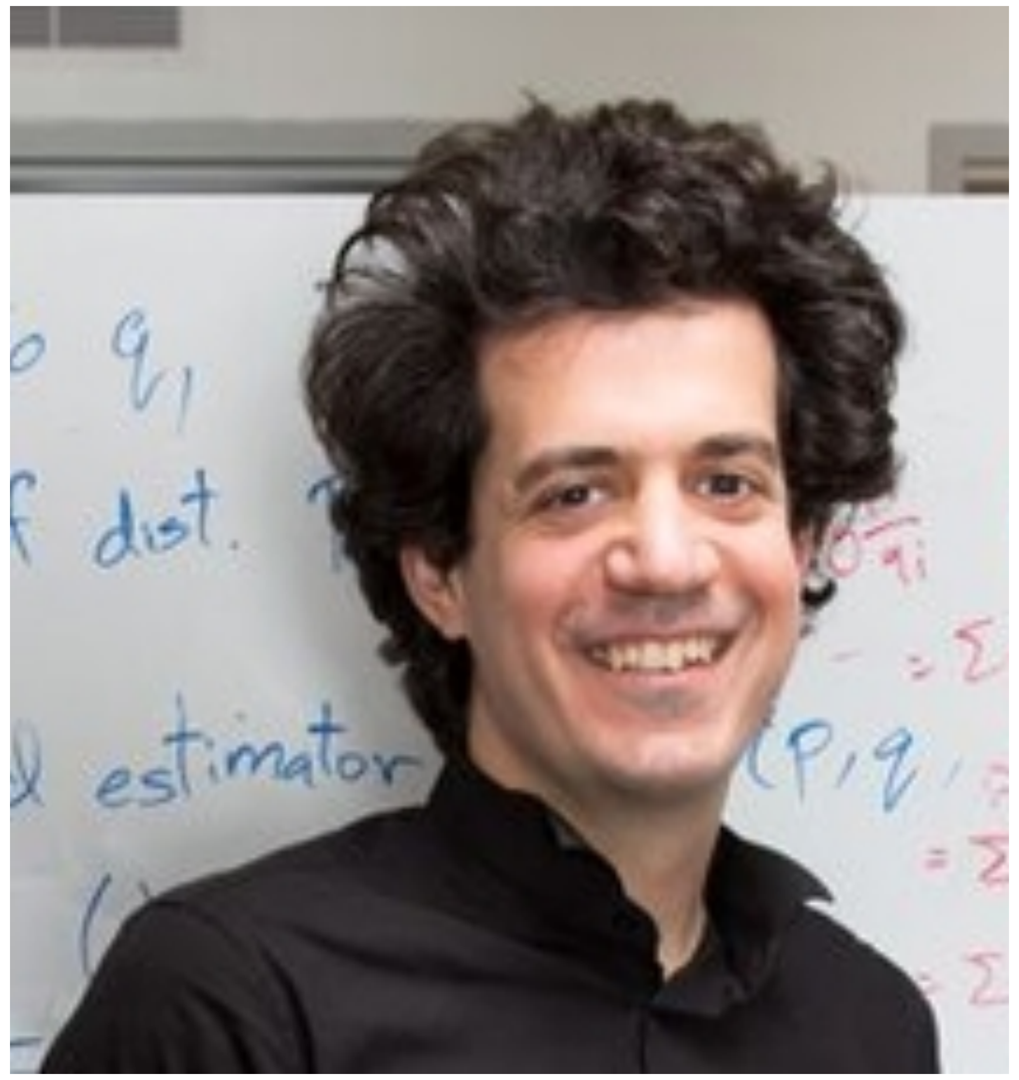
From External to Swap Regret 2.0: An Efficient Reduction for Large Action Spaces

Presentation by: Maxwell Fishelson

Collaborators

Yuval Dagan



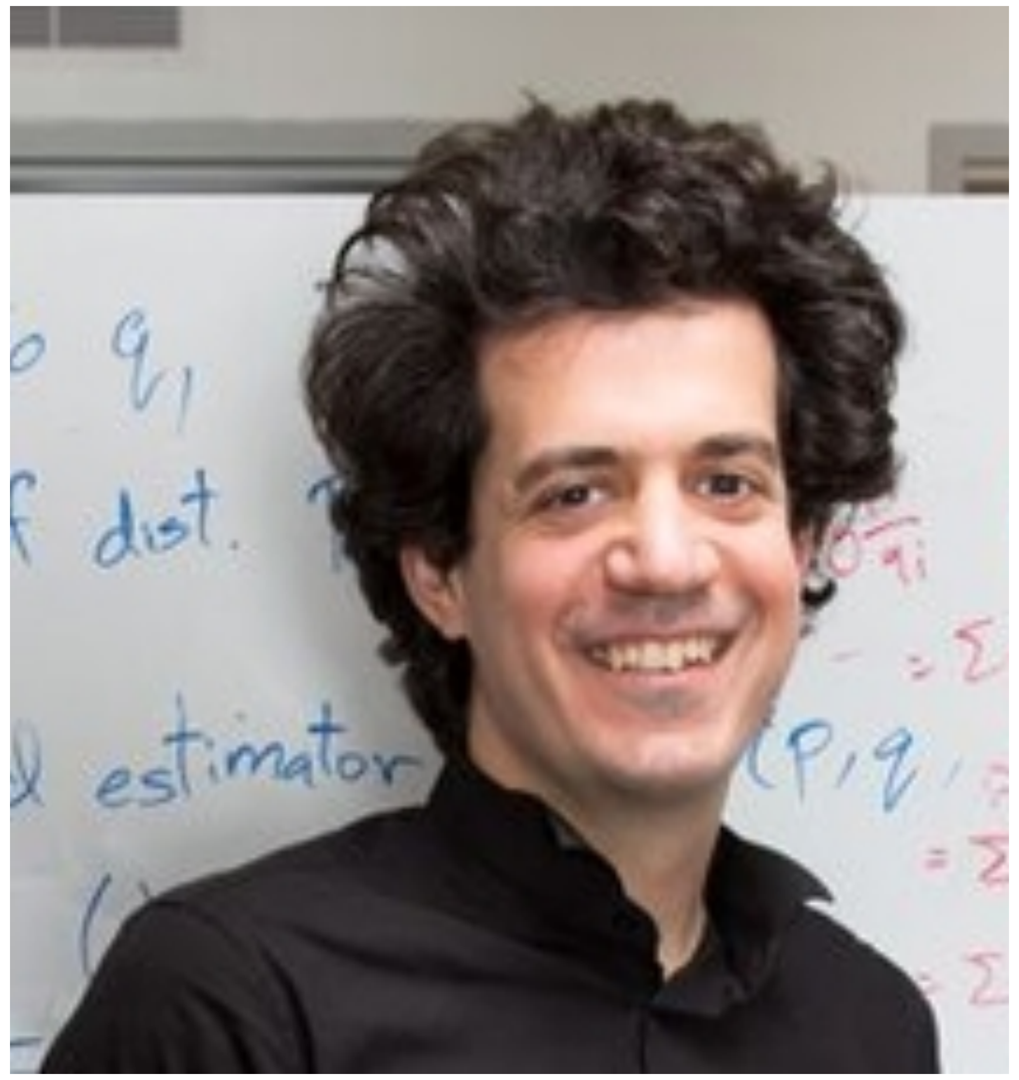
Costis Daskalakis

Maxwell Fishelson

Noah Golowich

Collaborators

Yuval Dagan



Costis Daskalakis

Maxwell Fishelson

Noah Golowich

Concurrent work: Binghui Peng, Aviad Rubinstein [PR23]

No Swap Regret

No Swap Regret
Online Learning

No Swap Regret Online Learning with many actions

No Swap Regret ?
Online Learning
with many actions

No Swap Regret ?
Online Learning ?
with many actions

No Swap Regret ?
Online Learning ?
with many actions ?

Online Learning

Online Learning

A	B	C	D	E

Online Learning

	A	B	C	D	E
1					

Online Learning

	A	B	C	D	E
1					

Online Learning

	A	B	C	D	E
1	0.1	0.3	0.9	0.5	0.7

Online Learning

	A	B	C	D	E
1	0.1	0.3	0.9	0.5	0.7
2					

Online Learning

	A	B	C	D	E
1	0.1	0.3	0.9	0.5	0.7
2					

Online Learning

	A	B	C	D	E
1	0.1	0.3	0.9	0.5	0.7
2	0.8	0.4	0	0.2	0.3

Online Learning

	A	B	C	D	E
1	0.1	0.3	0.9	0.5	0.7
2	0.8	0.4	0	0.2	0.3
3					

Online Learning

	A	B	C	D	E
1	0.1	0.3	0.9	0.5	0.7
2	0.8	0.4	0	0.2	0.3
3					

Online Learning

	A	B	C	D	E
1	0.1	0.3	0.9	0.5	0.7
2	0.8	0.4	0	0.2	0.3
3	0.1	0.2	0.3	0.4	0.5

Online Learning

	A	B	C	D	E
1	0.1	0.3	0.9	0.5	0.7
2	0.8	0.4	0	0.2	0.3
3	0.1	0.2	0.3	0.4	0.5

$$x^{(t)} \in \Delta_N$$

Online Learning

	A	B	C	D	E
1	0.1	0.3	0.9	0.5	0.7
2	0.8	0.4	0	0.2	0.3
3	0.1	0.2	0.3	0.4	0.5

$x^{(t)} \in \Delta_N$

$u^{(t)} \in [0,1]^N$

Online Learning

	A	B	C	D	E
1	0.1	0.3	0.9	0.5	0.7
2	0.8	0.4	0	0.2	0.3
3	0.1	0.2	0.3	0.4	0.5

$$x^{(t)} \in \Delta_N$$

$$u^{(t)} \in [0,1]^N$$

$$\text{Total Reward} = \sum_{t=1}^{T} x^{(t)} \cdot u^{(t)}$$

Online Learning

	A	B	C	D	E
1	0.1	0.3	0.9	0.5	0.7
2	0.8	0.4	0	0.2	0.3
3	0.1	0.2	0.3	0.4	0.5

$$x^{(t)} \in \Delta_N$$

$$u^{(t)} \in [0,1]^N$$

$$\text{Total Reward} = \sum_{t=1}^{T} x^{(t)} \cdot u^{(t)}$$

$$\text{Average Reward} = \frac{1}{T} \sum_{t=1}^{T} x^{(t)} \cdot u^{(t)}$$

Online vs Offline

Online vs Offline

Online: $u^{(t)} \leftarrow$ 😈

Online vs Offline

Online: $u^{(t)} \leftarrow$ 😈

Offline: $u^{(t)} \sim D$

Online vs Offline

Online: $u^{(t)} \leftarrow$ 😈

Offline: $u^{(t)} \sim D$

Games:

Online vs Offline

Online: $u^{(t)} \leftarrow$ 😈 **Offline:** $u^{(t)} \sim D$

Games:

Online vs Offline

Online: $u^{(t)} \leftarrow$ 😈 **Offline:** $u^{(t)} \sim D$

Games:

🪨:100

🤔 📄:120 😈

✂️:110

Online vs Offline

Online: $u^{(t)} \leftarrow$ 😈

Offline: $u^{(t)} \sim D$

Games:

$x^{(t)} =$ ✂️

🤔

🪨**:100**

📄**:120**

✂️**:110**

😈

Online vs Offline

Online: $u^{(t)} \leftarrow$ 😈

Offline: $u^{(t)} \sim D$

Games:

$x^{(t)} =$ ✂️ 🤔

🪨:100

📄:120

✂️:110

😈 $u^{(t)} =$ 🪨

Regret

Regret

regret = reward obtained vs benchmark

Regret

regret = reward obtained vs benchmark

$x^{(1)}$ $u^{(1)}$

Regret

regret = reward obtained vs benchmark

$x^{(1)}$ $u^{(1)}$

$x^{(2)}$ $u^{(2)}$

Regret

regret = reward obtained vs benchmark

$x^{(1)}$	$u^{(1)}$
$x^{(2)}$	$u^{(2)}$
$x^{(3)}$	$u^{(3)}$

Regret

regret = reward obtained vs benchmark

$x^{(1)}$	$u^{(1)}$
$x^{(2)}$	$u^{(2)}$
$x^{(3)}$	$u^{(3)}$
$\vdots$	
$x^{(T)}$	$u^{(T)}$

Regret

regret = reward obtained vs benchmark

$$
\begin{array}{cc}
x^{(1)} & u^{(1)} \\
x^{(2)} & u^{(2)} \\
x^{(3)} & u^{(3)} \\
\vdots & \\
x^{(T)} & u^{(T)}
\end{array}
$$

$$\frac{1}{T}\sum_{t=1}^{T} x^{(t)} \cdot u^{(t)}$$

Regret

regret = reward obtained vs benchmark

$x^{(1)}$	$u^{(1)}$
$x^{(2)}$	$u^{(2)}$
$x^{(3)}$	$u^{(3)}$
$\vdots$	
$x^{(T)}$	$u^{(T)}$

$$\frac{1}{T}\sum_{t=1}^{T} x^{(t)} \cdot u^{(t)}$$

ϕ →

Regret

regret = reward obtained vs benchmark

$x^{(1)}$	$u^{(1)}$
$x^{(2)}$	$u^{(2)}$
$x^{(3)}$	$u^{(3)}$
$\vdots$	
$x^{(T)}$	$u^{(T)}$

