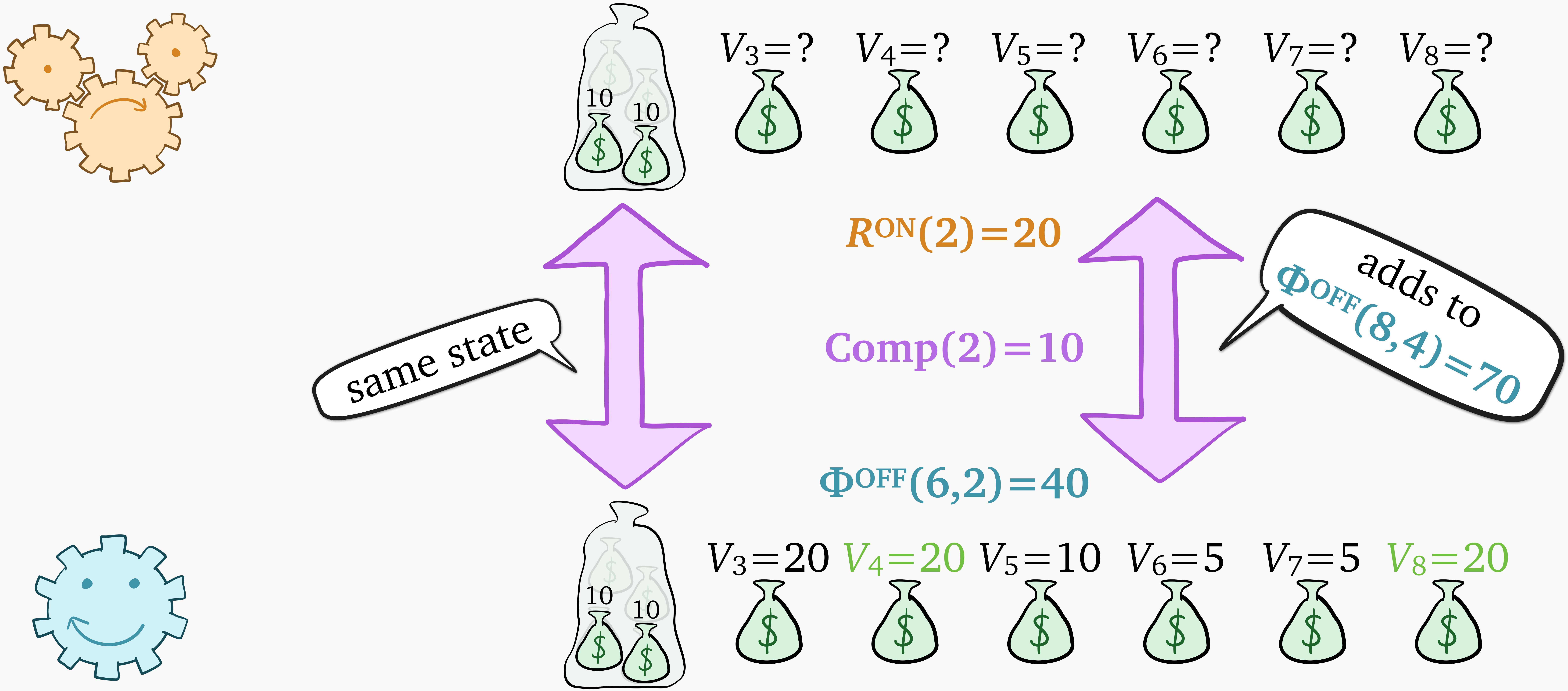
Coupling to OFFLINE with compensations



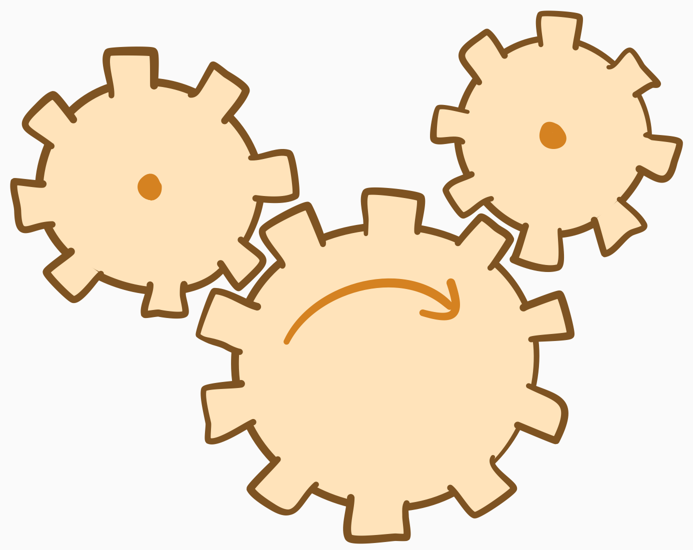
Coupling to OFFLINE with compensations

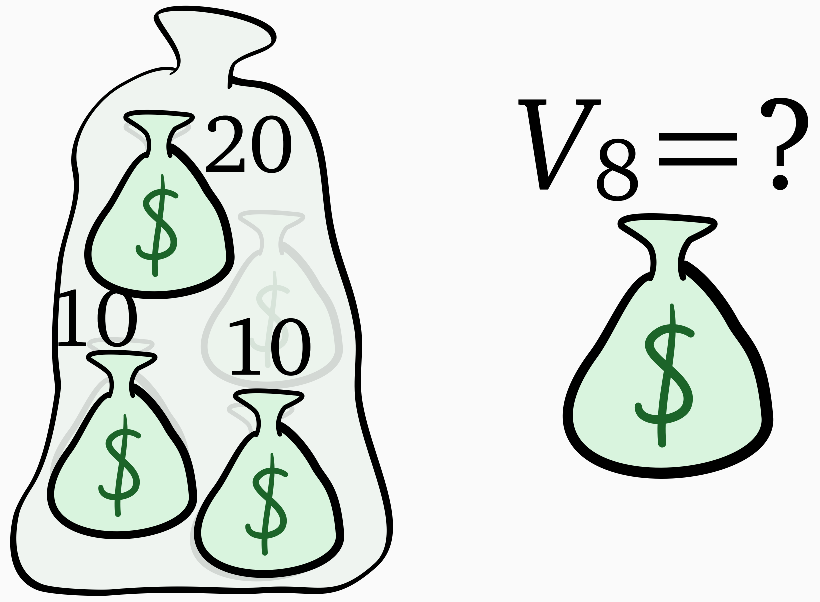


Coupling to OFFLINE with compensations



Coupling to OFFLINE with compensations



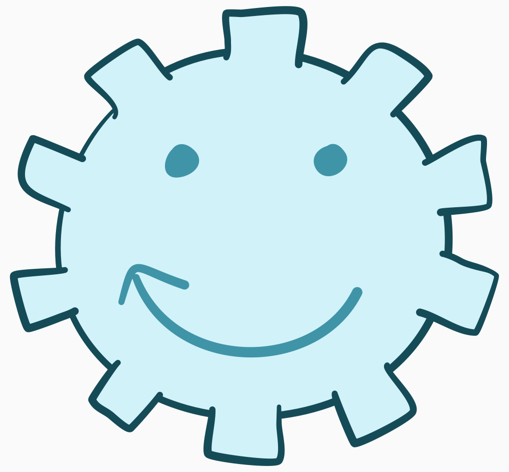


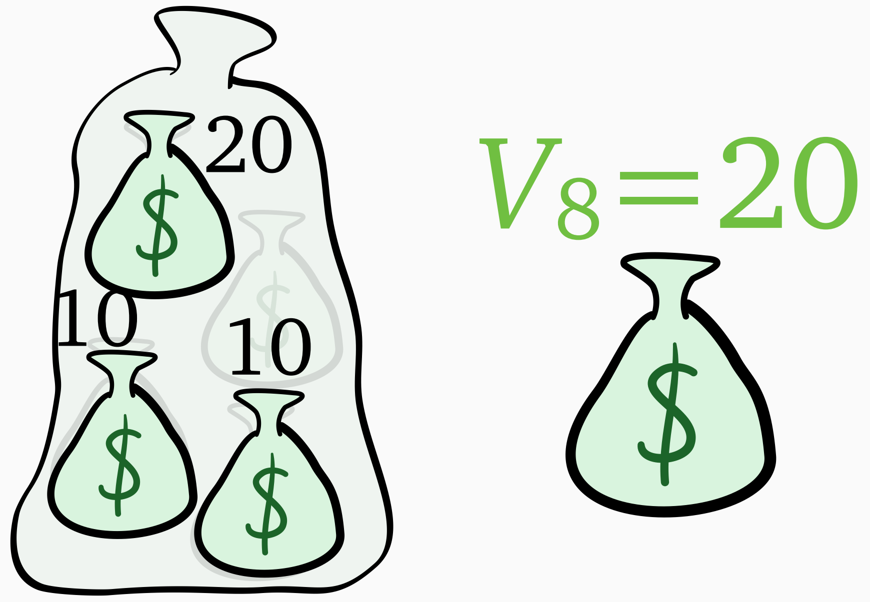
adds to
$\Phi^{\text{OFF}}(8,4) = 70$

$R^{\text{ON}}(7) = 40$

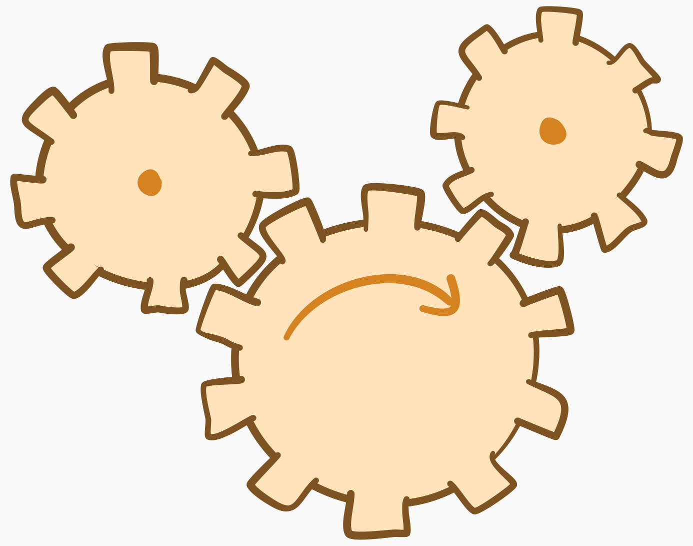
$\text{Comp}(1) + \ldots + \text{Comp}(7) = 10$

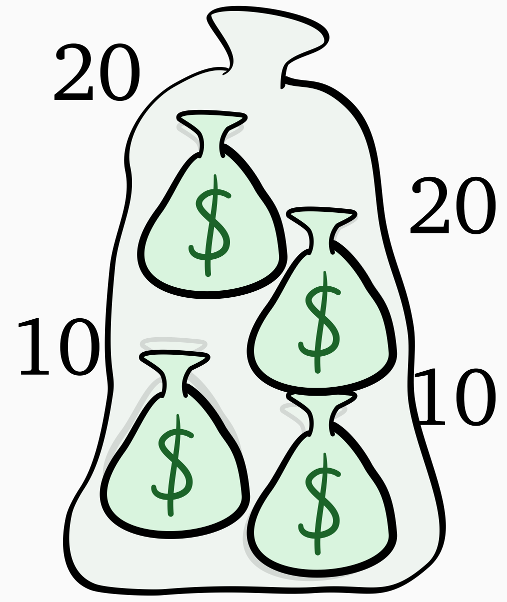
$\Phi^{\text{OFF}}(1,1) = 20$





Coupling to OFFLINE with compensations



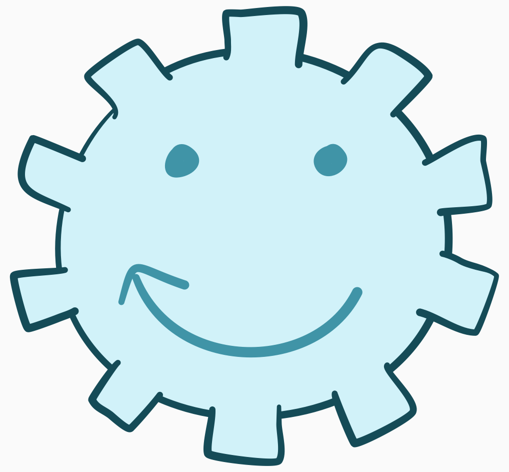


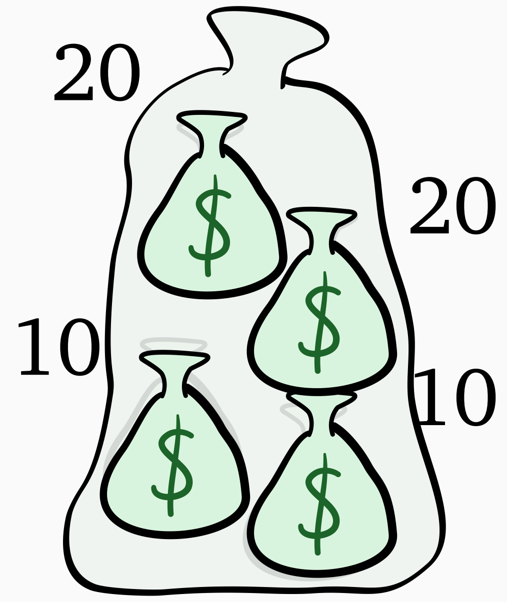
$R^{\mathrm{ON}}(8) = 60$

adds to
$\Phi^{\mathrm{OFF}}(8,4) = 70$

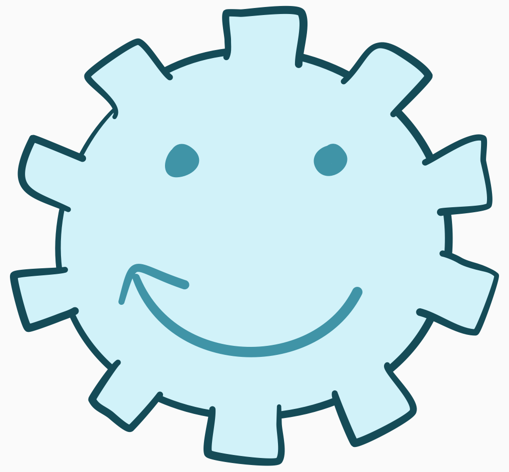
$\mathrm{Comp}(1) + \ldots + \mathrm{Comp}(8) = 10$

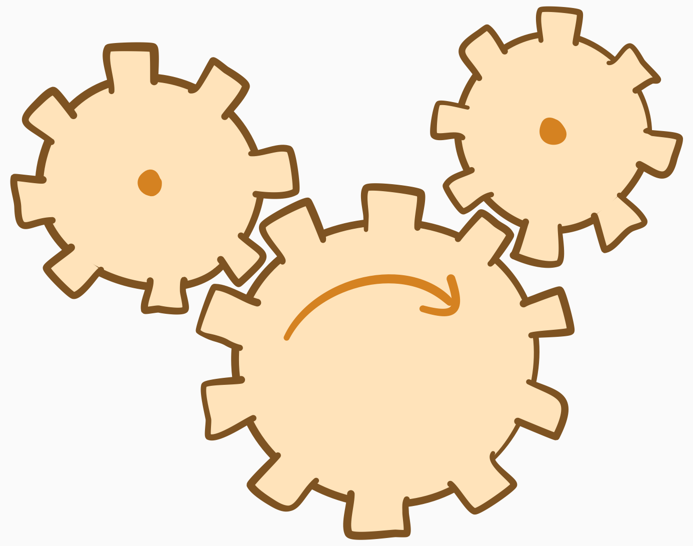
$\Phi^{\mathrm{OFF}}(0,0) = 0$





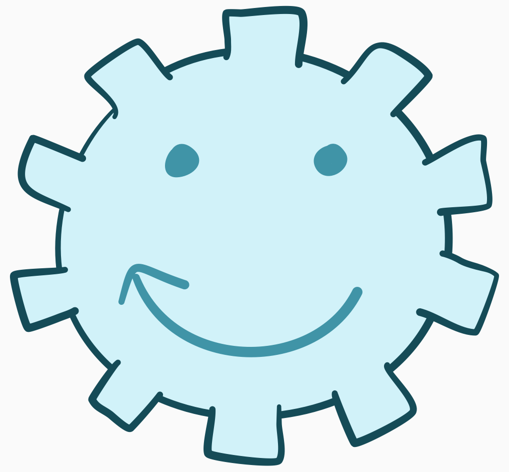
Compensated coupling

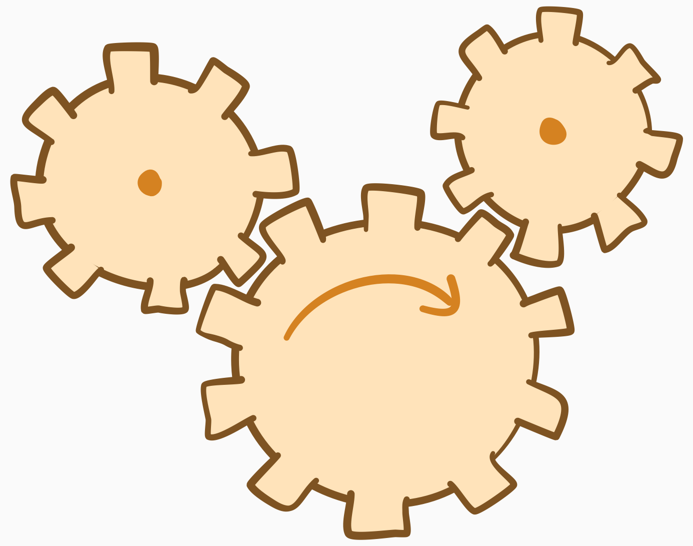




$$\Phi^{OFF}(8,4) - \Phi^{OFF}(0,0) = R^{ON}(8) + (\mathrm{Comp}(1)+\mathrm{Comp}(2)+\ldots+\mathrm{Comp}(8))$$

Compensated coupling



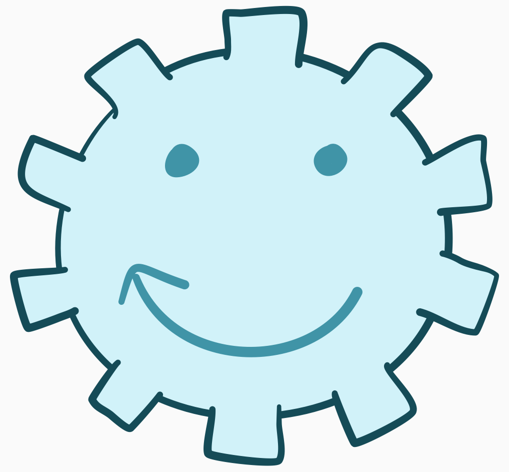


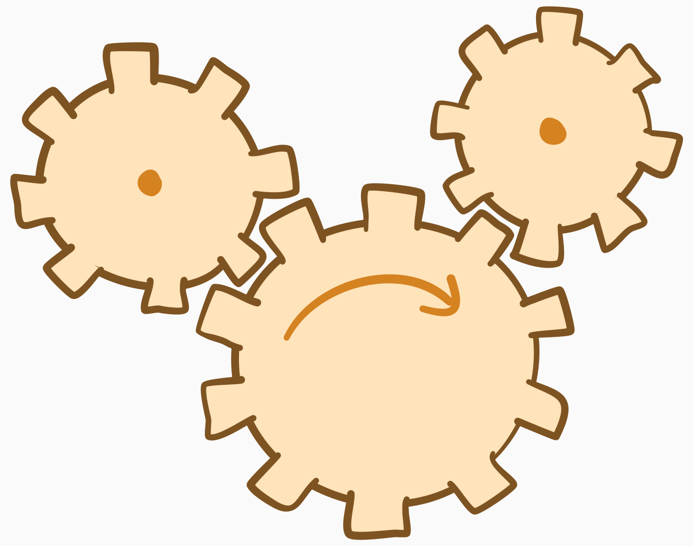
$$\Phi^{OFF}(8,4) - \Phi^{OFF}(0,0) = R^{ON}(8) + (\text{Comp}(1)+\text{Comp}(2)+\ldots+\text{Comp}(8))$$

ONLINE state

where $\text{Comp}(t) = \Phi^{OFF}(T-t, B_t) - R^{ON}(t)$

Compensated coupling





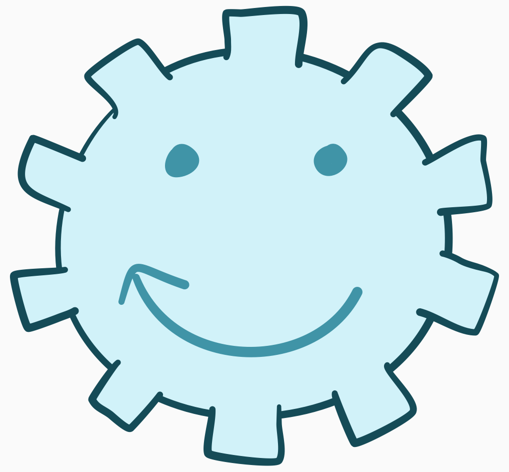
$$\Phi^{OFF}(8,4) - \Phi^{OFF}(0,0) = R^{ON}(8) + (\mathrm{Comp}(1)+\mathrm{Comp}(2)+\ldots+\mathrm{Comp}(8))$$

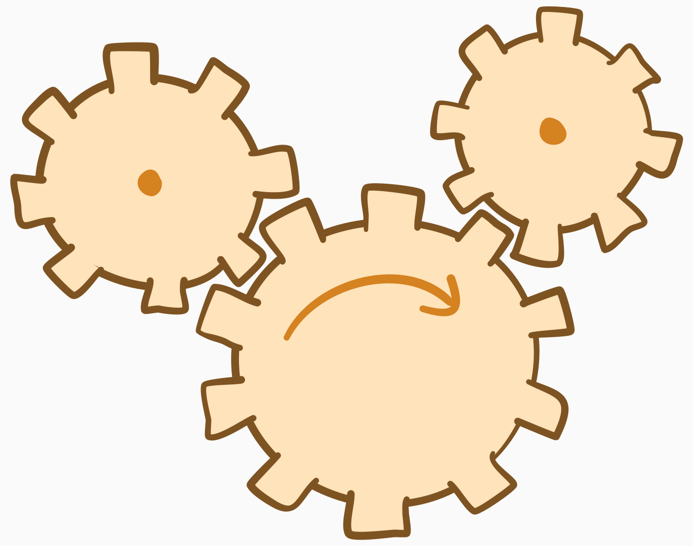
ONLINE state

where $\mathrm{Comp}(t) = \Phi^{OFF}(T-t, B_t) - R^{ON}(t)$

= "Compensation" provided to **OFFLINE** for following **ONLINE**

Compensated coupling





$$\Phi^{\mathrm{OFF}}(8,4) - \Phi^{\mathrm{OFF}}(0,0) = R^{\mathrm{ON}}(8) + (\mathrm{Comp}(1)+\mathrm{Comp}(2)+\ldots+\mathrm{Comp}(8))$$

ONLINE state

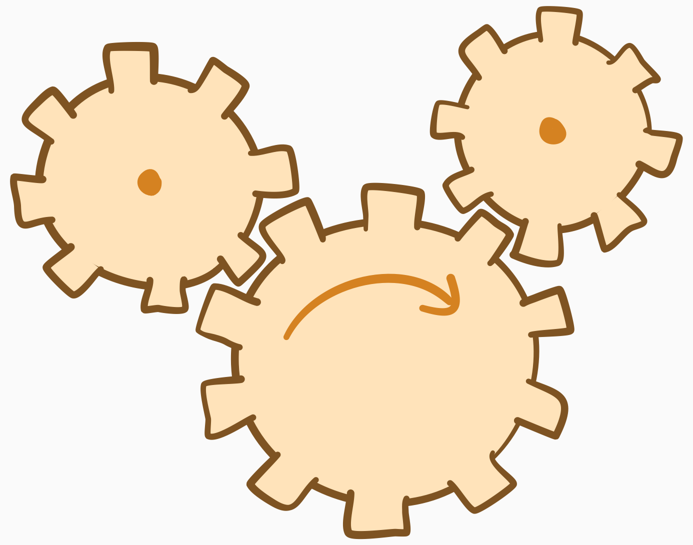
where $\mathrm{Comp}(t) = \Phi^{\mathrm{OFF}}(T-t, B_t) - R^{\mathrm{ON}}(t)$

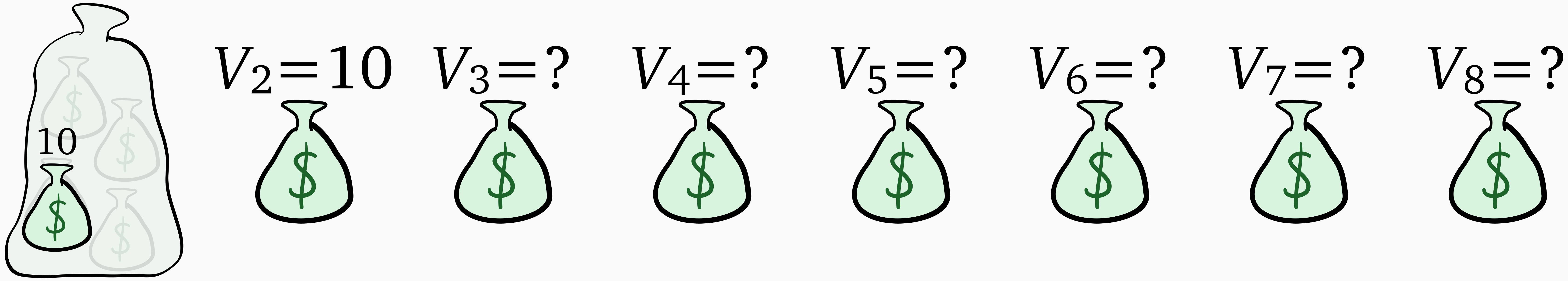
= "Compensation" provided to **OFFLINE** for following **ONLINE**

Notes: In computing $\mathrm{Comp}(t)$

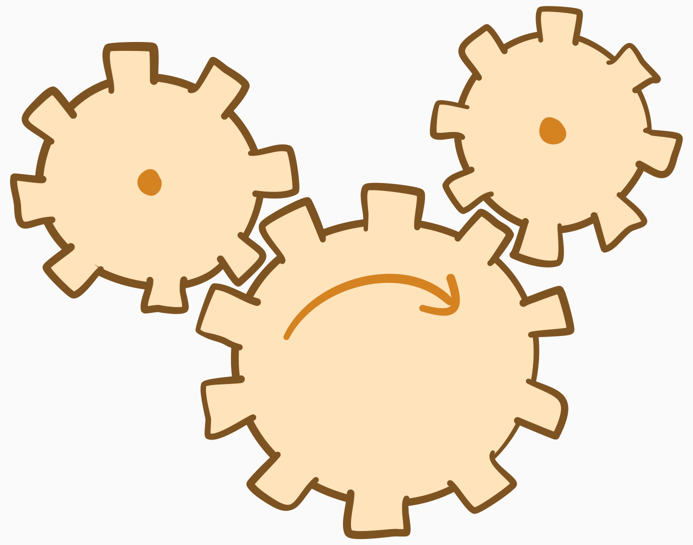
- Past arrivals/actions are forgotten (only **ONLINE**'s current state matters)
- Future arrivals are incorporated via **OFFLINE** ($\mathrm{Comp}(t)$ is a rand. var.)