$\xrightarrow{\phi}$

$\phi(x^{(1)})$	$u^{(1)}$
$\phi(x^{(2)})$	$u^{(2)}$
$\phi(x^{(3)})$	$u^{(3)}$
$\vdots$	
$\phi(x^{(T)})$	$u^{(T)}$

$$\frac{1}{T}\sum_{t=1}^{T} x^{(t)} \cdot u^{(t)}$$

Regret

regret = reward obtained vs benchmark

$x^{(1)}$	$u^{(1)}$
$x^{(2)}$	$u^{(2)}$
$x^{(3)}$	$u^{(3)}$
$\vdots$	
$x^{(T)}$	$u^{(T)}$

$$\frac{1}{T}\sum_{t=1}^{T} x^{(t)} \cdot u^{(t)}$$

$\xrightarrow{\ \phi\ }$

$\phi(x^{(1)})$	$u^{(1)}$
$\phi(x^{(2)})$	$u^{(2)}$
$\phi(x^{(3)})$	$u^{(3)}$
$\vdots$	
$\phi(x^{(T)})$	$u^{(T)}$

$$\frac{1}{T}\sum_{t=1}^{T} \phi(x^{(t)}) \cdot u^{(t)}$$

Regret

regret = reward obtained vs benchmark

$x^{(1)}$	$u^{(1)}$
$x^{(2)}$	$u^{(2)}$
$x^{(3)}$	$u^{(3)}$
$\vdots$	
$x^{(T)}$	$u^{(T)}$

$$\frac{1}{T}\sum_{t=1}^{T} x^{(t)} \cdot u^{(t)}$$

ϕ →

Benchmark:
Best $\phi \in \Phi$

$\phi(x^{(1)})$	$u^{(1)}$
$\phi(x^{(2)})$	$u^{(2)}$
$\phi(x^{(3)})$	$u^{(3)}$
$\vdots$	
$\phi(x^{(T)})$	$u^{(T)}$

$$\frac{1}{T}\sum_{t=1}^{T} \phi(x^{(t)}) \cdot u^{(t)}$$

Regret

regret = reward obtained vs benchmark

$x^{(1)}$	$u^{(1)}$
$x^{(2)}$	$u^{(2)}$
$x^{(3)}$	$u^{(3)}$
$\vdots$	
$x^{(T)}$	$u^{(T)}$

$\xrightarrow{\phi}$

Benchmark:
Best $\phi \in \Phi$

$\phi(x^{(1)})$	$u^{(1)}$
$\phi(x^{(2)})$	$u^{(2)}$
$\phi(x^{(3)})$	$u^{(3)}$
$\vdots$	
$\phi(x^{(T)})$	$u^{(T)}$

$$\frac{1}{T}\sum_{t=1}^{T} x^{(t)} \cdot u^{(t)}$$

$$\frac{1}{T}\sum_{t=1}^{T} \phi(x^{(t)}) \cdot u^{(t)}$$

$$\text{Regret}\left(x^{(1:T)}, u^{(1:T)}\right) = \max_{\phi \in \Phi}\left(\frac{1}{T}\sum_{t=1}^{T} \phi(x^{(t)}) \cdot u^{(t)}\right) - \left(\frac{1}{T}\sum_{t=1}^{T} x^{(t)} \cdot u^{(t)}\right)$$

Regret

regret = reward obtained vs benchmark

$x^{(1)}$	$u^{(1)}$
$x^{(2)}$	$u^{(2)}$
$x^{(3)}$	$u^{(3)}$
$\vdots$	
$x^{(T)}$	$u^{(T)}$

ϕ $\longrightarrow$

Benchmark:
Best $\phi \in \Phi$

$\phi(x^{(1)})$	$u^{(1)}$
$\phi(x^{(2)})$	$u^{(2)}$
$\phi(x^{(3)})$	$u^{(3)}$
$\vdots$	
$\phi(x^{(T)})$	$u^{(T)}$

$$\frac{1}{T}\sum_{t=1}^{T} x^{(t)} \cdot u^{(t)} \qquad \frac{1}{T}\sum_{t=1}^{T} \phi(x^{(t)}) \cdot u^{(t)}$$

$$\text{Regret}\left(x^{(1:T)}, u^{(1:T)}\right) = \max_{\phi\in\Phi}\left(\frac{1}{T}\sum_{t=1}^{T}\phi(x^{(t)})\cdot u^{(t)}\right) - \left(\frac{1}{T}\sum_{t=1}^{T}x^{(t)}\cdot u^{(t)}\right)$$

Why?

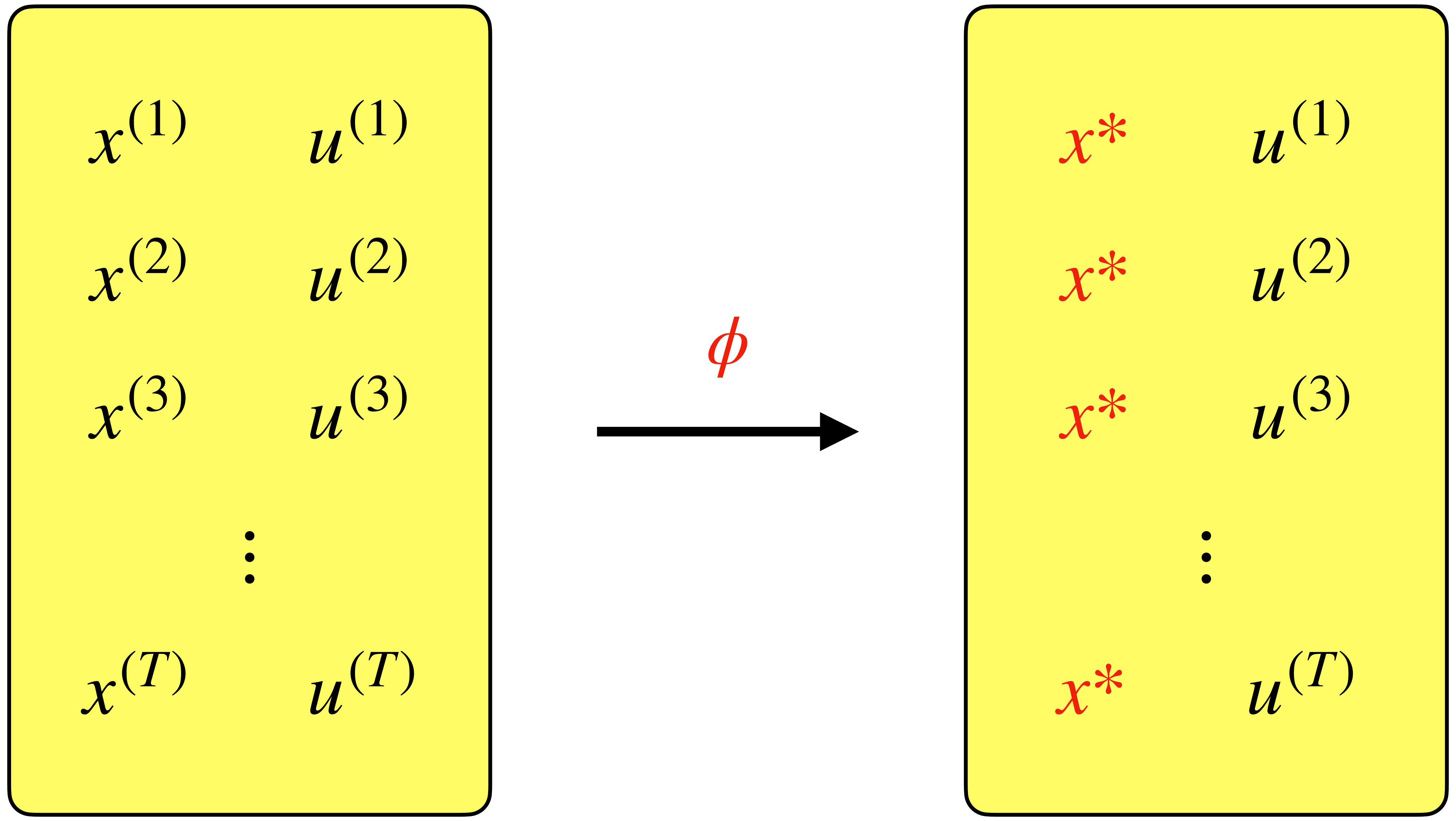
External Regret

External Regret

$\Phi =$ **all constant maps**

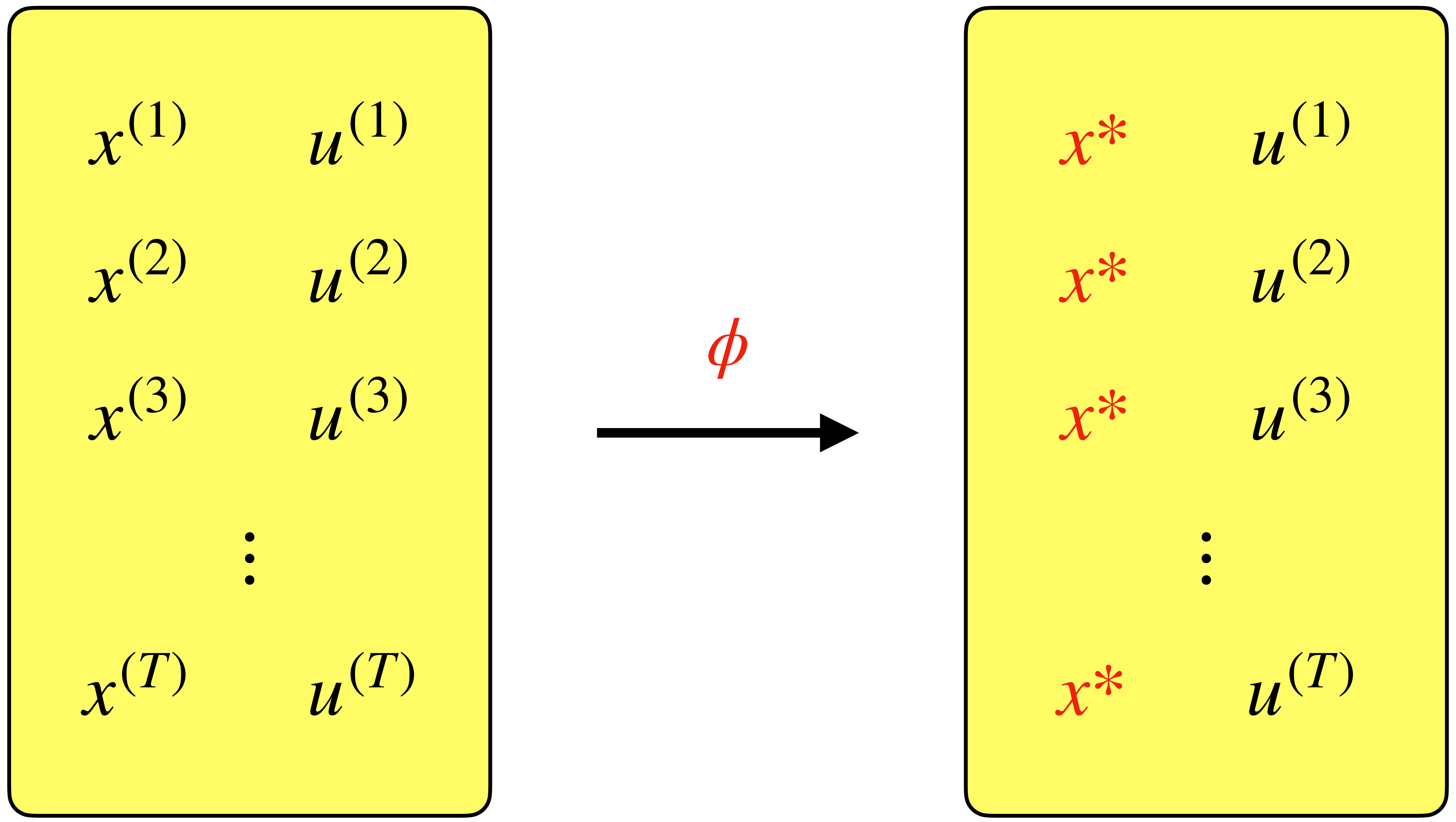
External Regret

Φ = all constant maps

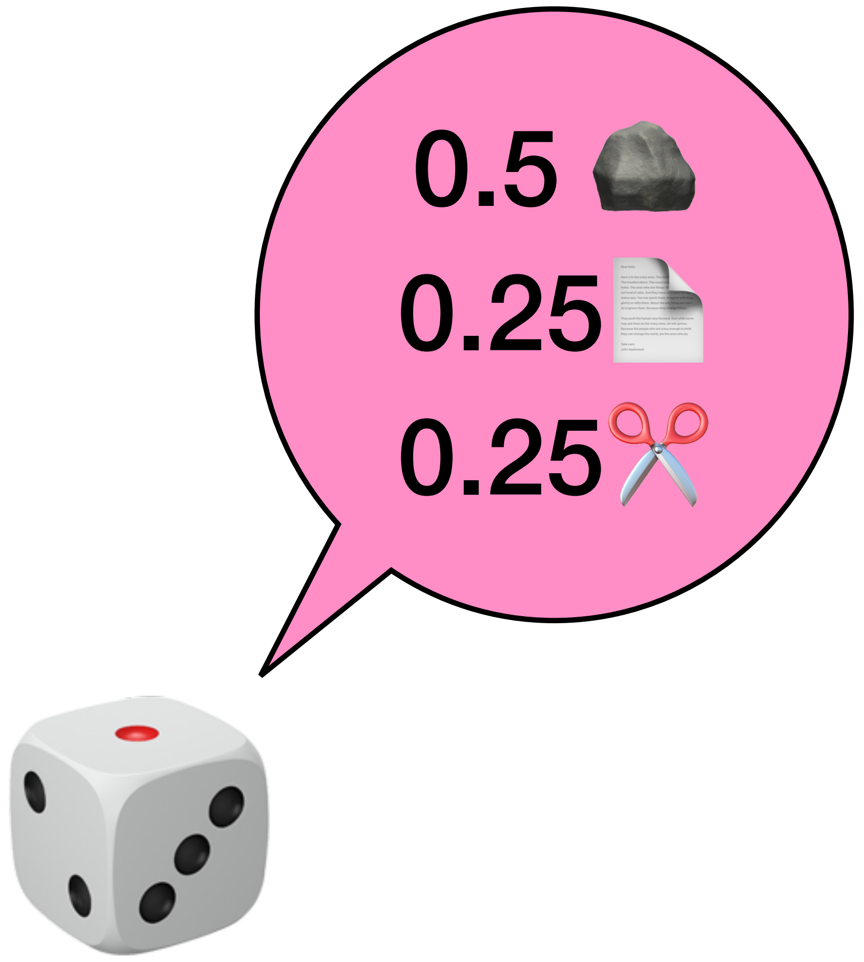


External Regret

Φ = all constant maps

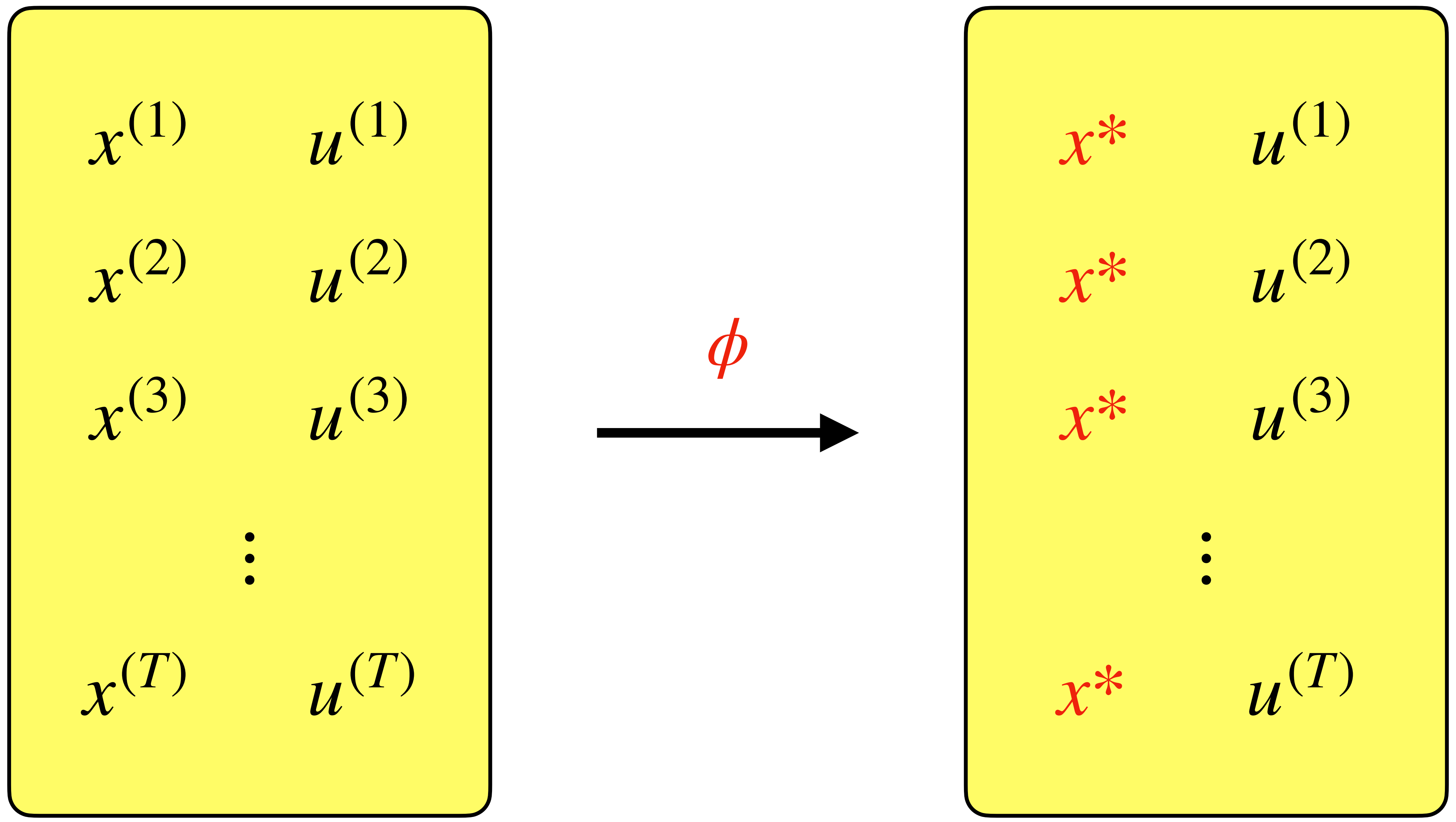


Offline:

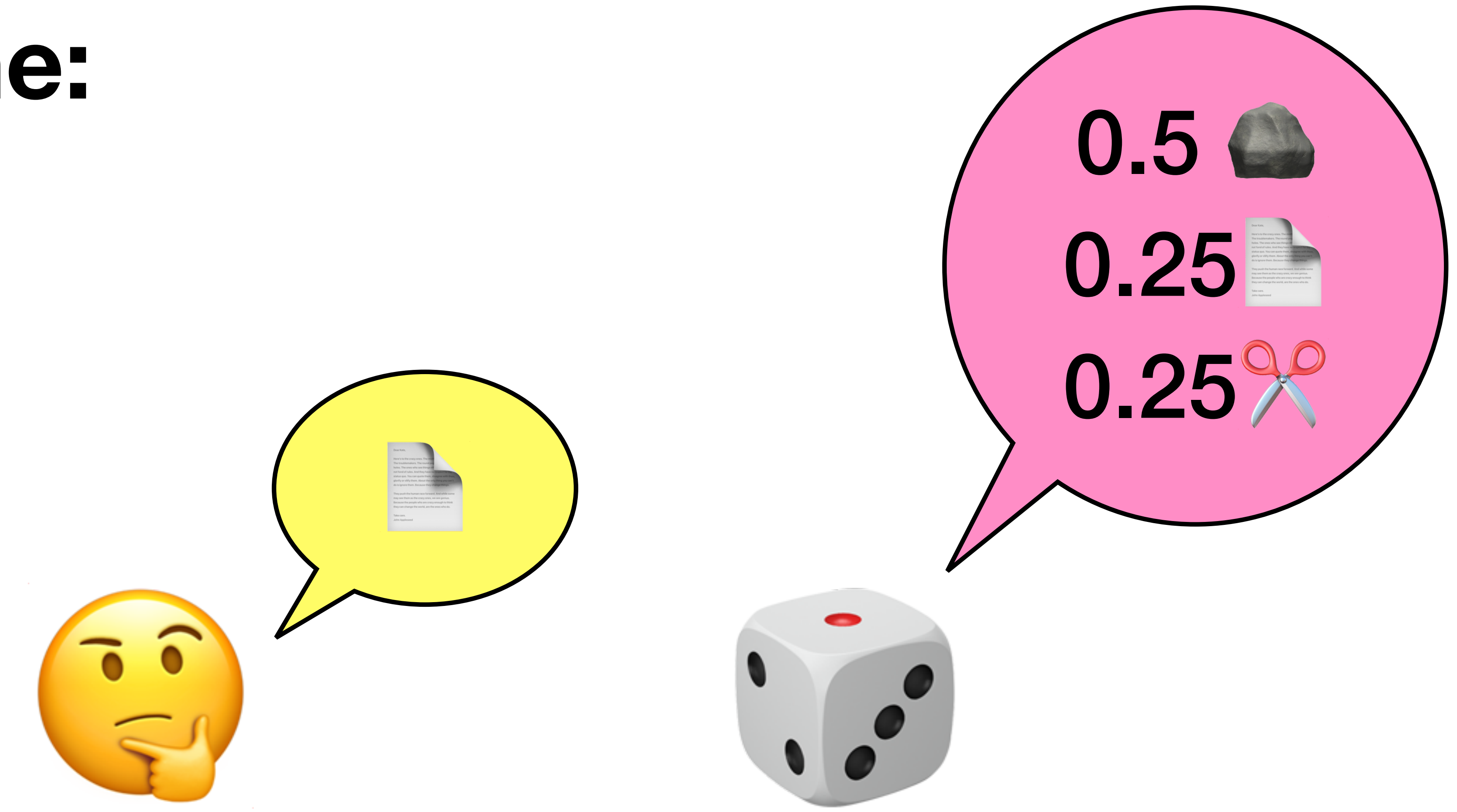


External Regret

Φ = all constant maps

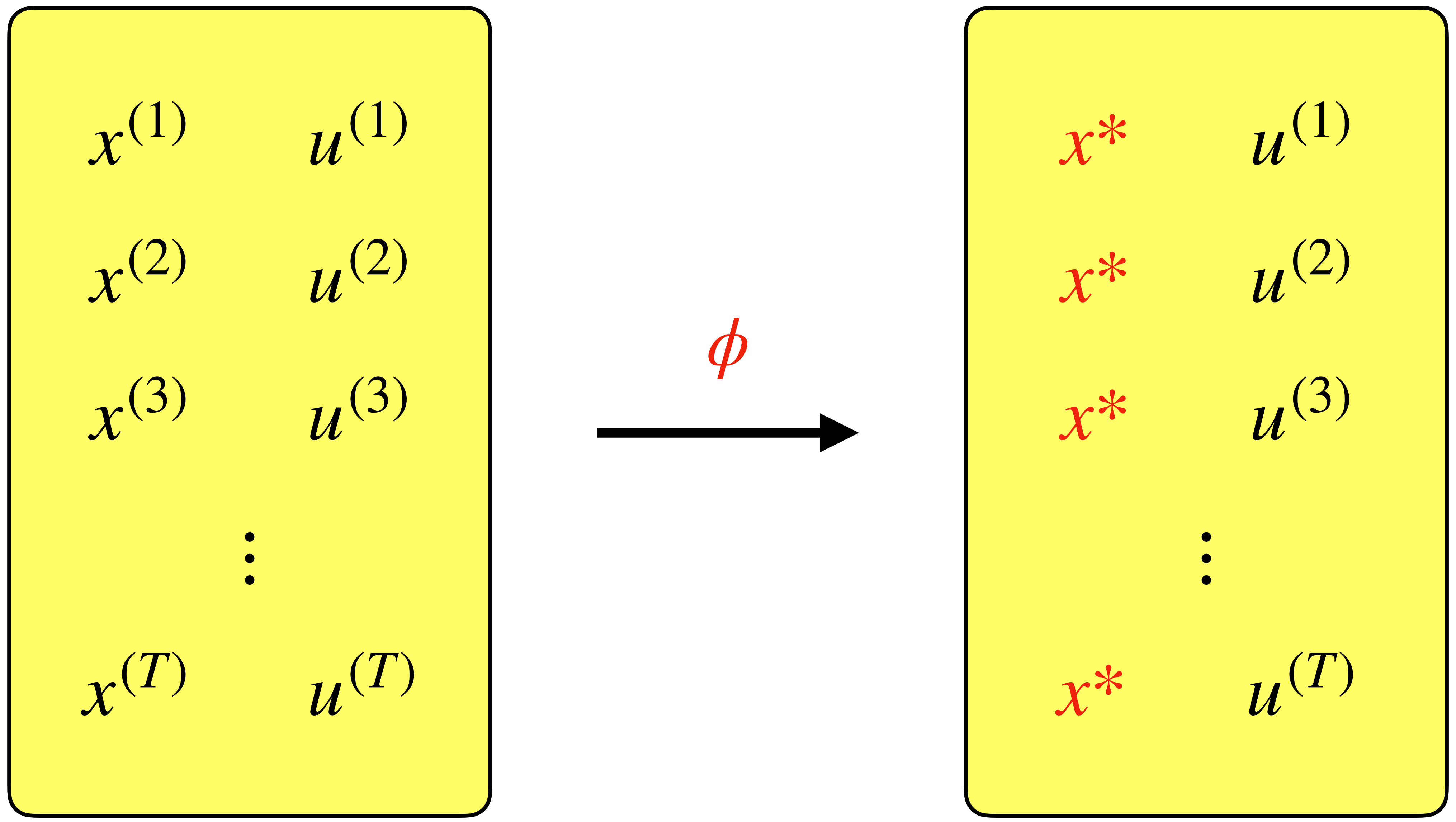


Offline:

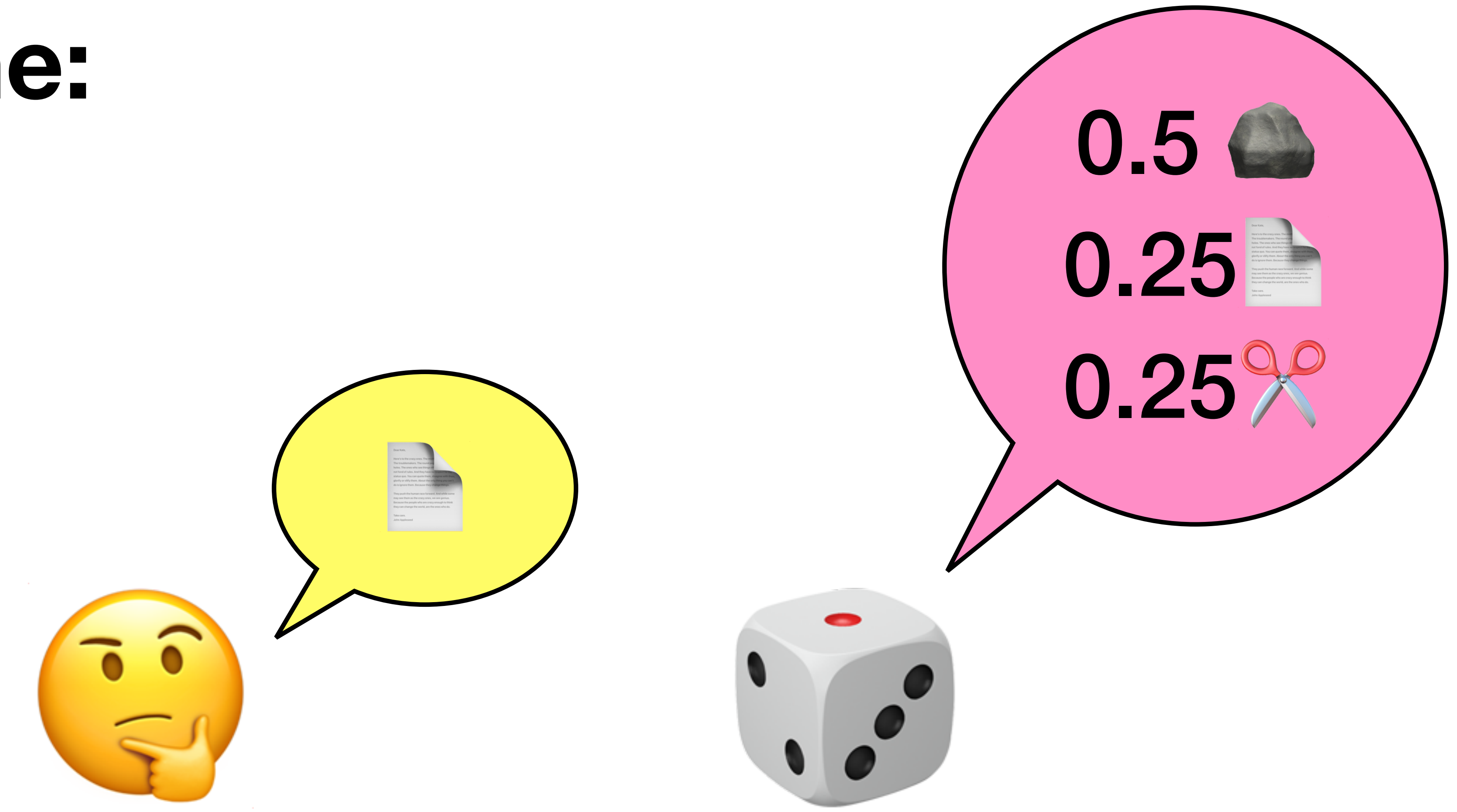


External Regret

Φ = all constant maps



Offline:



Online setting: we can do better!