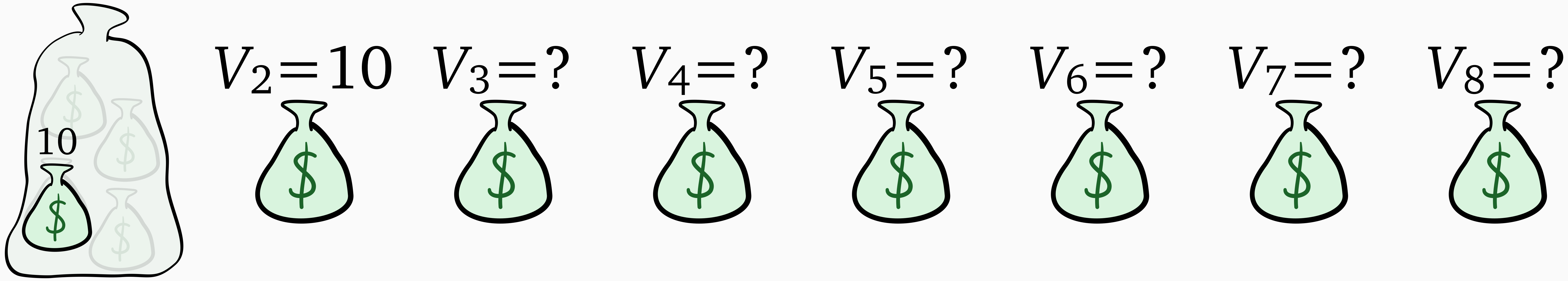
When do we compensate?





When do we compensate?

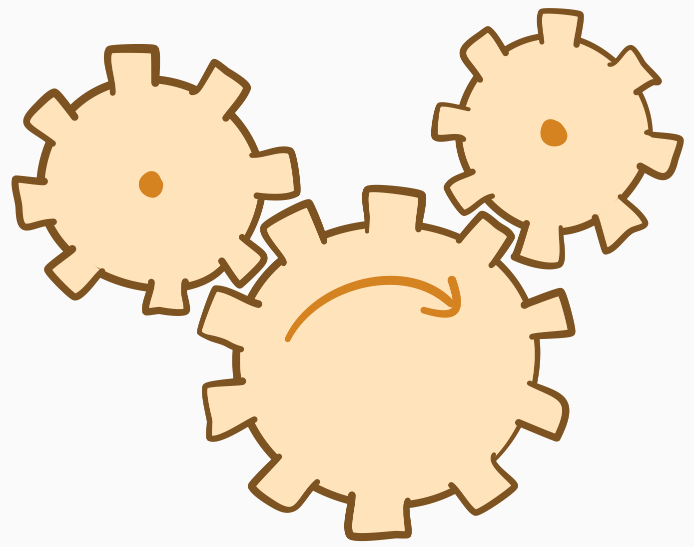


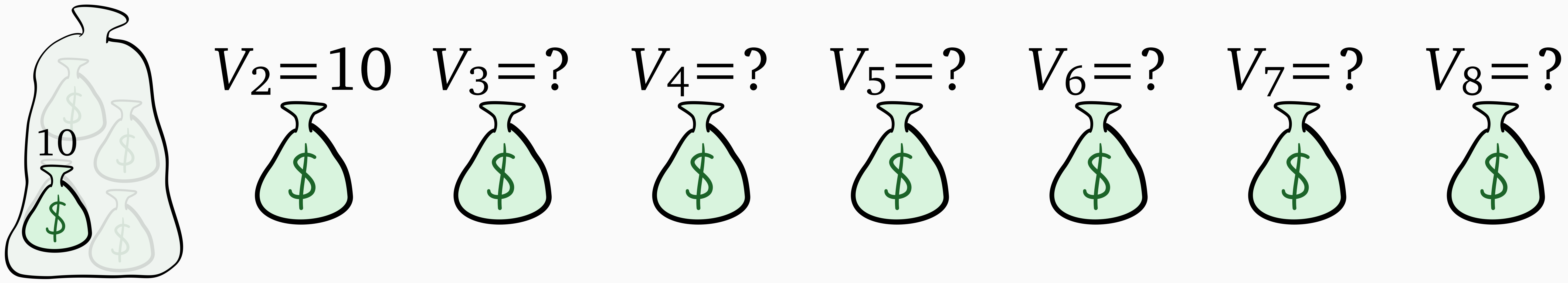


If **ONLINE** wants to accept V_2

- **Comp(2)** $= 10$ IFF at least 3 future arrivals have value 20
- **Comp(2)** $= 0$ otherwise

When do we compensate?





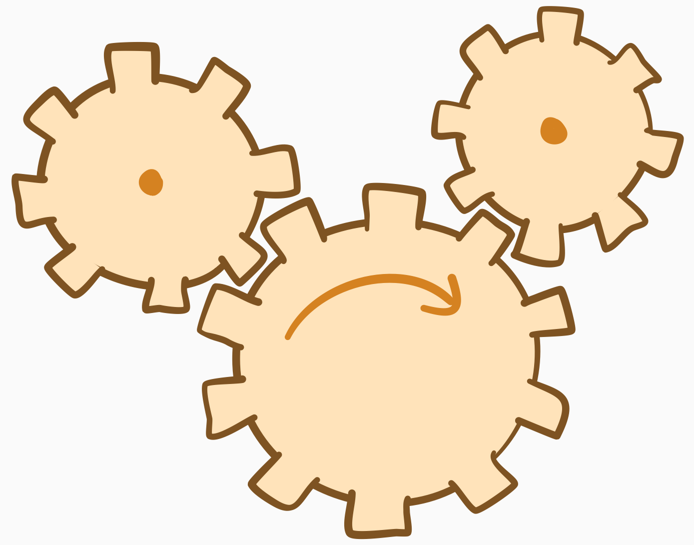
If **ONLINE** wants to accept V_2

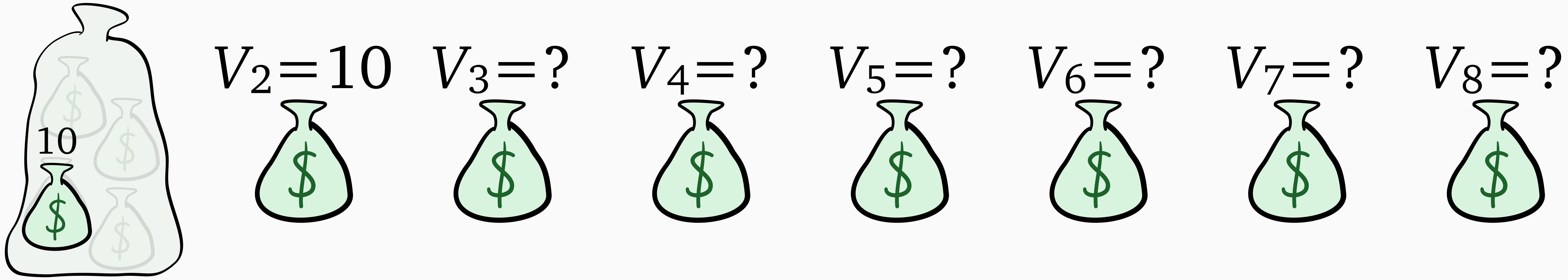
- **Comp(2)** $=$ 10 IFF at least 3 future arrivals have value 20
- **Comp(2)** $=$ 0 otherwise

If **ONLINE** wants to reject V_2

- **Comp(2)** $=$ 10 IFF at most 2 future arrivals have values in {10,20}
- **Comp(2)** $=$ 0 otherwise

When do we compensate?





If ONLINE wants to accept V_2

- Comp(2) $=$ 10 IFF at least 3 future arrivals have value 20
- Comp(2) $=$ 0 otherwise

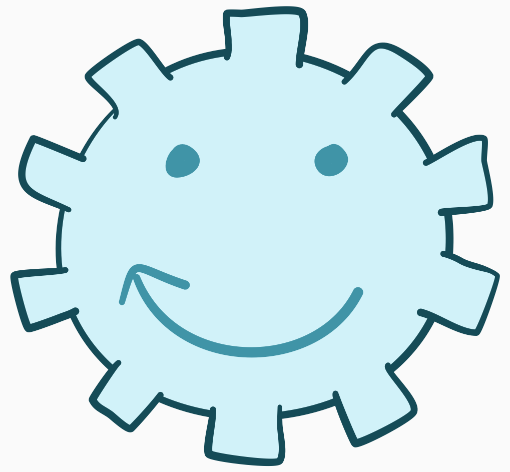
$\mathbf{P}[\mathrm{Bin}(6, p_{20}) \geq 3]$

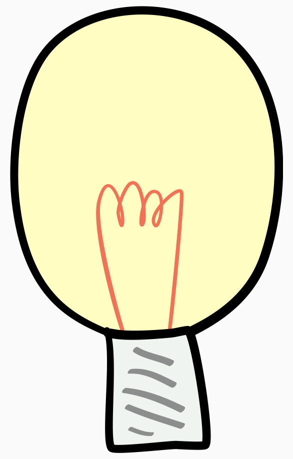
If ONLINE wants to reject V_2

- Comp(2) $=$ 10 IFF at most 2 future arrivals have values in {10,20}
- Comp(2) $=$ 0 otherwise

$\mathbf{P}[\mathrm{Bin}(6, p_{20} + p_{10}) \leq 2]$

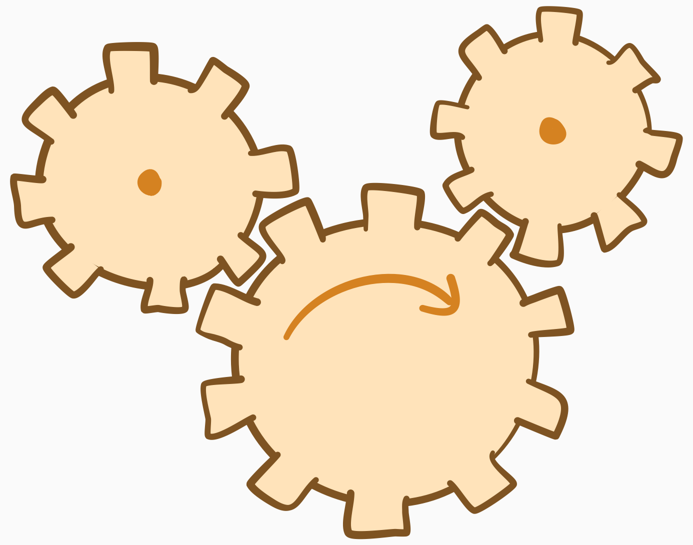
Compensated coupling: online knapsack





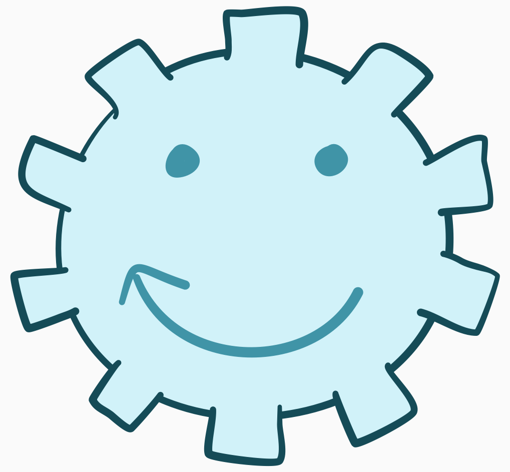
The Bayes selector:

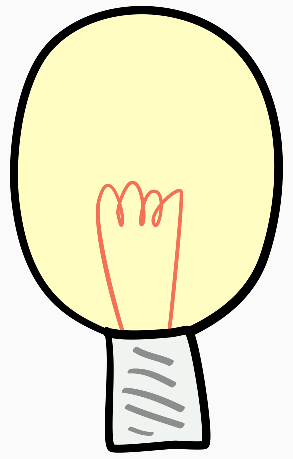
ONLINE chooses action with least compensation



Vera and Banerjee: The Bayesian Prophet (2020)