Swap Regret

Swap Regret

$\Phi =$ **all maps** $[N] \to [N]$

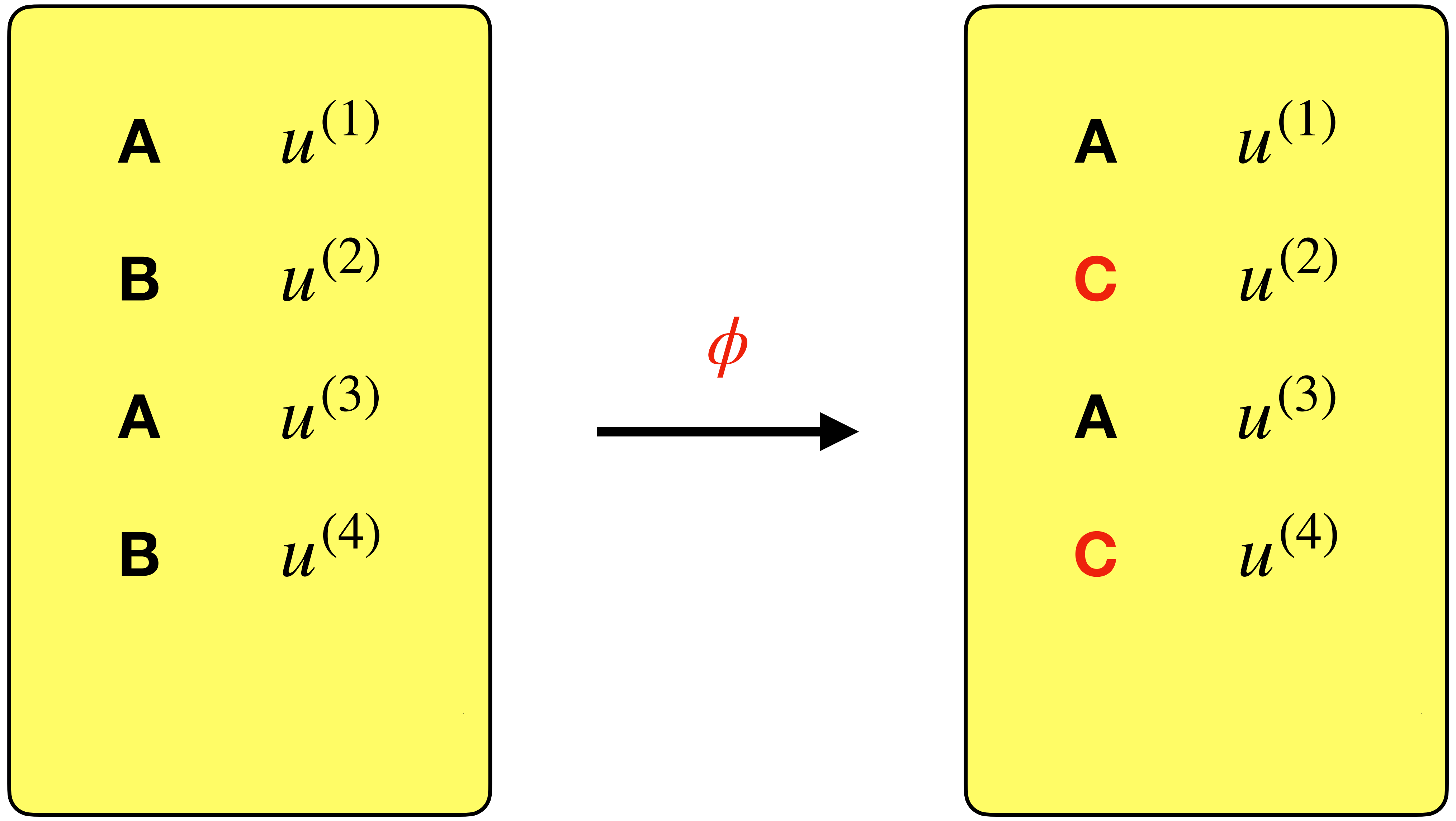
Swap Regret

Φ = **all maps** $[N] \to [N]$

A	$u^{(1)}$
B	$u^{(2)}$
A	$u^{(3)}$
B	$u^{(4)}$

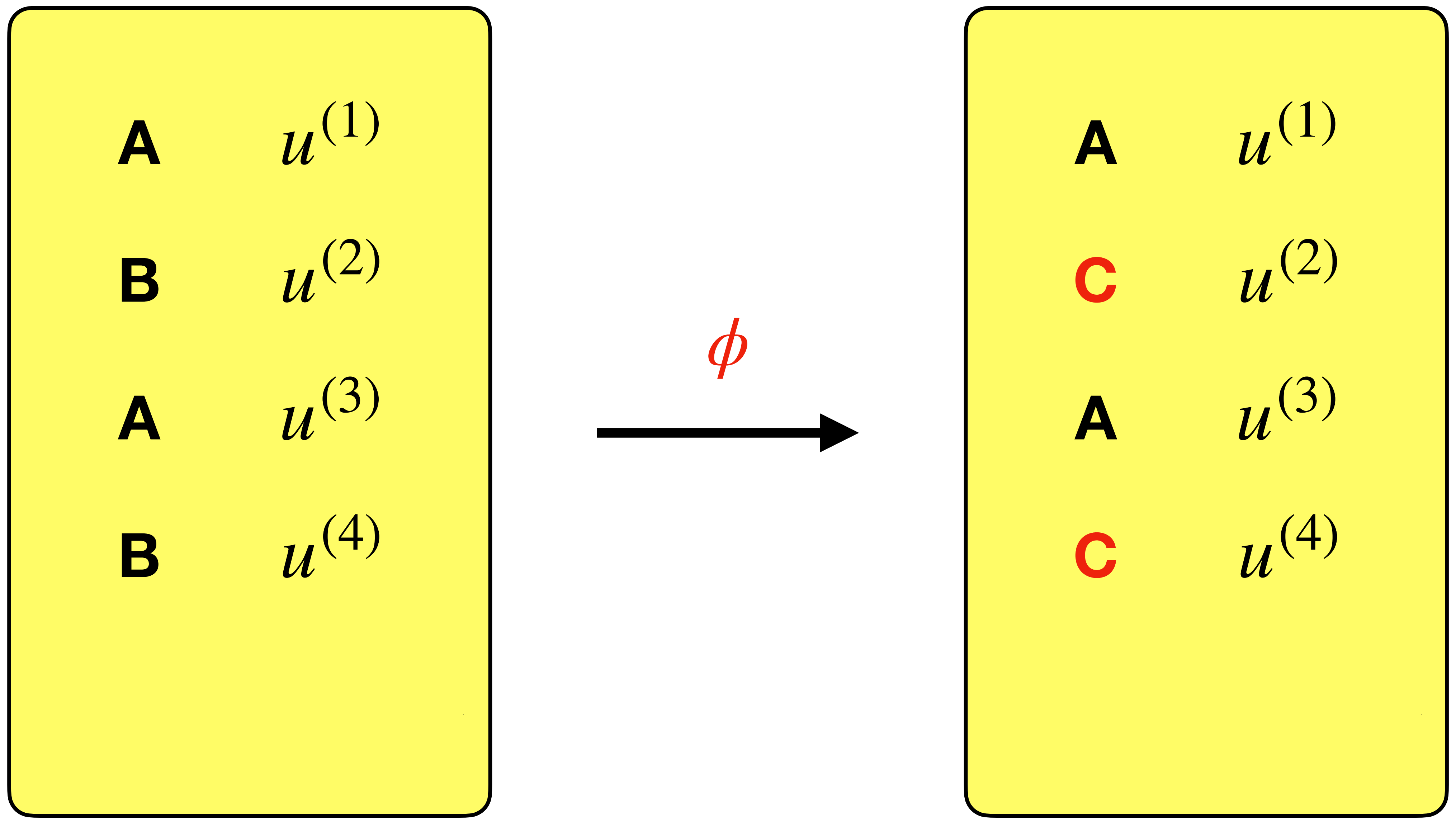
Swap Regret

Φ = all maps $[N] \to [N]$



Swap Regret

Φ = all maps $[N] \to [N]$

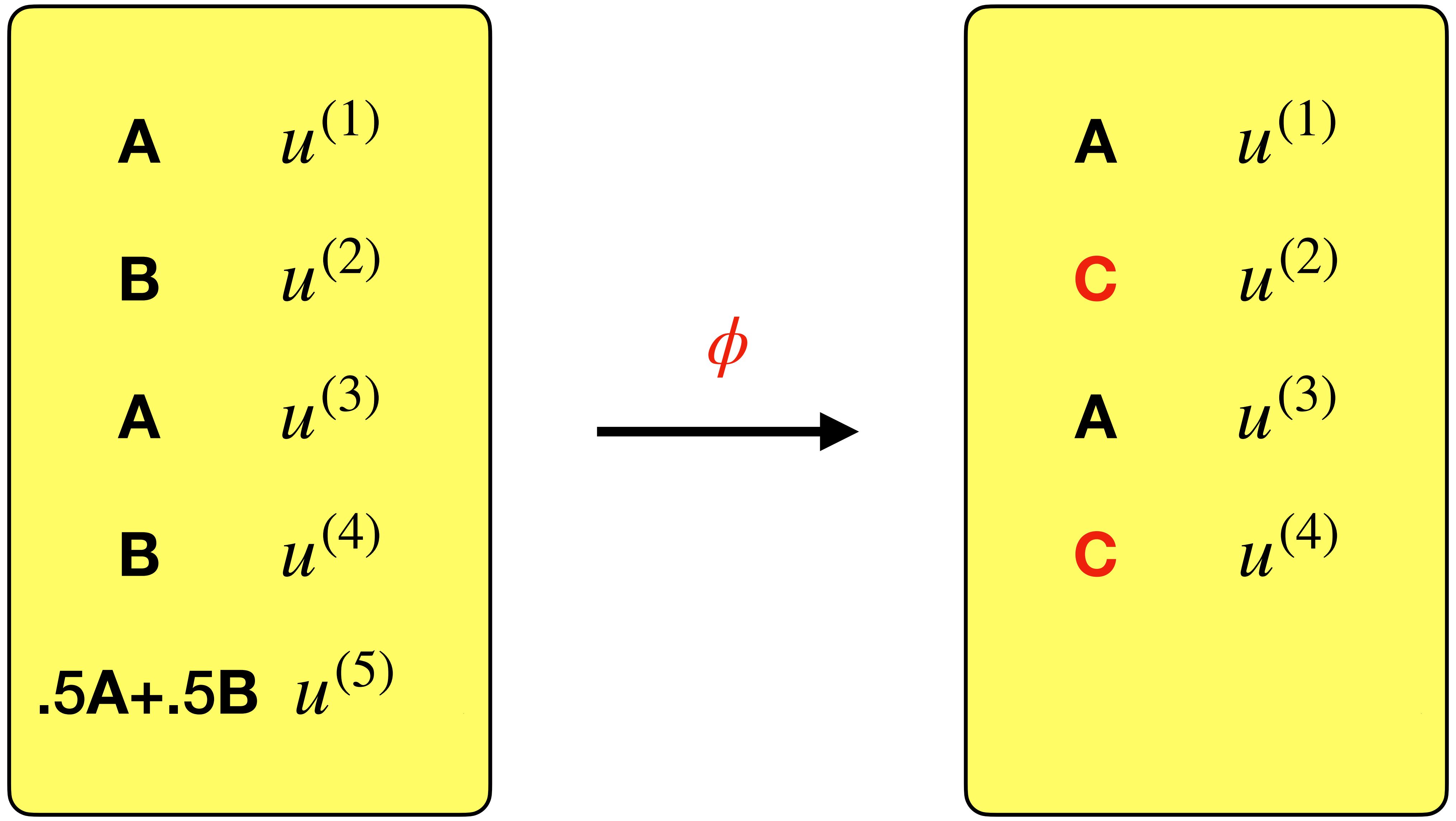


ϕ

A	A
B	C
C	C
⋮	⋮

Swap Regret

Φ = all maps $[N] \to [N]$



ϕ

A	A
B	C
C	C
⋮	⋮

Swap Regret

$\Phi =$ all maps $[N] \to [N]$

A	$u^{(1)}$
B	$u^{(2)}$
A	$u^{(3)}$
B	$u^{(4)}$
.5A+.5B	$u^{(5)}$

$\xrightarrow{\phi}$

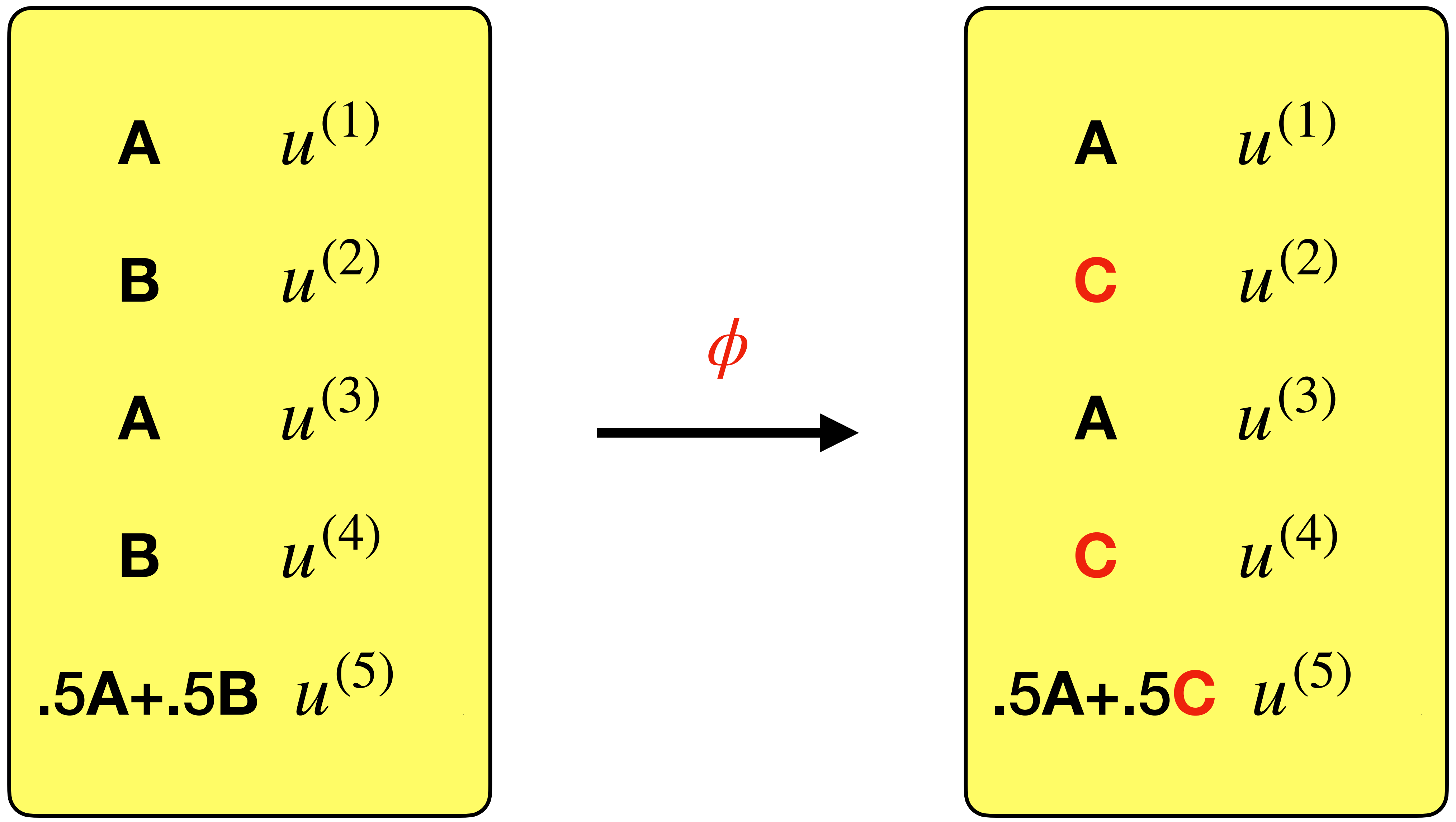
A	$u^{(1)}$
C	$u^{(2)}$
A	$u^{(3)}$
C	$u^{(4)}$
.5A+.5C	$u^{(5)}$

ϕ

A	A
B	C
C	C
⋮	⋮

Swap Regret

Φ = all maps $[N] \to [N]$



ϕ

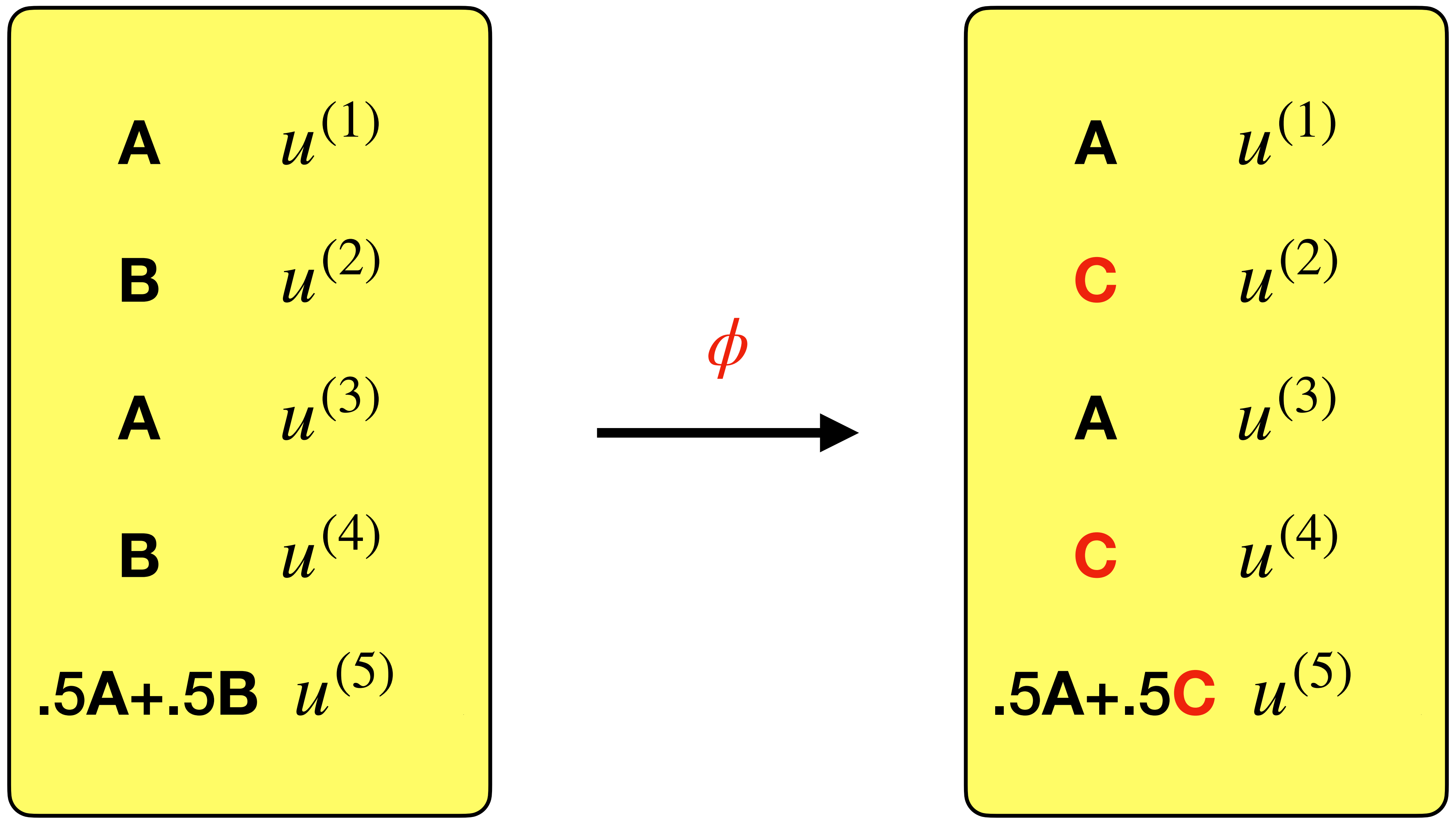
A	A
B	C
C	C
⋮	⋮

ϕ

A	D
B	D
C	D
⋮	⋮

Swap Regret

Φ = all maps $[N] \to [N]$



ϕ

A	A
B	C
C	C
⋮	⋮

ϕ

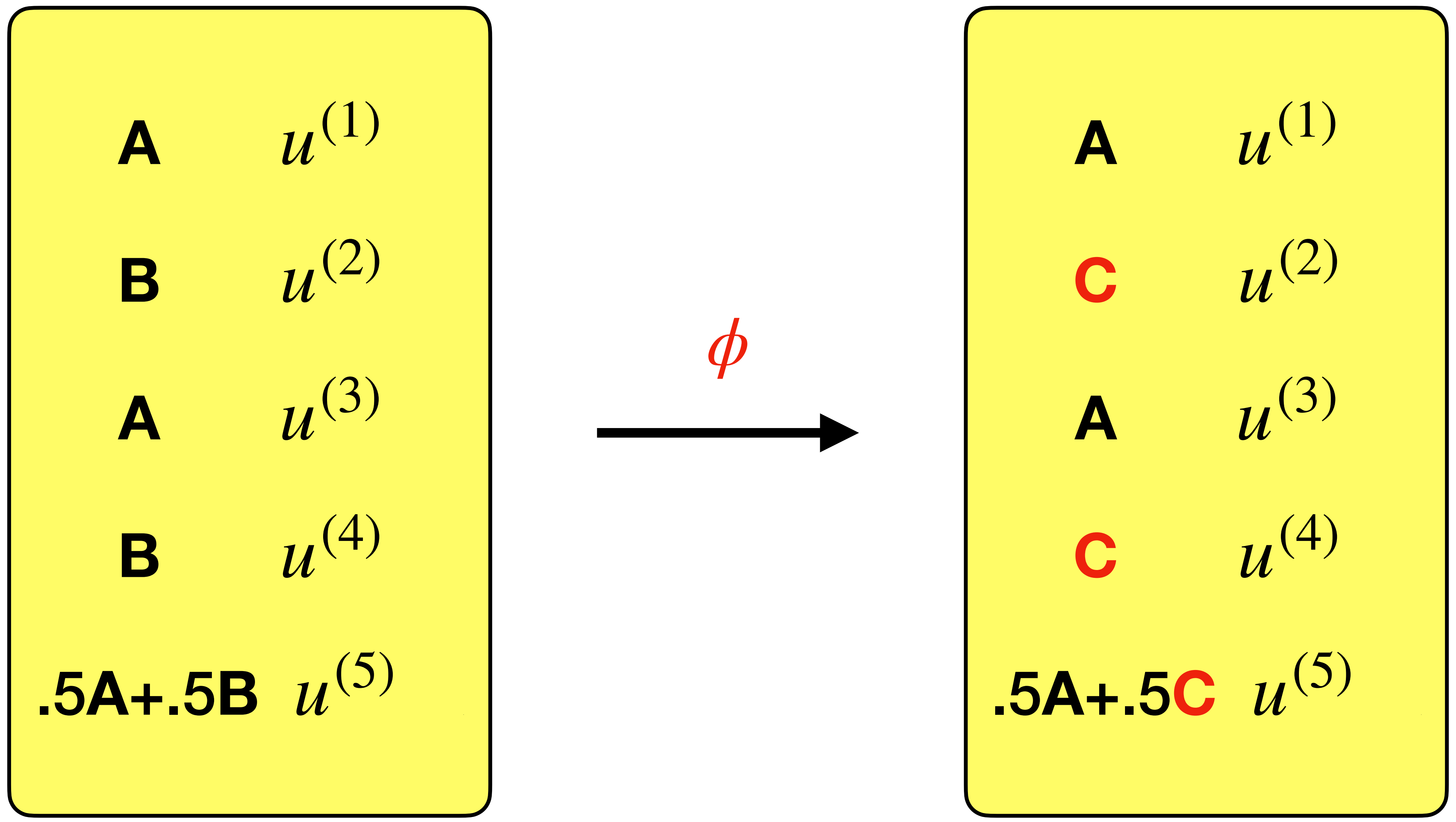
A	D
B	D
C	D
⋮	⋮

ϕ

A	X
B	Y
C	Z
⋮	⋮

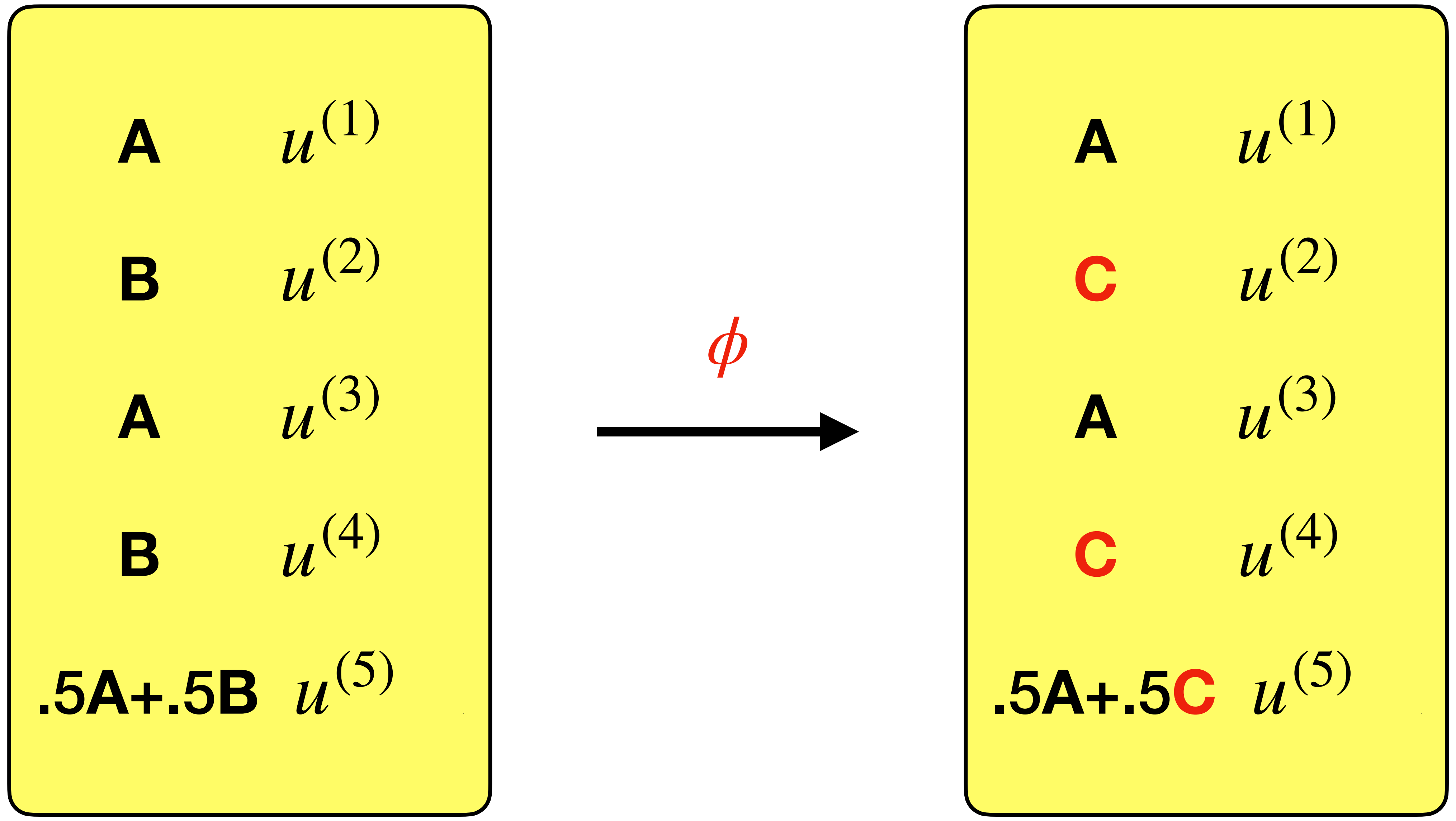
Swap Regret

Φ = all maps $[N] \to [N]$



Swap Regret

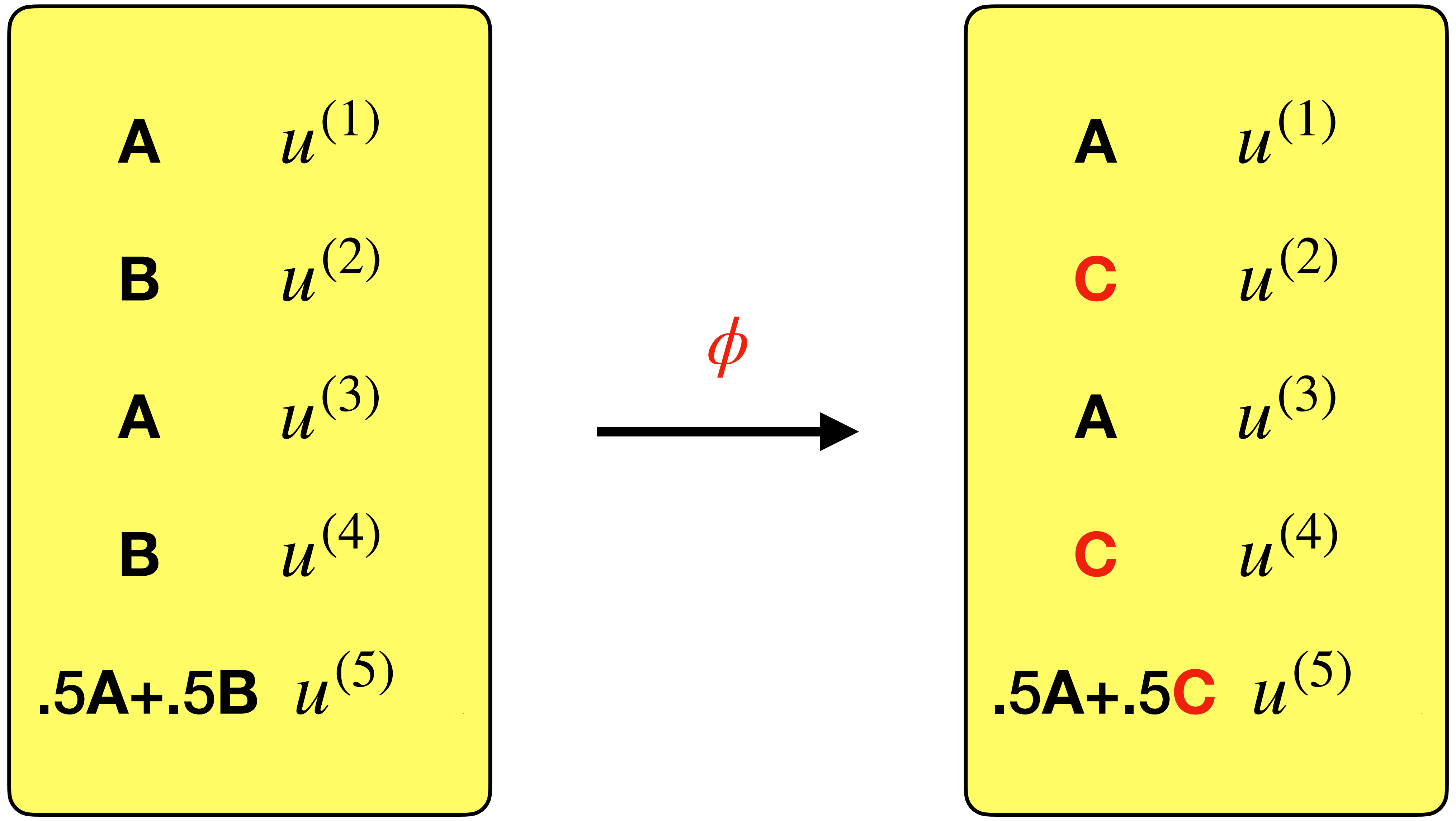
$\Phi =$ **all maps** $[N] \to [N]$



Why?

Swap Regret

$\Phi =$ all maps $[N] \to [N]$

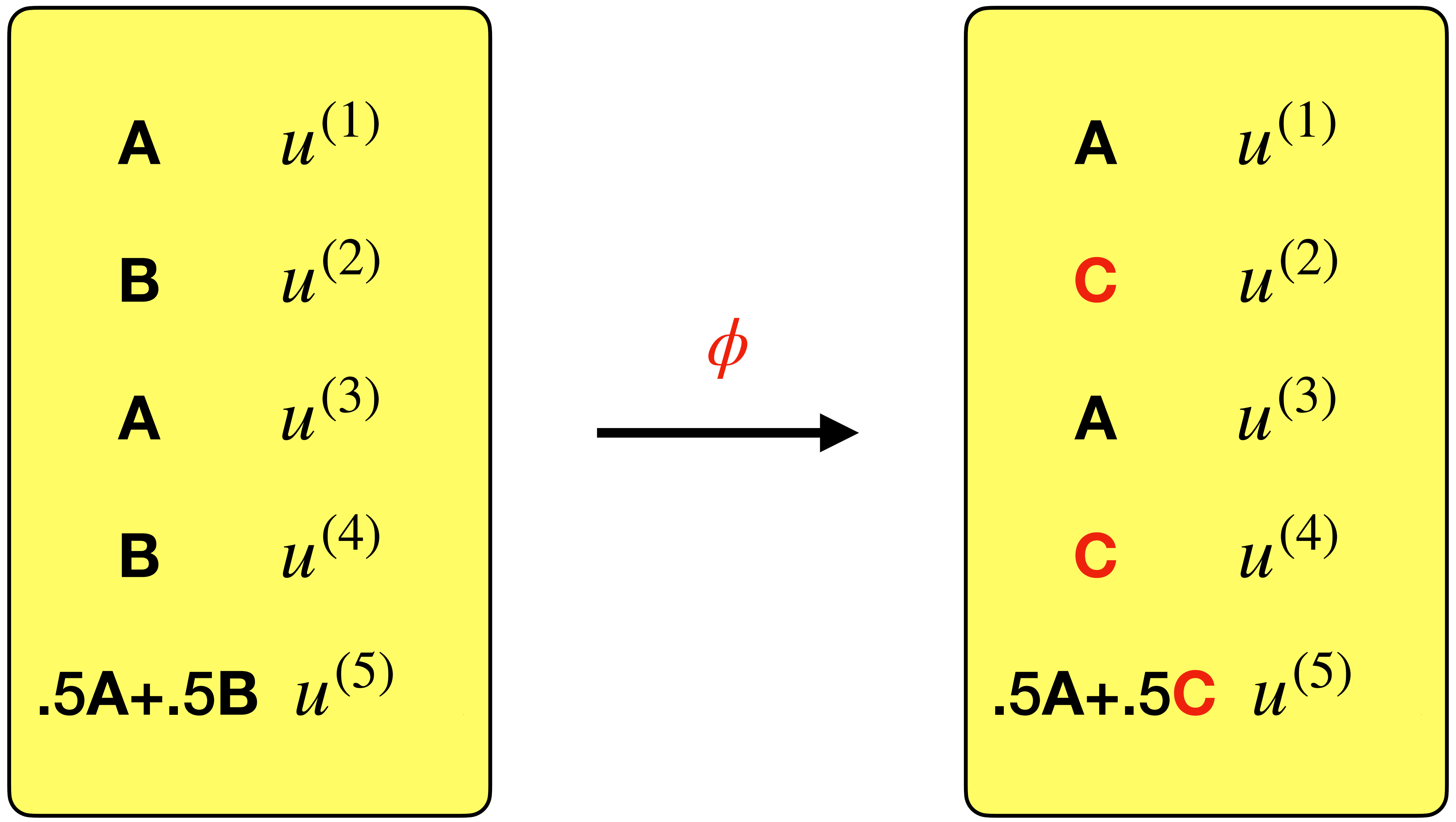


Why?

- L_2 Calibration

Swap Regret

$\Phi =$ **all maps** $[N] \to [N]$



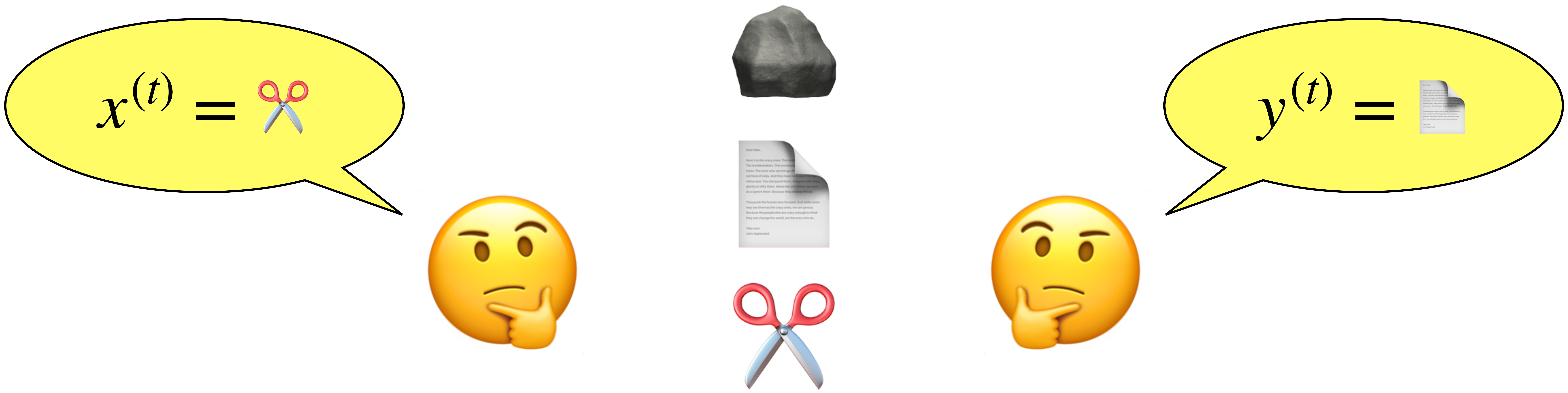
Why?

- L_2 Calibration
- Games!