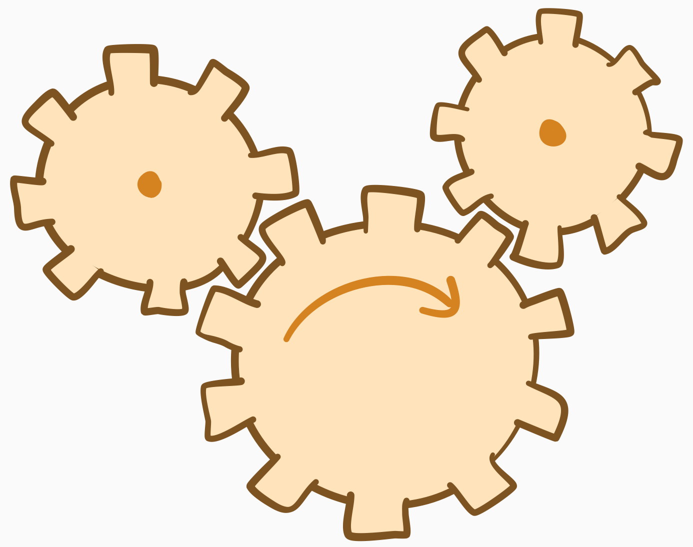
Compensated coupling: online knapsack





The Bayes selector:

ONLINE chooses action with least compensation

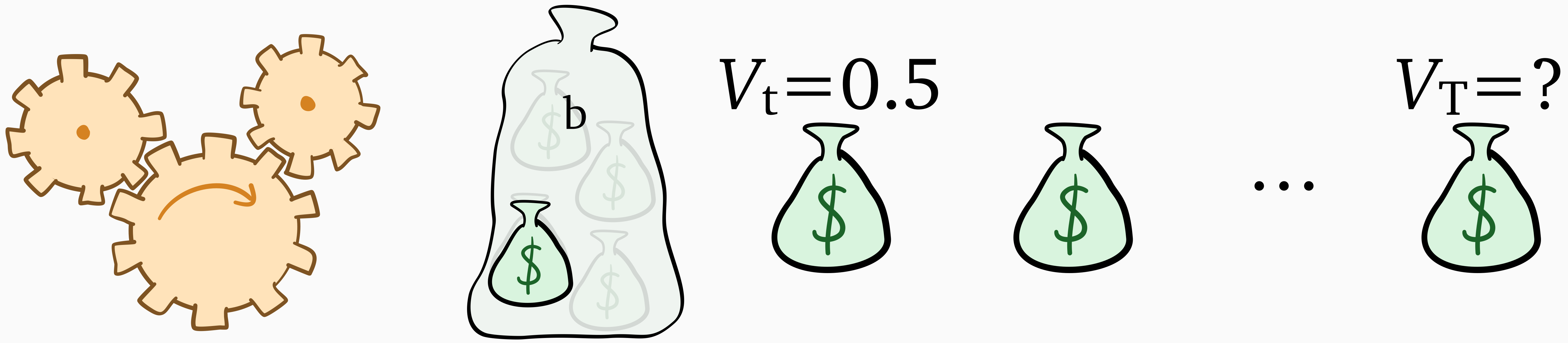


For online knapsack with k types, Bayes selector has

$$\mathbb{E}[\text{Regret}] \leq k v_{\max} / p_{\min}$$

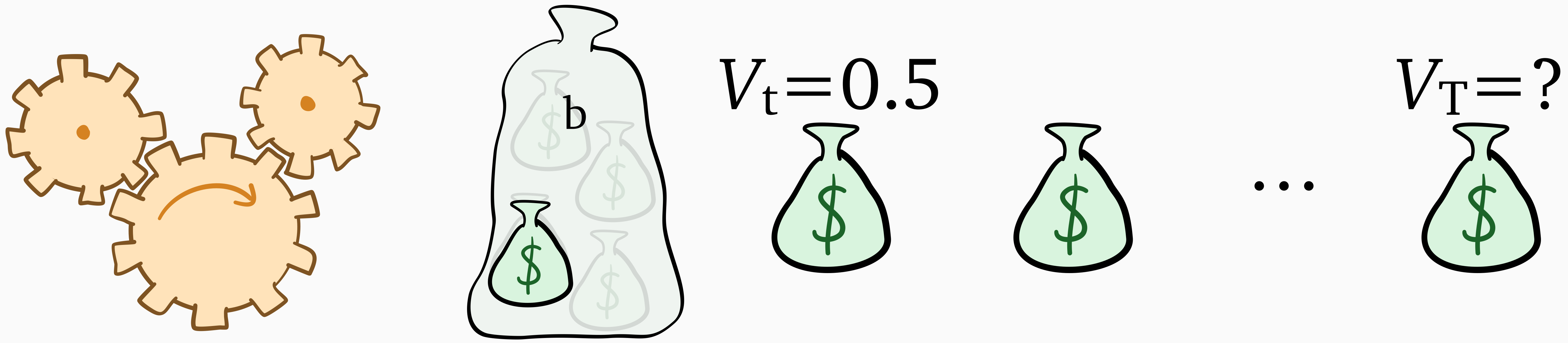
Vera and Banerjee: The Bayesian Prophet (2020)

Example: $V_t \sim U[0,1]$, i.i.d.

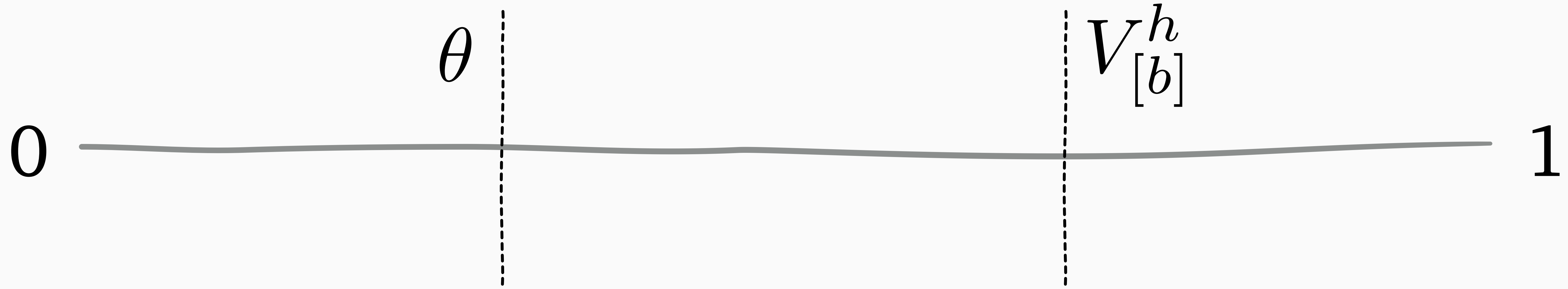


- **ONLINE** : Accept if $V_t \geq \theta$
- Let h=T-t, and define $V_{[b]}^{h} = \{b^{th} \text{ largest value in } V[t+1], V[t+2], \ldots, V[T]\}$

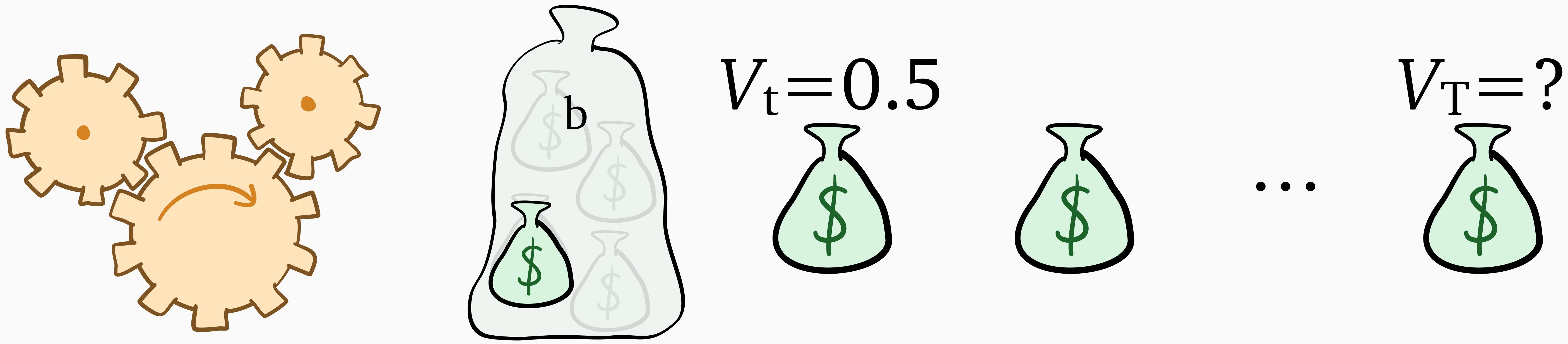
Example: $V_t \sim U[0,1]$, i.i.d.



- **ONLINE** : Accept if $V_t \geq \theta$
- Let h=T-t, and define $V_{[b]}^{h} = \{b^{th} \text{ largest value in } V[t+1], V[t+2], \ldots, V[T]\}$

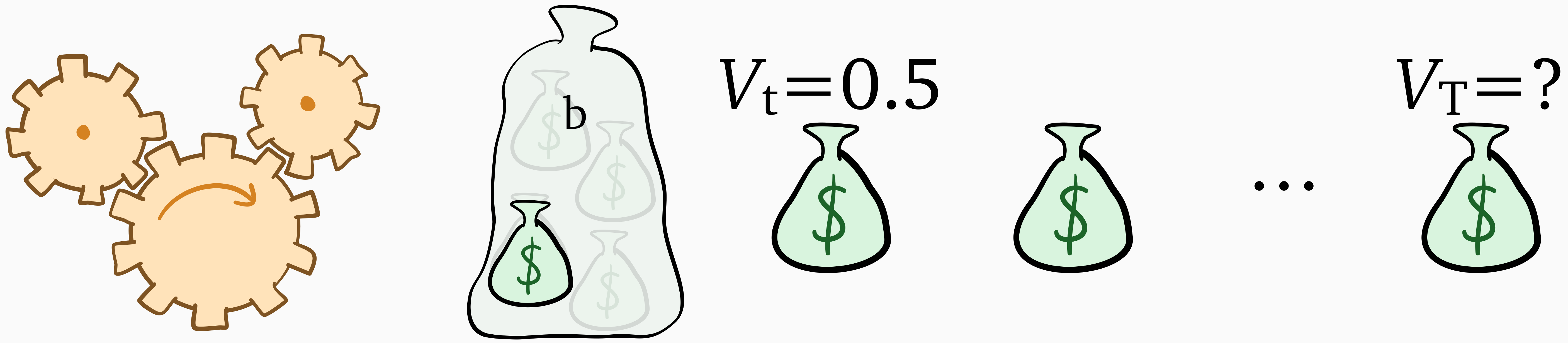


Example: $V_t \sim U[0,1]$, i.i.d.



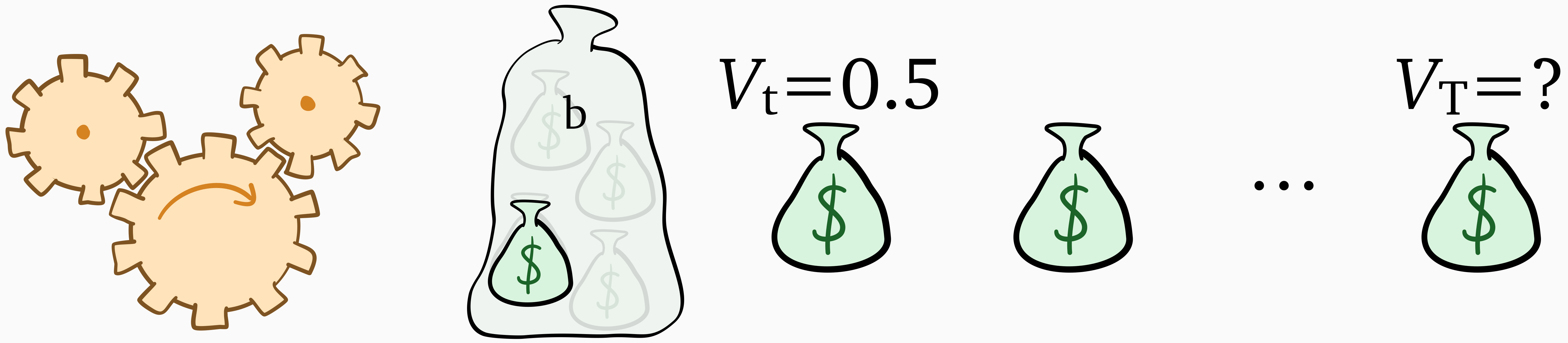
- **ONLINE** : Accept if $V_t \geq \theta$
- Let h=T-t, and define $V_{[b]}^h = \{b^{th} \text{ largest value in } V[t+1], V[t+2], \ldots, V[T]\}$

Example: $V_t \sim U[0,1]$, i.i.d.