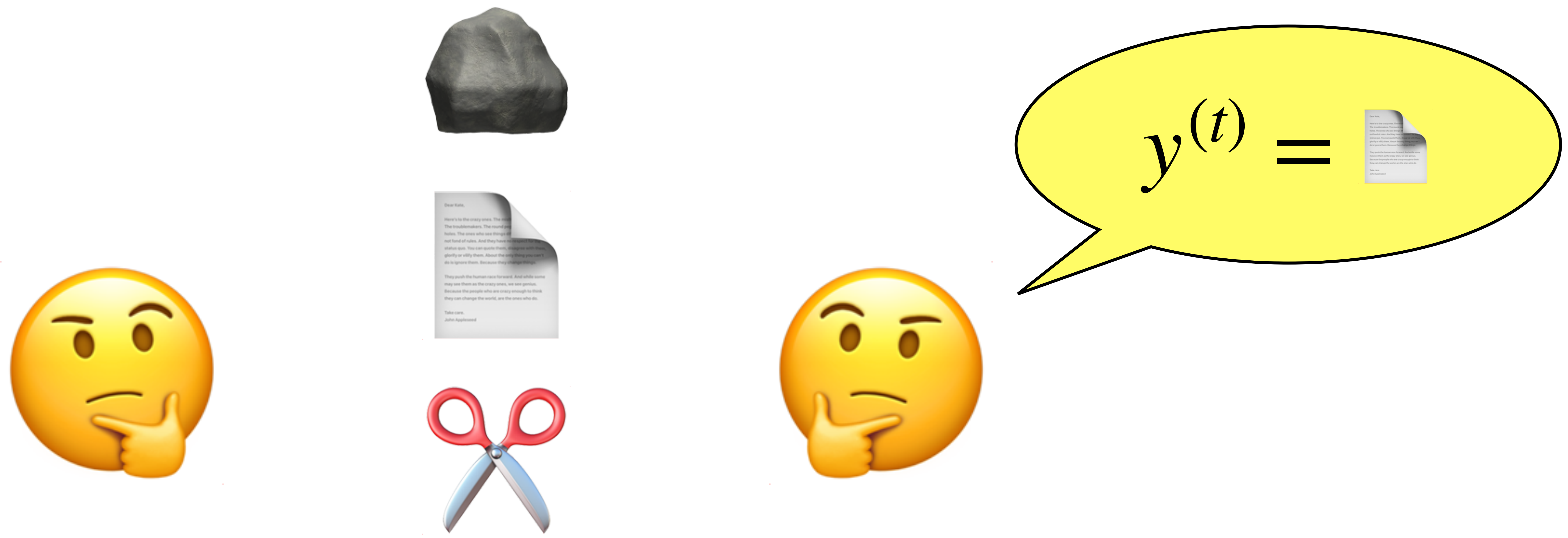
Learning in Games

Learning in Games

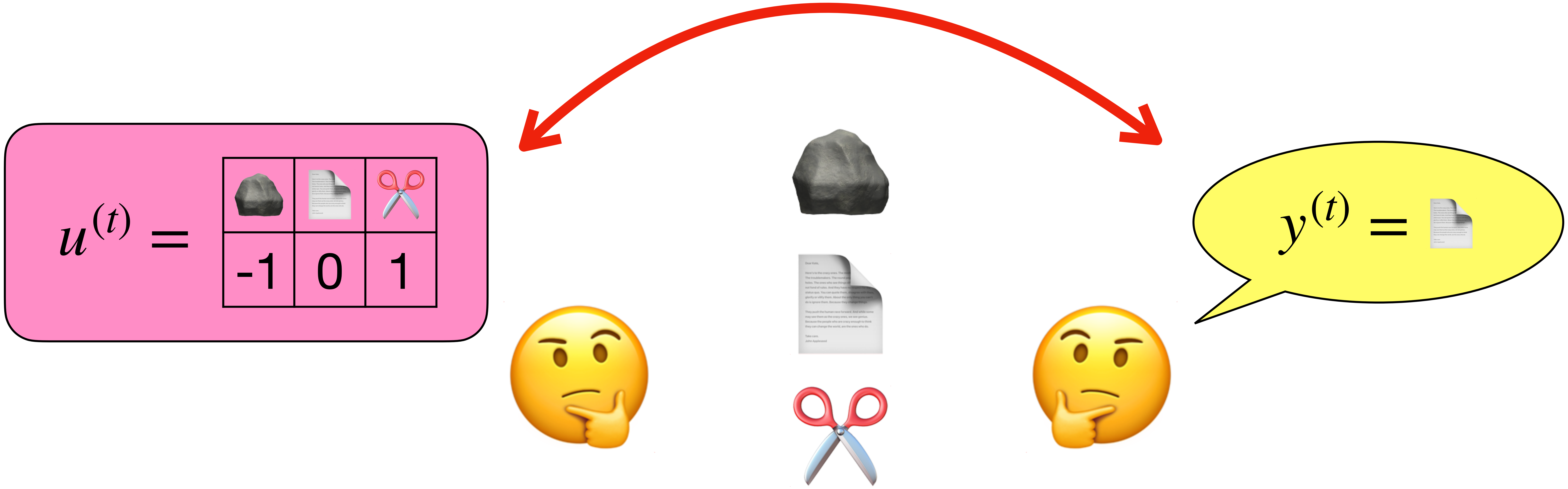
Learning in Games



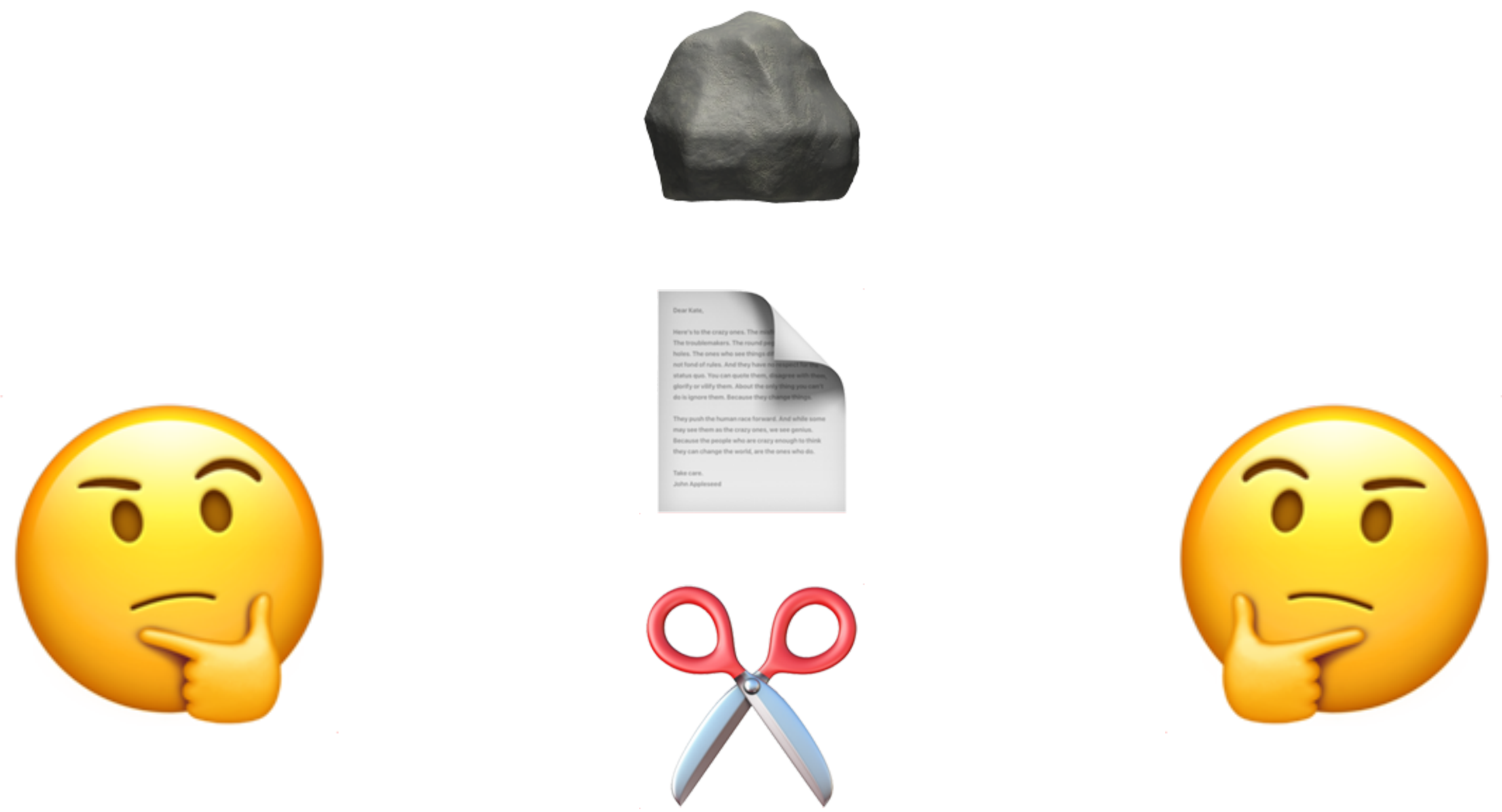
Learning in Games



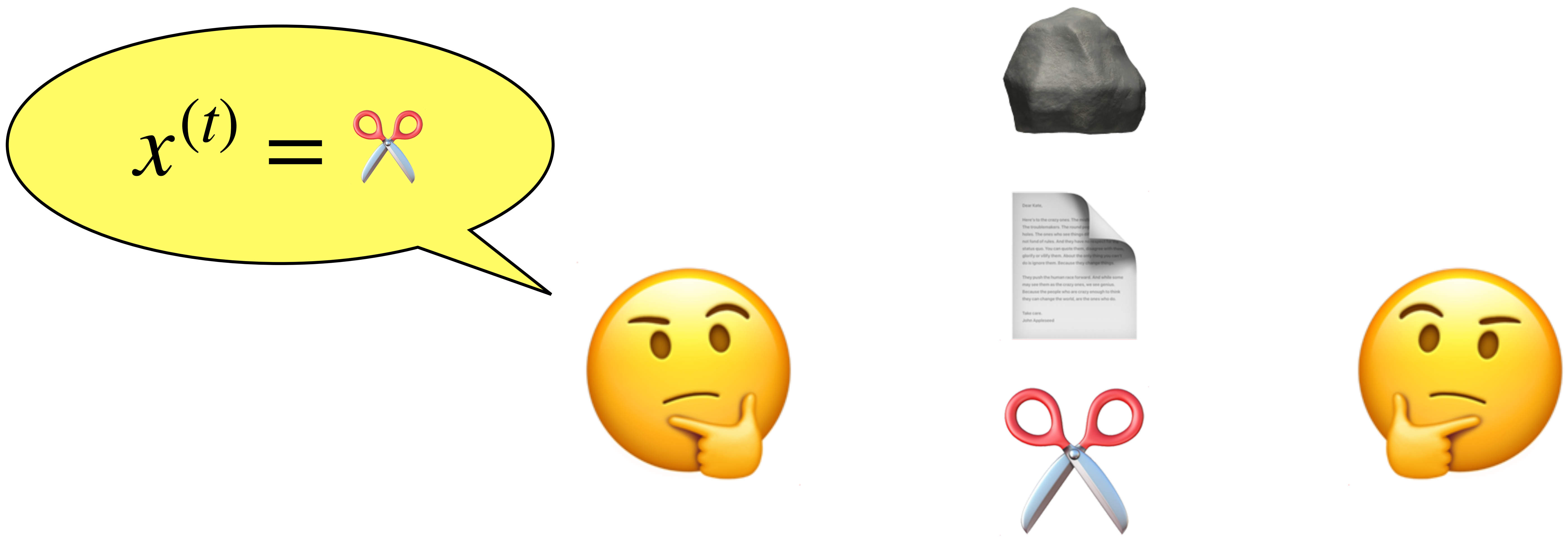
Learning in Games



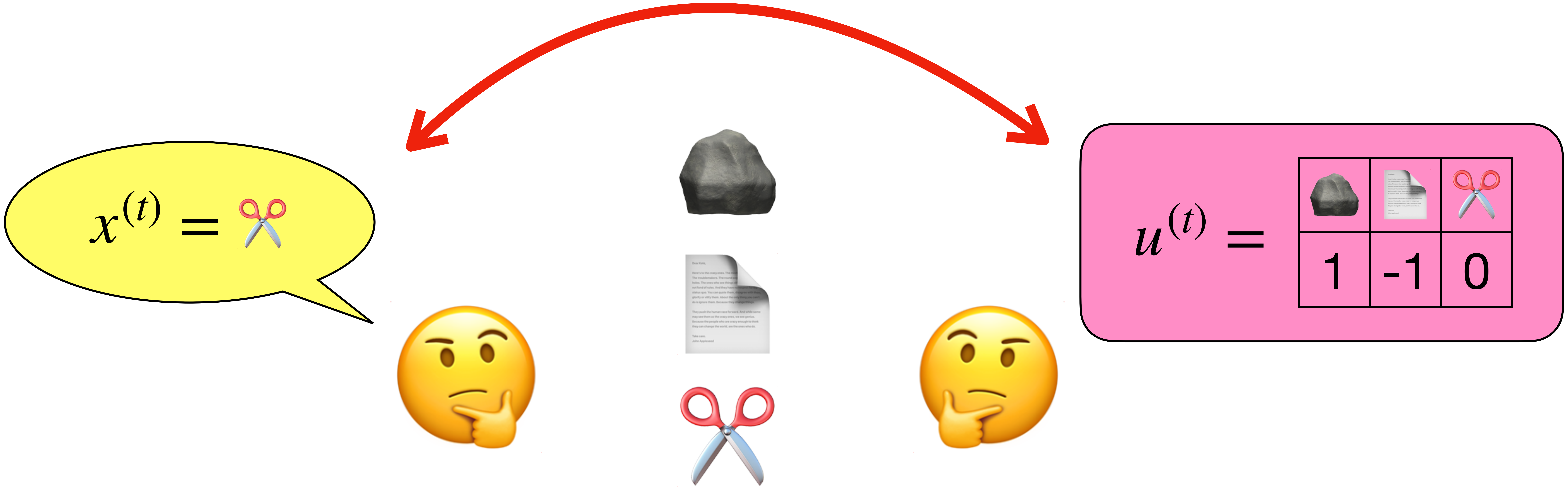
Learning in Games



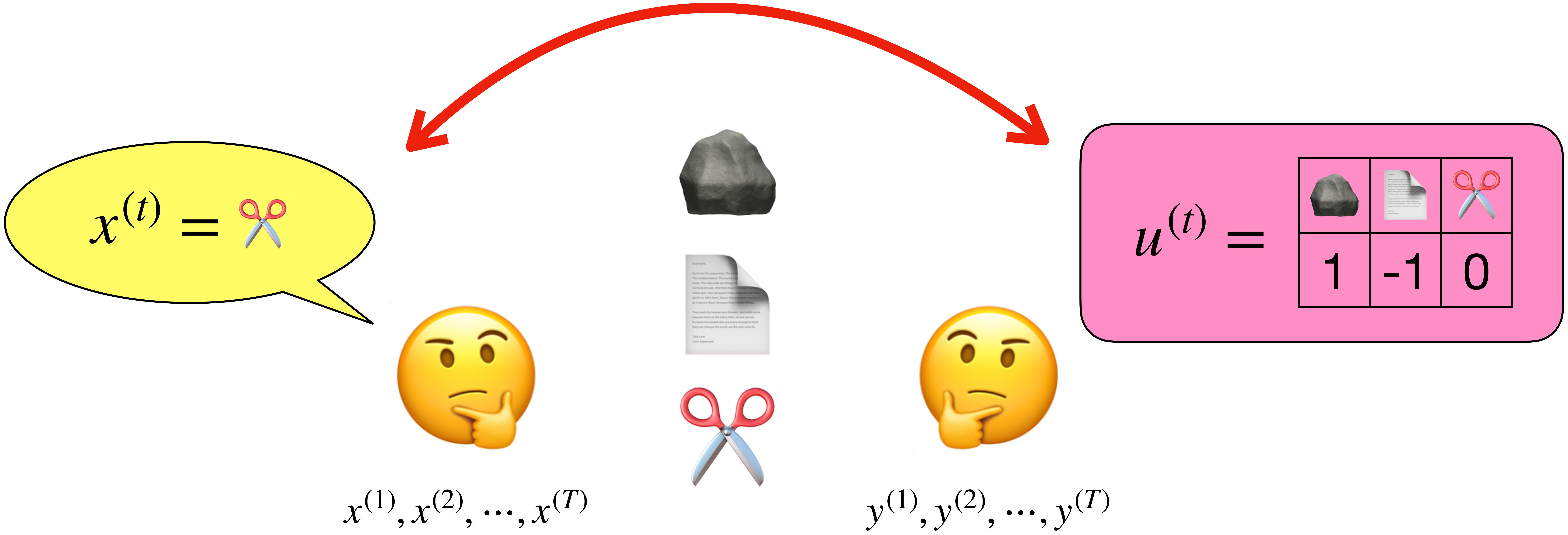
Learning in Games



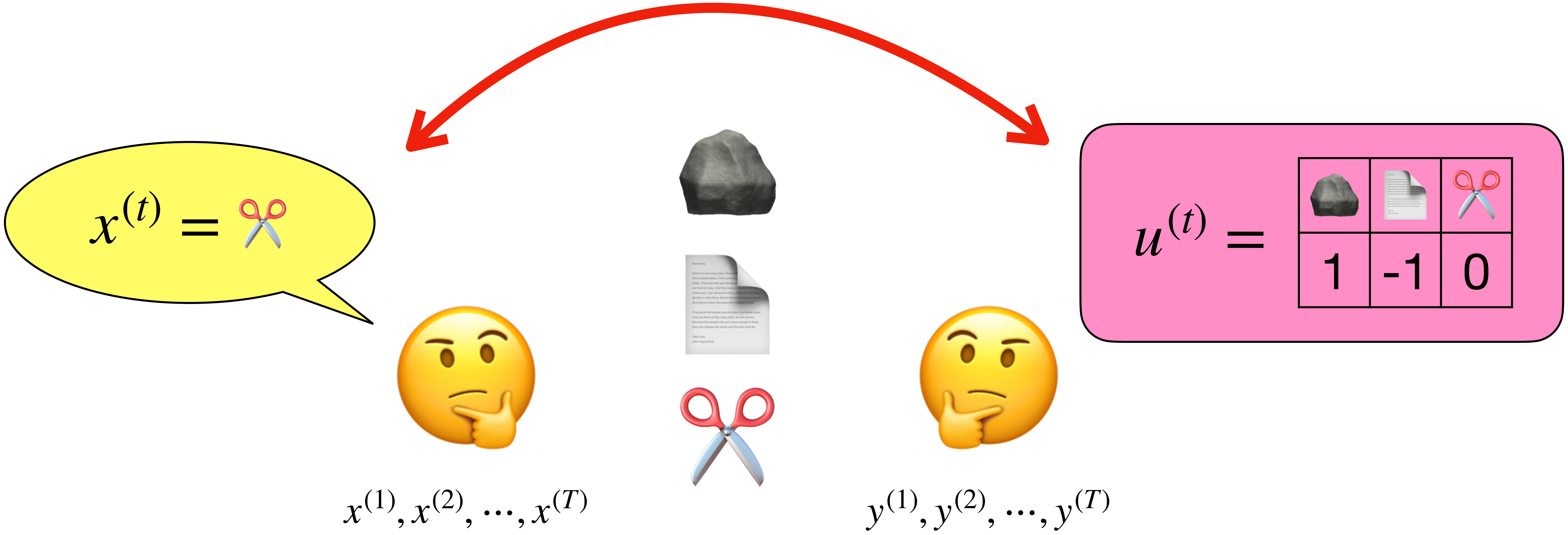
Learning in Games



Learning in Games

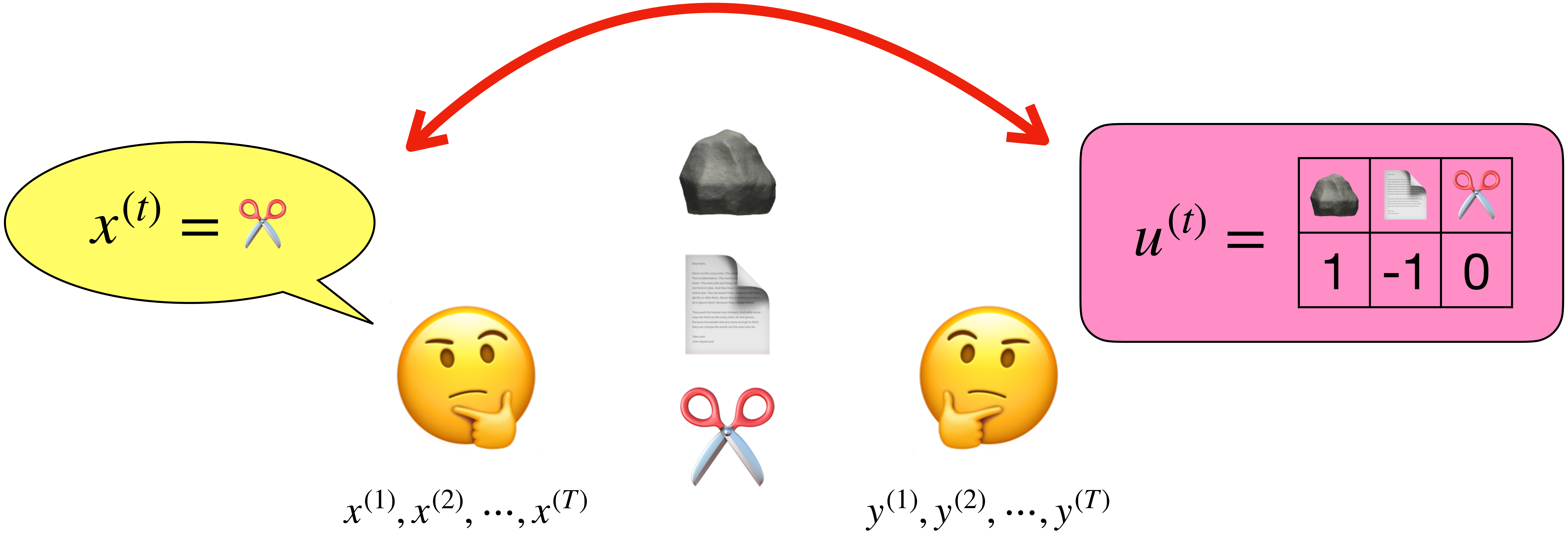


Learning in Games



ϵ external regret over T rounds $\rightarrow$ average strategies are ϵ coarse-correlated-equilibrium

Learning in Games



ϵ external regret over T rounds $\rightarrow$ average strategies are ϵ coarse-correlated-equilibrium

ϵ swap regret over T rounds $\rightarrow$ average strategies are ϵ correlated-equilibrium

Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate