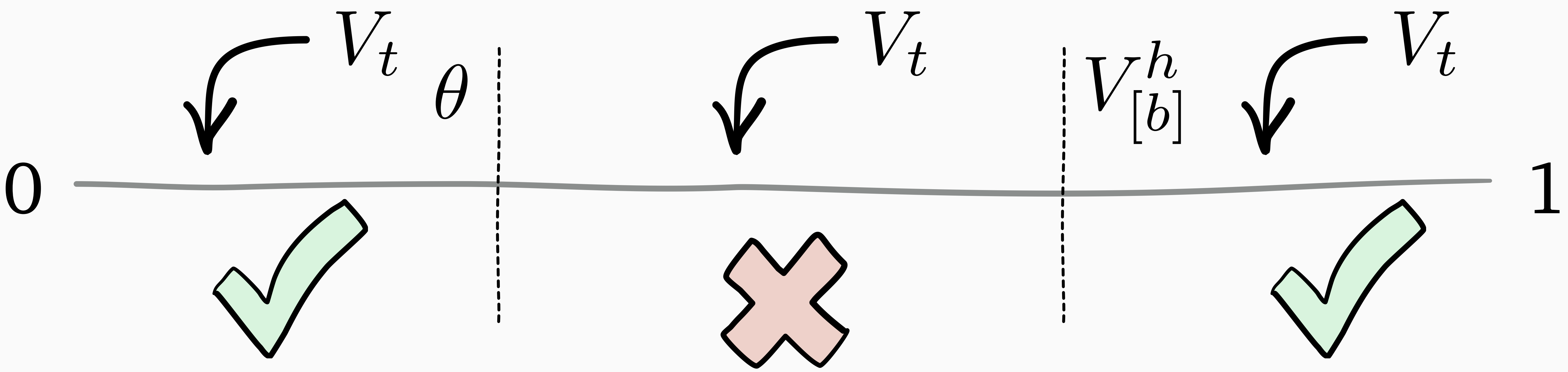


- **ONLINE** : Accept if $V_t \geq \theta$
- Let h=T-t, and define $V^h_{[b]} = \{b^{th}$ largest value in $V[t+1], V[t+2], \ldots, V[T]\}$

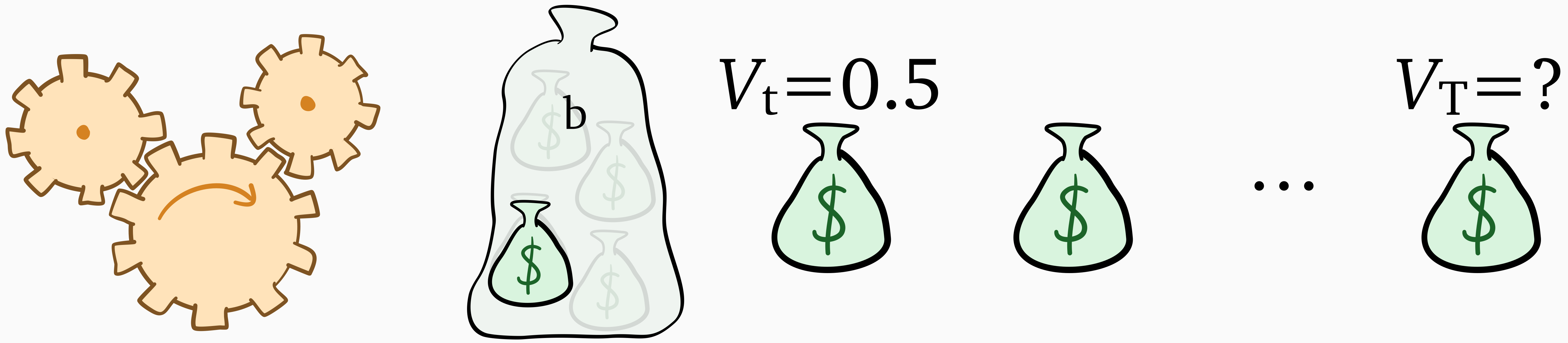
Example: $V_t \sim U[0,1]$, i.i.d.



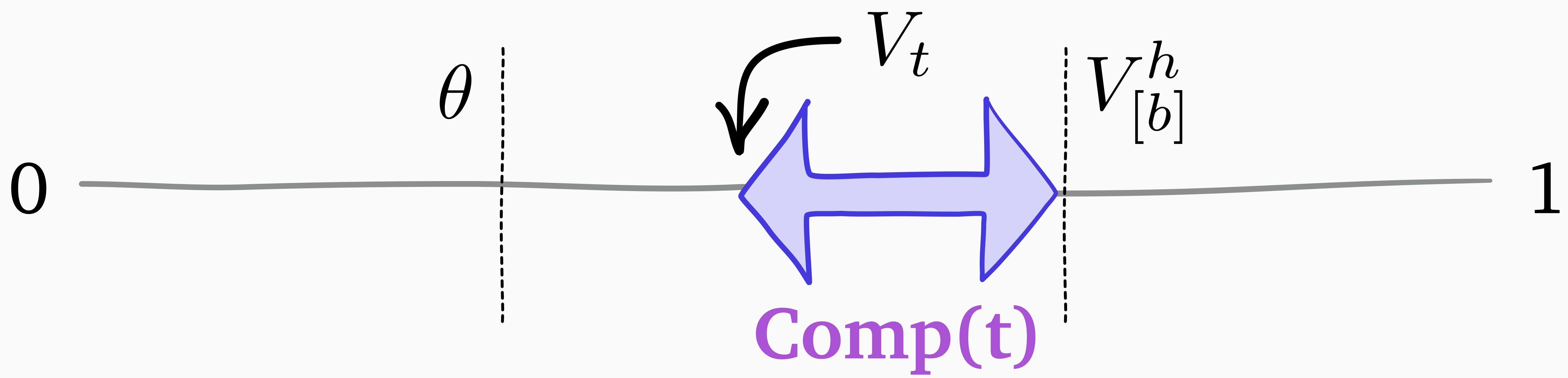
- **ONLINE** : Accept if $V_t \geq \theta$
- Let h=T-t, and define $V^h_{[b]} = \{b^{th} \text{ largest value in } V[t+1], V[t+2], \ldots, V[T]\}$



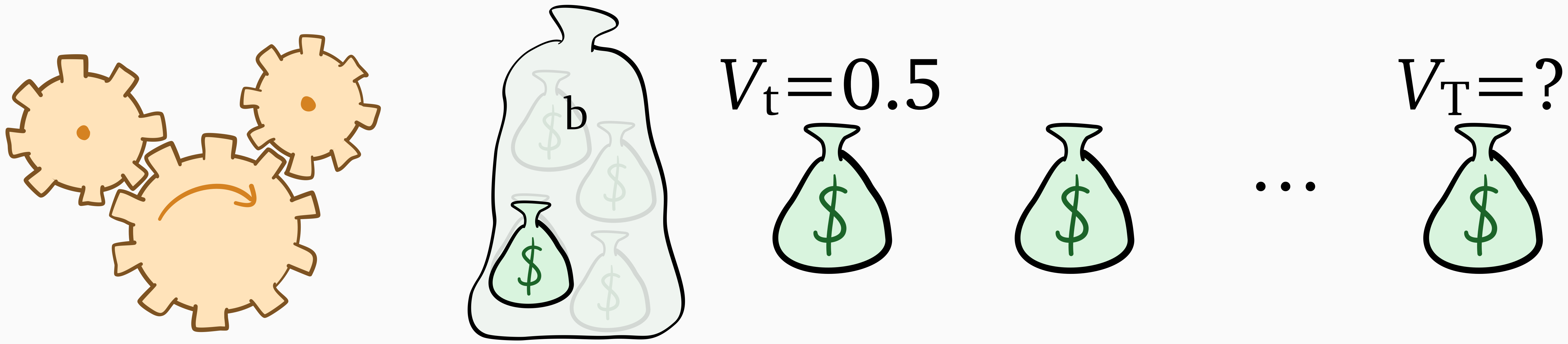
Example: $V_t \sim U[0,1]$, i.i.d.



- **ONLINE** : Accept if $V_t \geq \theta$
- Let h=T-t, and define $V^h_{[b]} = \{b^{th} \text{ largest value in } V[t+1], V[t+2], \ldots, V[T]\}$

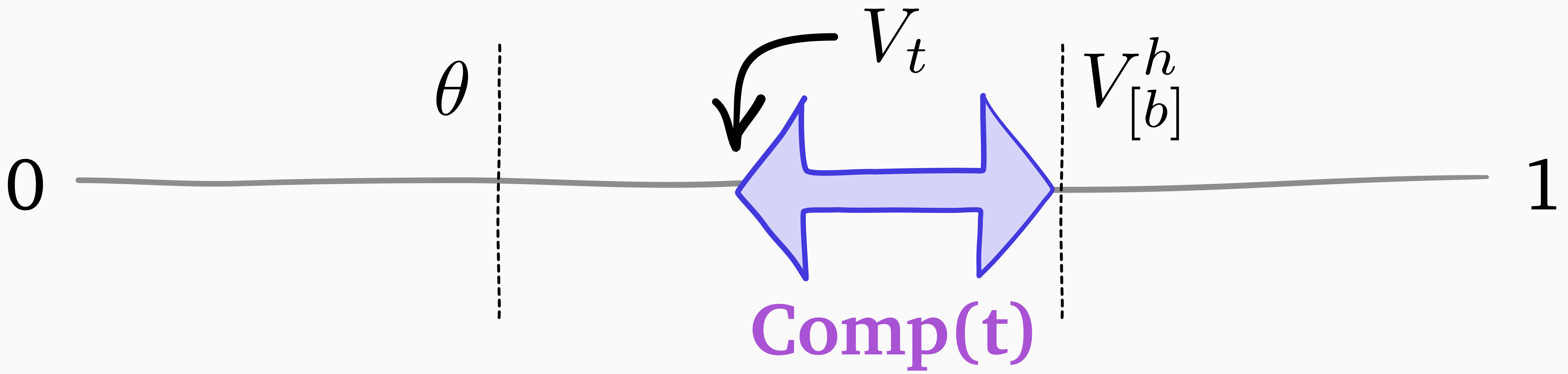


Example: $V_t \sim U[0, 1]$, i.i.d.

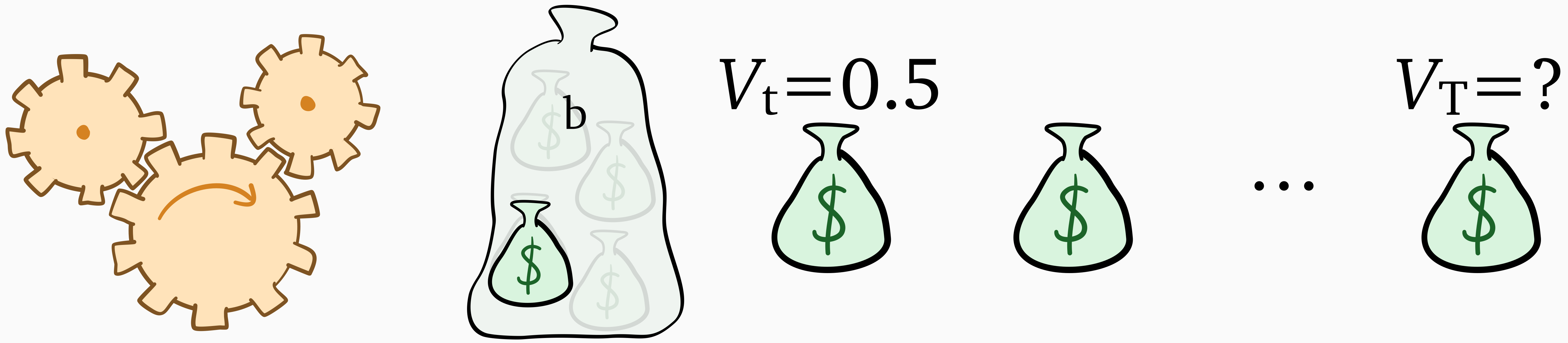


- **ONLINE** : Accept if $V_t \geq \theta$
- Let h=T-t, and define $V_{[b]}^h = \{b^{th} \text{ largest value in } V[t+1], V[t+2], \ldots, V[T]\}$

$$\text{Comp}(t) = |V_{[b]}^h - V_t|| \mathbb{1}_{V_t \in [\theta, V_{[b]}^h]}$$



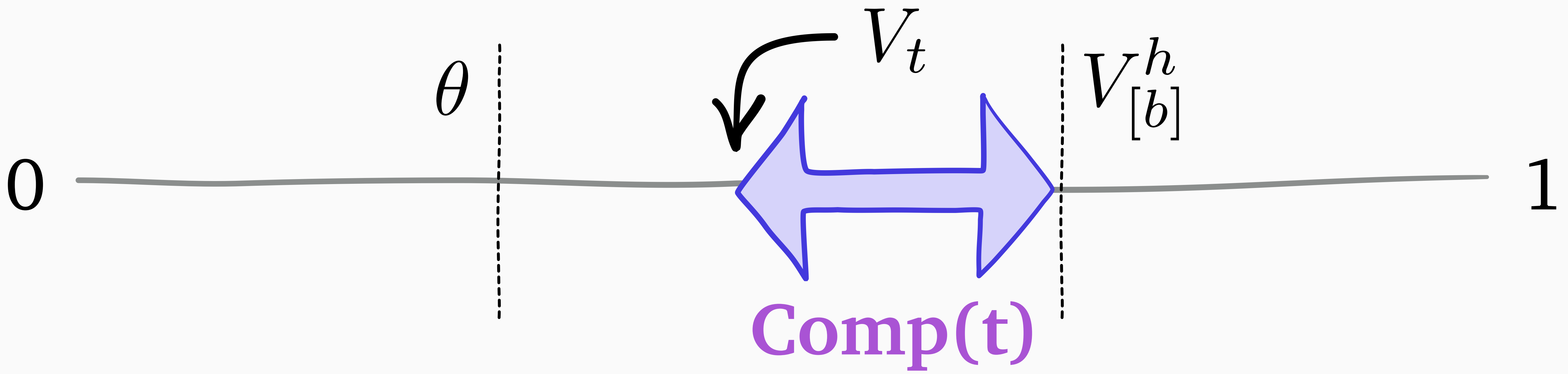
Example: $V_t \sim U[0,1]$, i.i.d.



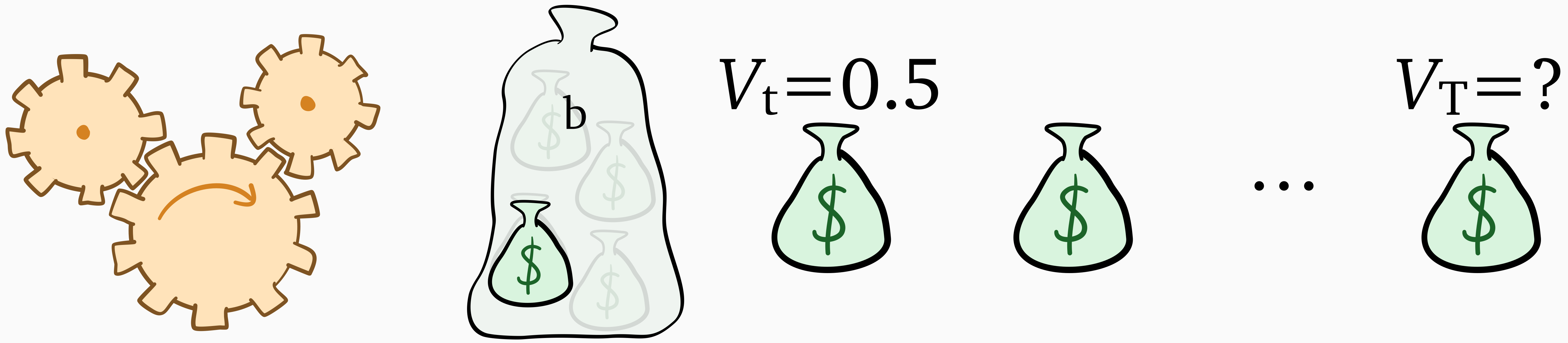
- **ONLINE** : Accept if $V_t \geq \theta$
- Let h=T-t, and define $V^h_{[b]} = \{b^{th} \text{ largest value in } V[t+1], V[t+2], \ldots, V[T]\}$

$$\mathbb{E}[\text{Comp}(t)] = \mathbb{E}\left[\mathbb{E}[|V^h_{[b]} - V_t| \mathbb{1}_{V_t \in [\theta, V^h_{[b]}]} | V_t] \middle| V^h_{[b]}\right]$$

$$= \mathbb{E}\left[(V^h_{[b]} - \theta)^2/2 | V^h_{[b]}\right] \leq Var(V^h_{[b]}) = \Theta(1/h)$$



Example: $V_t \sim U[0,1]$, i.i.d.



- **ONLINE** : Accept if $V_t \geq \theta$
- Let h=T-t, and define $V_{[b]}^h = \{b^{th} \text{ largest value in } V[t+1], V[t+2], \ldots, V[T]\}$

$$\mathbb{E}[\text{Comp}(t)] = \mathbb{E}\left[\mathbb{E}[|V_{[b]}^h - V_t|\mathbb{1}_{V_t \in [\theta, V_{[b]}^h]}|V_t]\middle| V_{[b]}^h\right]$$

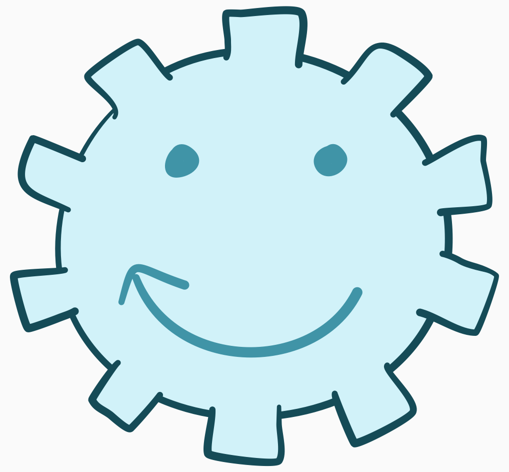
$$= \mathbb{E}\left[(V_{[b]}^h - \theta)^2/2|V_{[b]}^h\right] \leq Var(V_{[b]}^h) = \Theta(1/h)$$

$$\mathbb{E}[\text{Regret}] = \Theta(\log T)$$

Rob Bray: Does the Multisecretary Problem Always Have Bounded Regret? (2020)

Coupling to **easy system** with more information

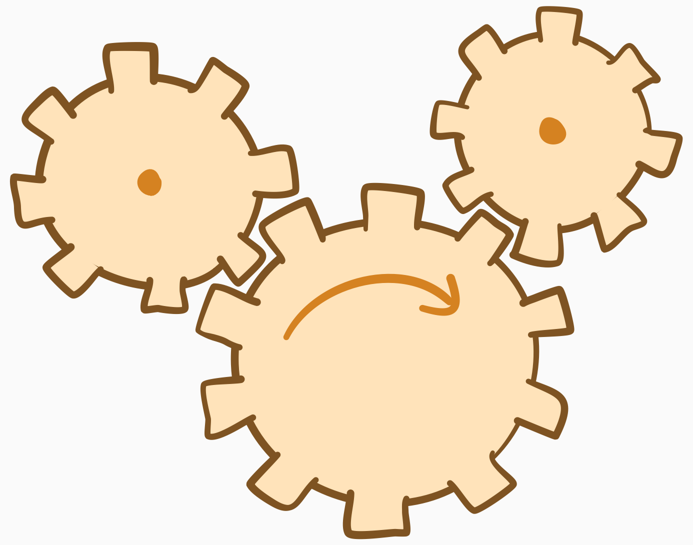
OFFLINE



$\mathcal{F}_t$ = 'natural' filtration

$\mathcal{G}_t$ = 'richer' filtration (i.e., $\mathcal{F}_t \subseteq \mathcal{G}_t$)

ONLINE



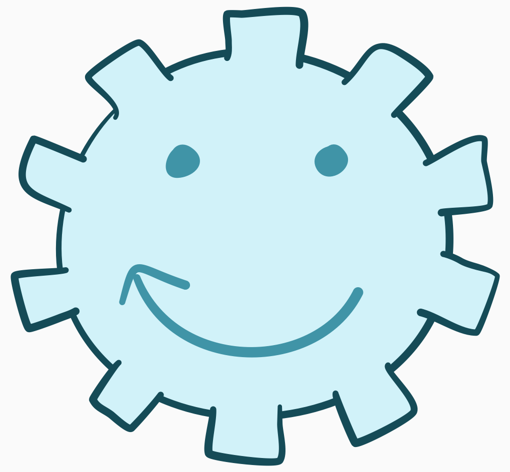
$\Phi^{\mathrm{OFF}}(\ )$, $Q^{\mathrm{OFF}}(\ ,\ )$: value/Q functions for **OFFLINE**

$\Phi^{\mathrm{ON}}(\ )$, $Q^{\mathrm{ON}}(\ ,\ )$: value/Q functions for **ONLINE**

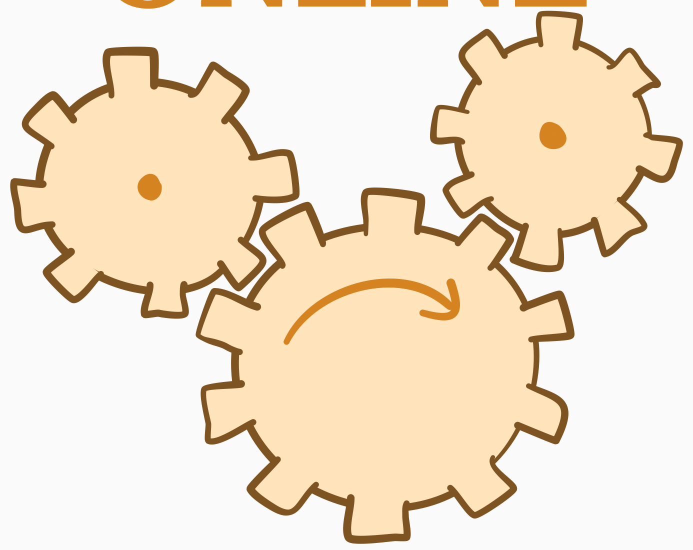
S_t , A_t : state/action taken by **ONLINE** in time slot t

Compensated coupling

OFFLINE

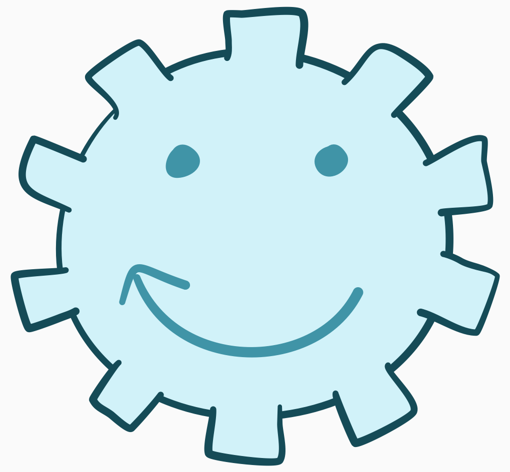


ONLINE

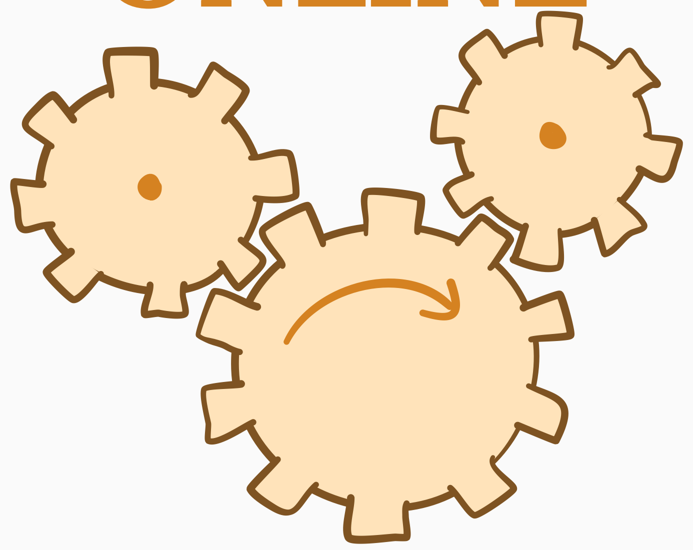


Compensated coupling

OFFLINE



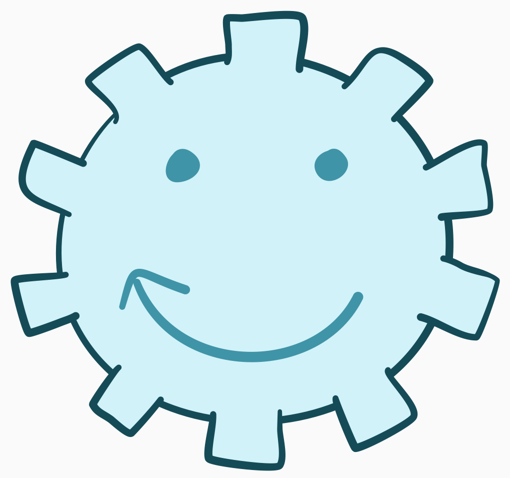
ONLINE



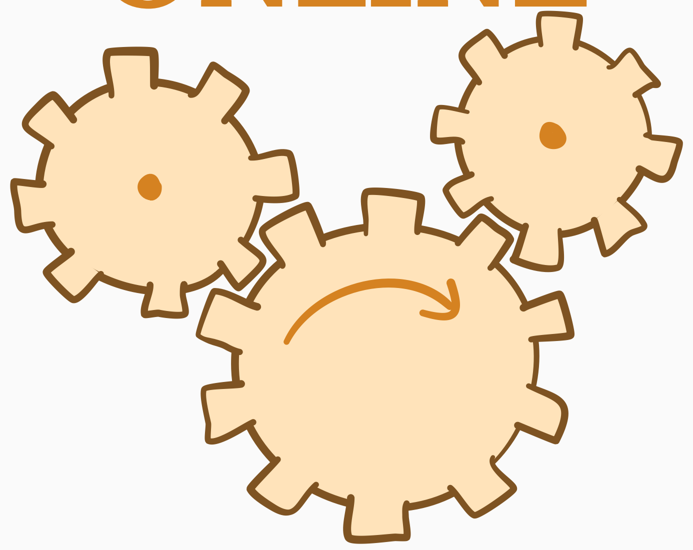
$$\Phi^{\mathrm{OFF}}(S_1) = R_1(S_1, A_1) + \Phi^{\mathrm{OFF}}(S_2) + \mathrm{Comp}(1)$$

Compensated coupling

OFFLINE



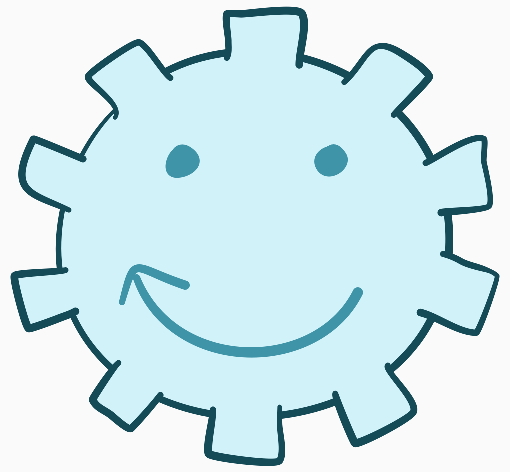
ONLINE

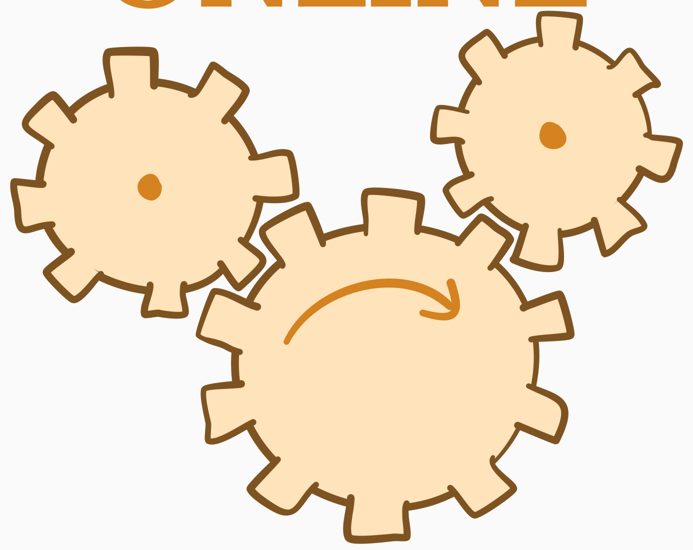


$$\Phi^{\mathrm{OFF}}(S_1) = R_1(S_1, A_1) + \Phi^{\mathrm{OFF}}(S_2) + \mathrm{Comp}(1)$$

$$\Phi^{\mathrm{OFF}}(S_2) = R_2(S_2, A_2) + \Phi^{\mathrm{OFF}}(S_3) + \mathrm{Comp}(2)$$

Compensated coupling





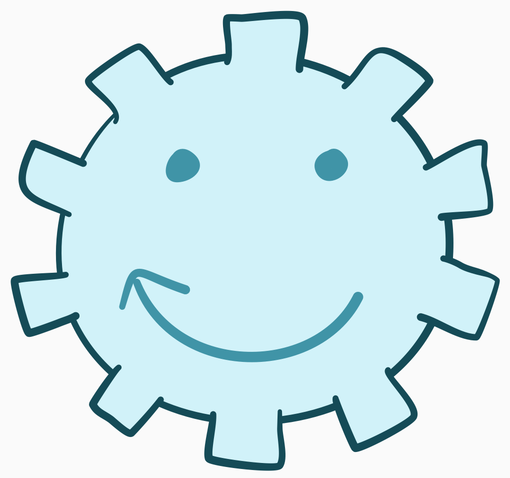
$$\Phi^{\mathrm{OFF}}(S_1) = R_1(S_1, A_1) + \Phi^{\mathrm{OFF}}(S_2) + \mathrm{Comp}(1)$$

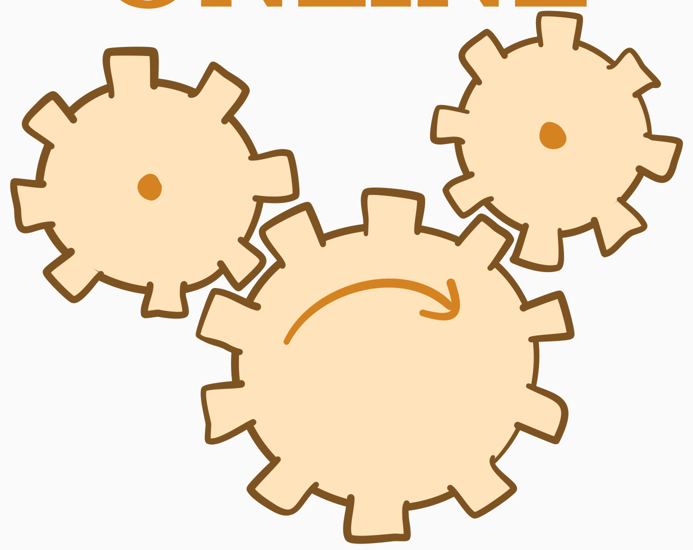
$$\Phi^{\mathrm{OFF}}(S_2) = R_2(S_2, A_2) + \Phi^{\mathrm{OFF}}(S_3) + \mathrm{Comp}(2)$$

$$\vdots$$

$$\Phi^{\mathrm{OFF}}(S_T) = R_T(S_T, A_T) + \Phi^{\mathrm{OFF}}(S_{T+1}) + \mathrm{Comp}(T)$$

Compensated coupling





$$\Phi^{\mathrm{OFF}}(S_1) = \overbrace{R_1(S_1, A_1) + \Phi^{\mathrm{OFF}}(S_2)}^{Q^{\mathrm{OFF}}(S_1, A_1)} + \mathrm{Comp}(1)$$

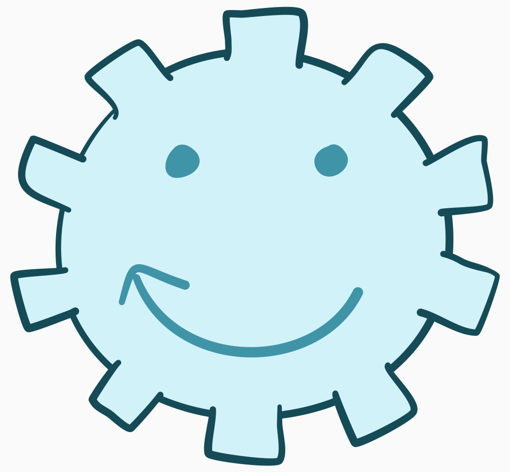
$$\Phi^{\mathrm{OFF}}(S_2) = R_2(S_2, A_2) + \Phi^{\mathrm{OFF}}(S_3) + \mathrm{Comp}(2)$$

$$\vdots$$

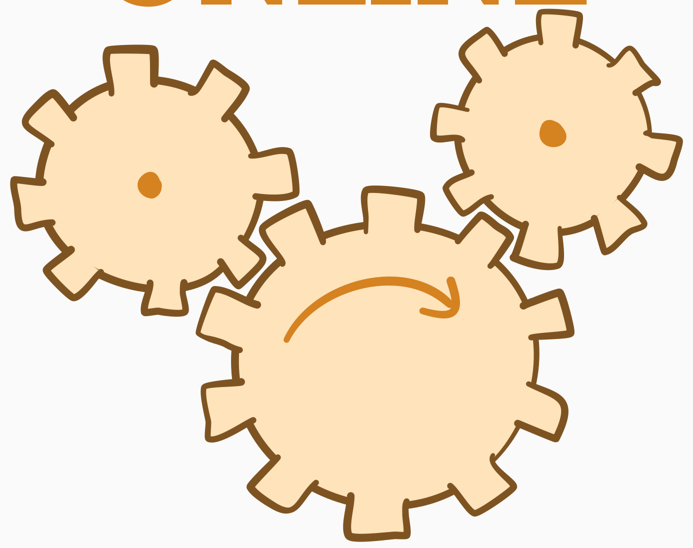
$$\Phi^{\mathrm{OFF}}(S_T) = R_T(S_T, A_T) + \Phi^{\mathrm{OFF}}(S_{T+1}) + \mathrm{Comp}(T)$$

Compensated coupling

OFFLINE



ONLINE



$$\underbrace{}_{}\quad Q^{\text{OFF}}(S_1, A_1)$$

$$\Phi^{\text{OFF}}(S_1) = \overbrace{R_1(S_1, A_1) + \Phi^{\text{OFF}}(S_2)}^{Q^{\text{OFF}}(S_1, A_1)} + \text{Comp}(1)$$

$$\Phi^{\text{OFF}}(S_2) = R_2(S_2, A_2) + \Phi^{\text{OFF}}(S_3) + \text{Comp}(2)$$

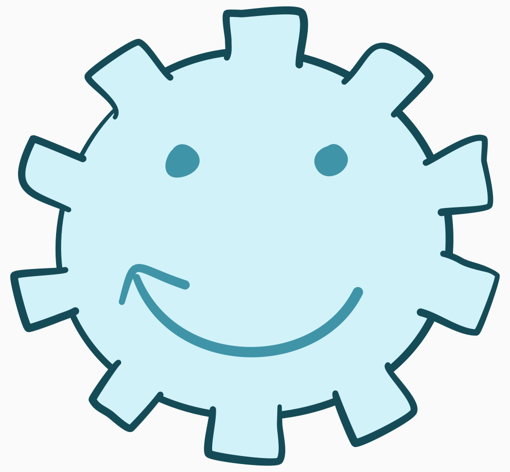
$$\vdots$$

$$\Phi^{\text{OFF}}(S_T) = R_T(S_T, A_T) + \Phi^{\text{OFF}}(S_{T+1}) + \text{Comp}(T)$$

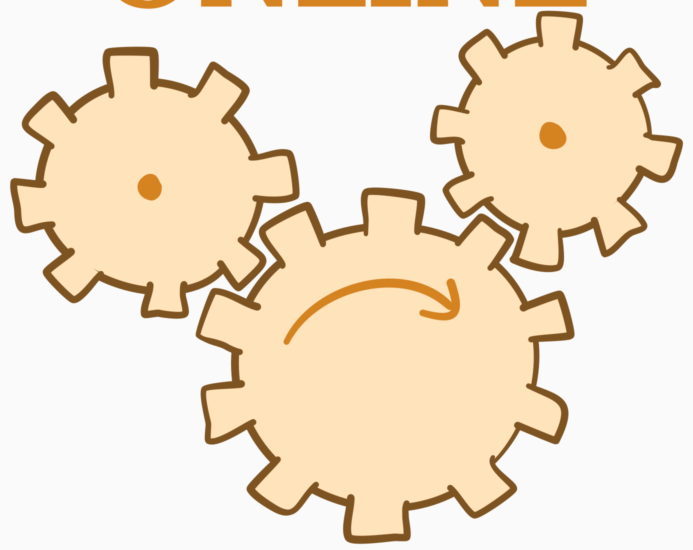
$$\Phi^{\text{OFF}}(S_1) = (R_1(S_1, A_1) + \ldots + R_T(S_T, A_T)) + (\text{Comp}(1) + \text{Comp}(T))$$

Compensated coupling

OFFLINE



ONLINE



$$\overbrace{R_1(S_1, A_1) + \Phi^{\text{OFF}}(S_2)}^{Q^{\text{OFF}}(S_1, A_1)}$$

$$\Phi^{\text{OFF}}(S_1) = R_1(S_1, A_1) + \Phi^{\text{OFF}}(S_2) + \text{Comp}(1)$$

$$\Phi^{\text{OFF}}(S_2) = R_2(S_2, A_2) + \Phi^{\text{OFF}}(S_3) + \text{Comp}(2)$$

$$\vdots$$

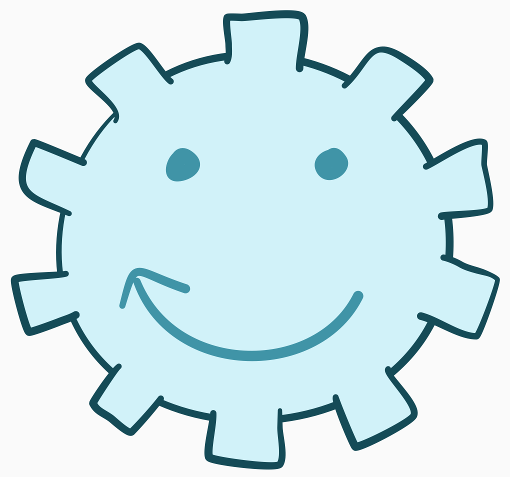
$$\Phi^{\text{OFF}}(S_T) = R_T(S_T, A_T) + \Phi^{\text{OFF}}(S_{T+1}) + \text{Comp}(T)$$

future value

$$\Phi^{\text{OFF}}(S_1) = R^{\text{ON}}(S_1) + \text{Total Comp}$$

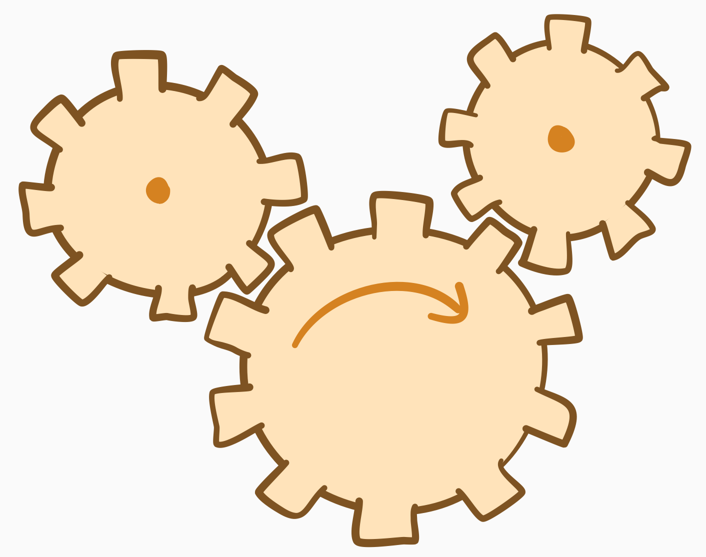
realized reward

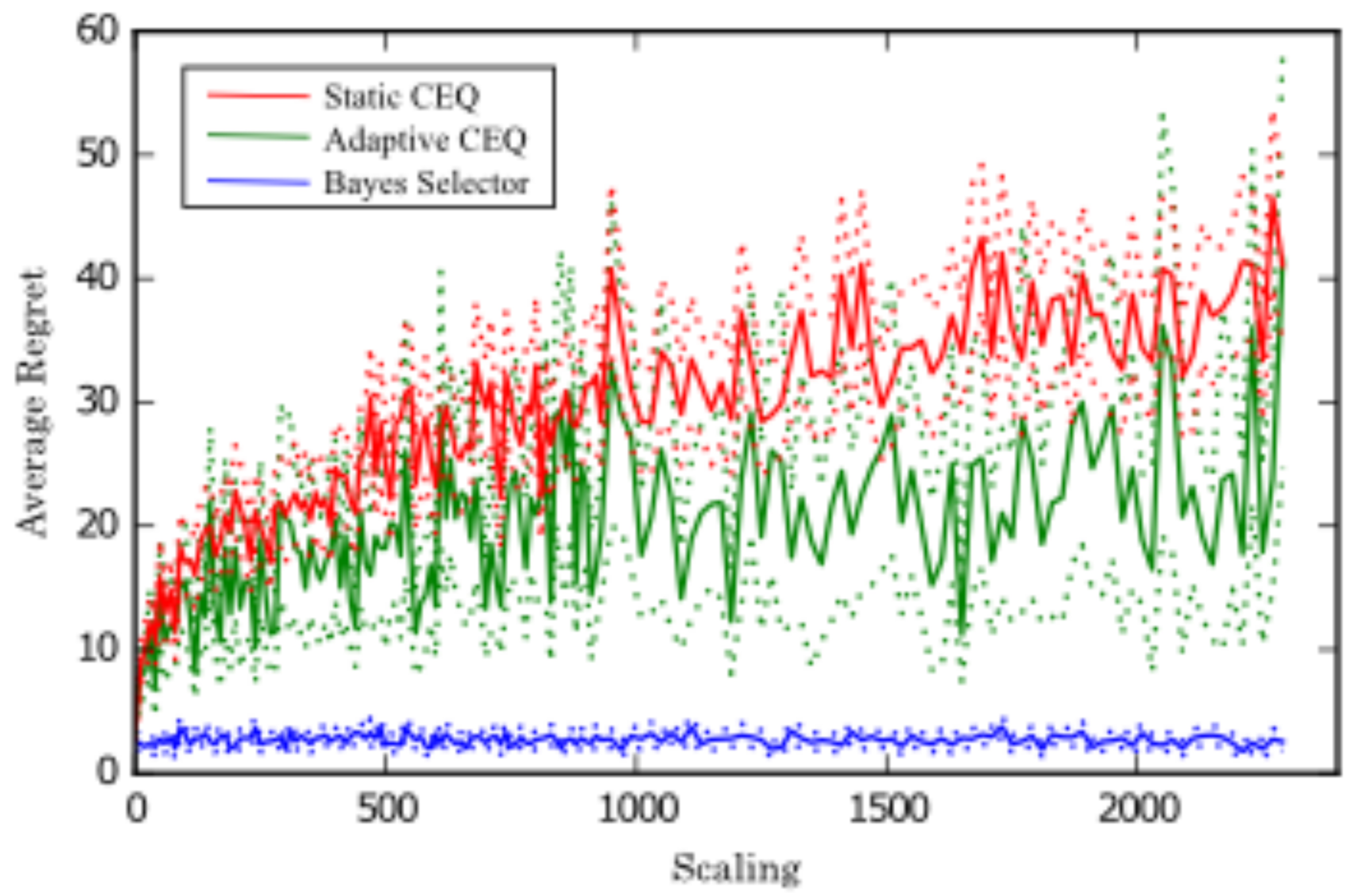
Compensated coupling: network revenue management



For NRM with d resources, finite types

$$\mathbb{E}[\text{Regret}] = \Theta(d)$$

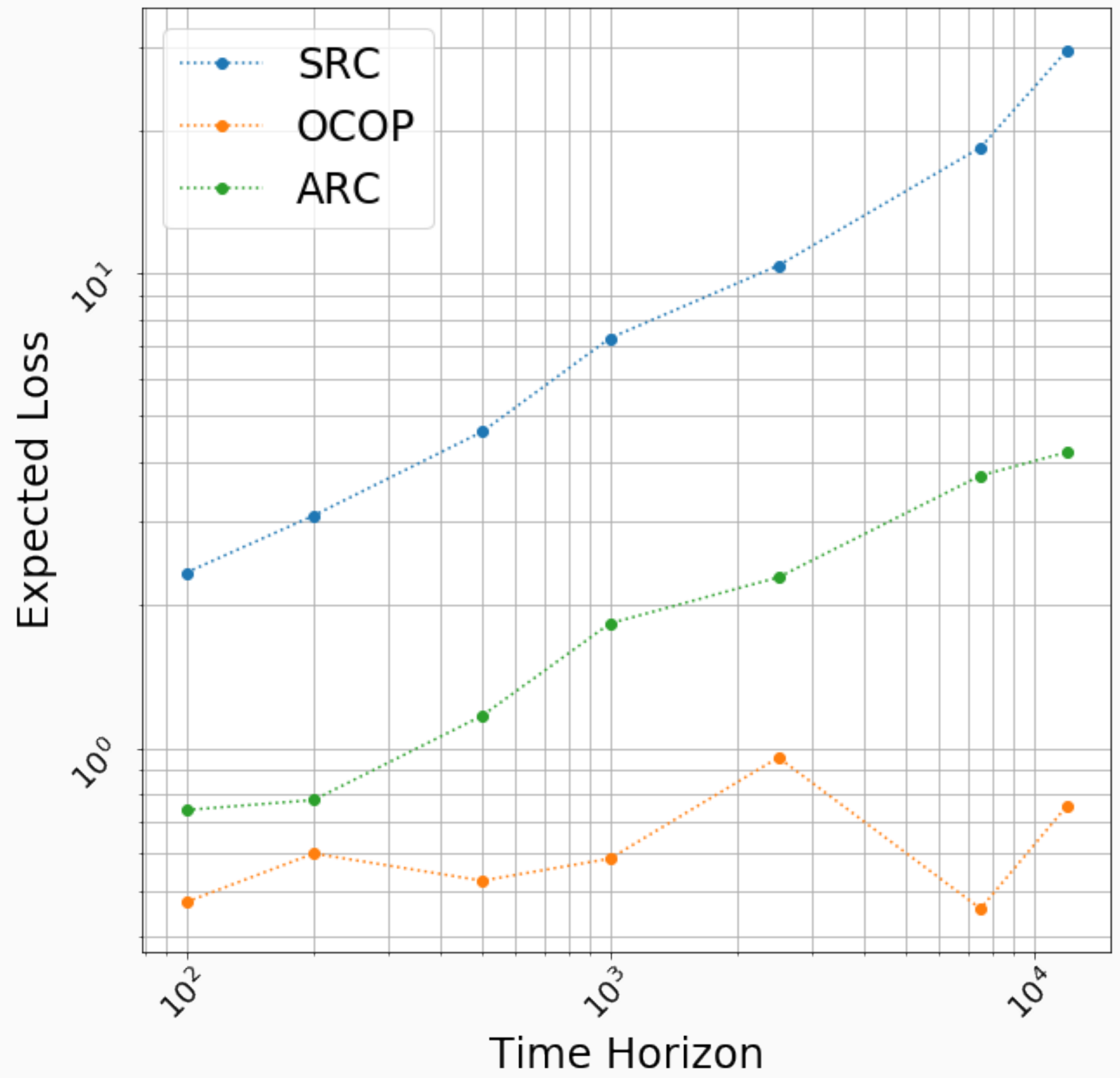




Vera and Banerjee: The Bayesian Prophet (2020)

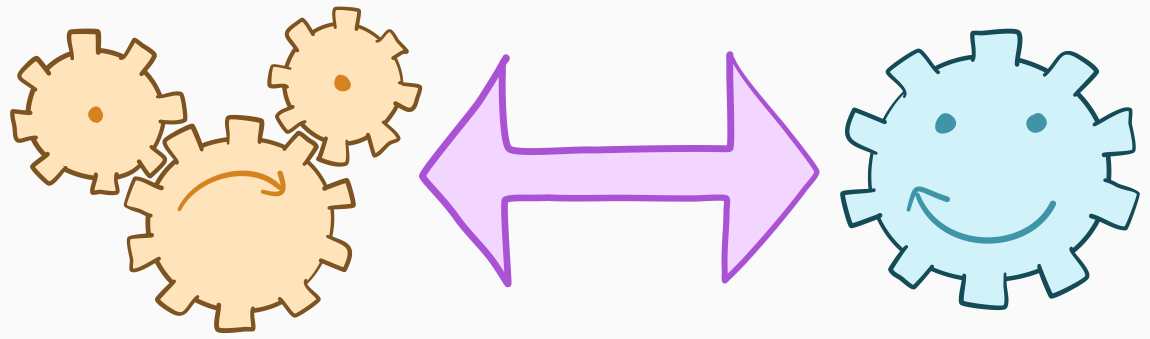
Compensated coupling: online bin packing

For online bin-packing with finite types
$\mathbb{E}[\text{Regret}] \leq \Theta(1)$



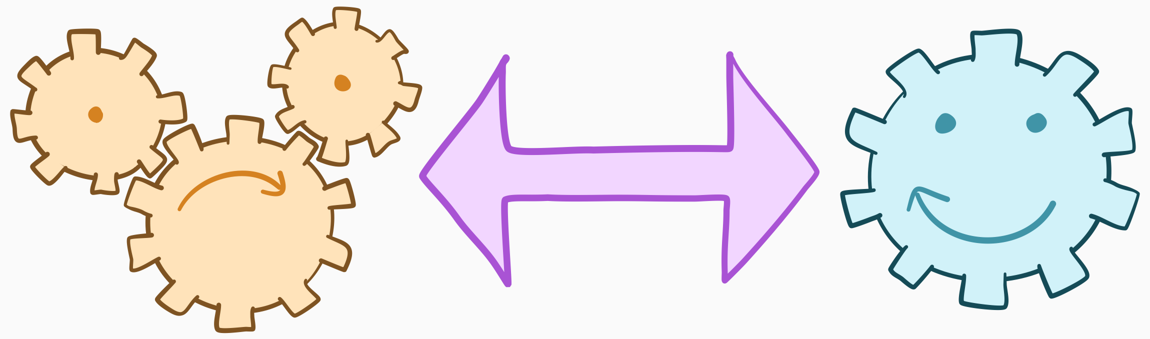
Banerjee and Freund: Uniform Loss Algorithms for Online Stochastic Decision-Making (2020)

Classifying coupling techniques



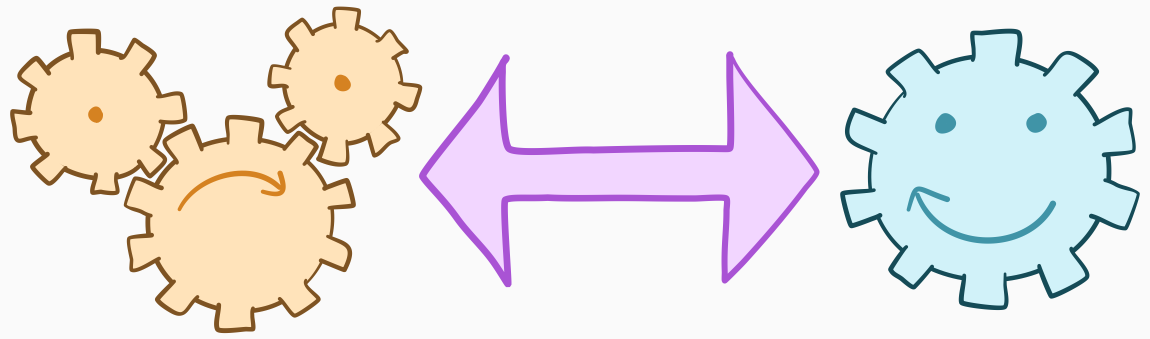
	A. Every sample path	B. Steady-state distribution
1. More information	A1	B1
2. Fewer constraints	A2 M/M/k vs. M/M/1 Online knapsack (via constraints-to-costs)	B2
3. Simpler dynamics	A3 SIS epidemics Queues with redundancy	B3

Classifying coupling techniques



	A. Every sample path	B. Steady-state distribution
1. More information	A1 BIG online knapsack (via compensated coupling)	B1
2. Fewer constraints	A2 M/M/k vs. M/M/1 Online knapsack (via constraints-to-costs)	B2
3. Simpler dynamics	A3 SIS epidemics Queues with redundancy	B3

Classifying coupling techniques



	A. Every sample path	B. Steady-state distribution
1. More information	A1 BIG online knapsack (via compensated coupling)	B1
2. Fewer constraints	A2 M/M/k vs. M/M/1 Online knapsack (via constraints-to-costs)	B2
3. Simpler dynamics	A3 SIS epidemics Queues with redundancy	B3

To be continued...

Overview

Part 1

Survey 1:
Sample-Path Coupling

In-Depth Study 1:
Online Resource Allocation

Part 2

Survey 2:
Steady-State Coupling

In-Depth Study 2:
Gittins in the M/G/k

Overview

✓ Part 1

Survey 1:
Sample-Path Coupling

In-Depth Study 1:
Online Resource Allocation

Part 2

Survey 2:
Steady-State Coupling

In-Depth Study 2:
Gittins in the M/G/k

Overview

✓ Part 1

Survey 1:
Sample-Path Coupling

In-Depth Study 1:
Online Resource Allocation

Part 2

tomorrow!

Survey 2:
Steady-State Coupling

In-Depth Study 2:
Gittins in the M/G/k