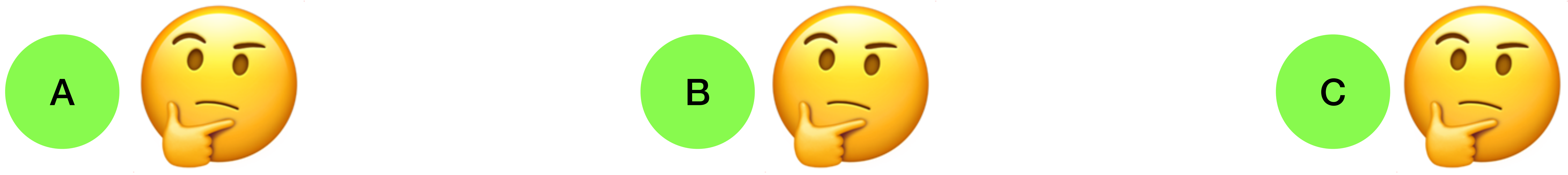
Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate

🤔 🤔 🤔

Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate



Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate

Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate

A 

B 

C  Z

Coarse Deviations

A	Z
B	Z
C	Z

Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate

Coarse Deviations

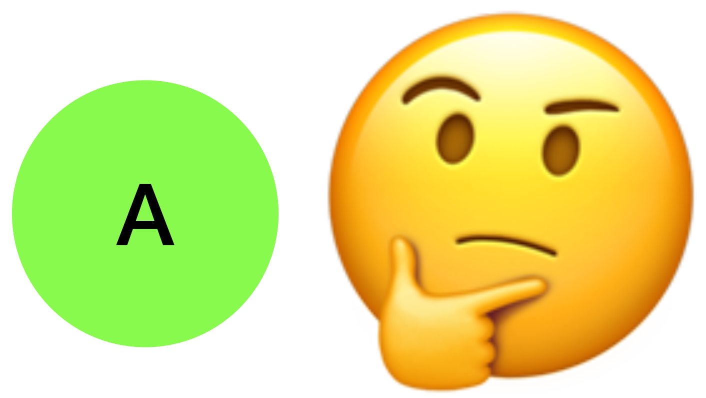
A	Z
B	Z
C	Z

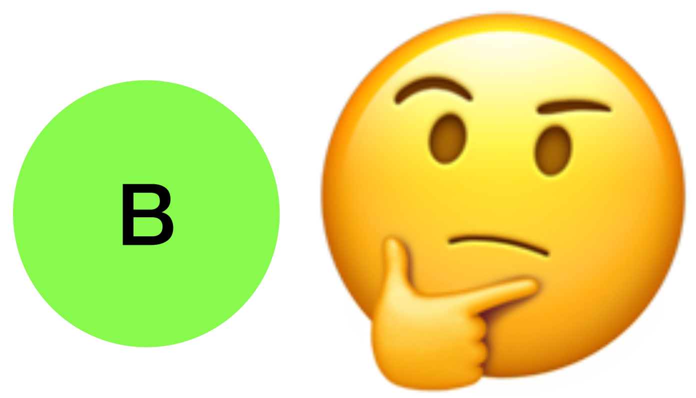
Swap Deviations

A	A
B	B
C	Z

Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate





Coarse Deviations

A	Z
B	Z
C	Z

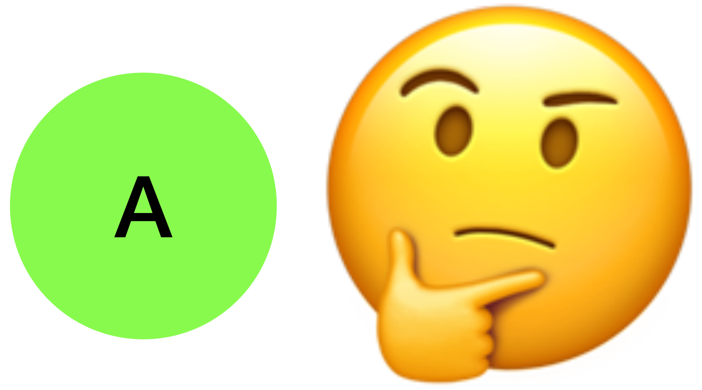
Swap Deviations

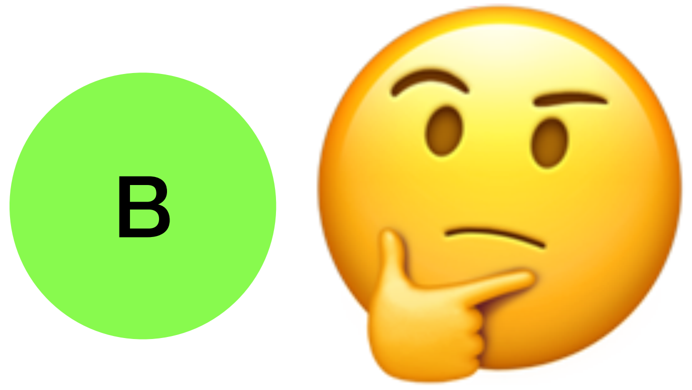
A	A
B	B
C	Z

CCE: No player wants to make a coarse deviation

Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate





Coarse Deviations

A	Z
B	Z
C	Z

Swap Deviations

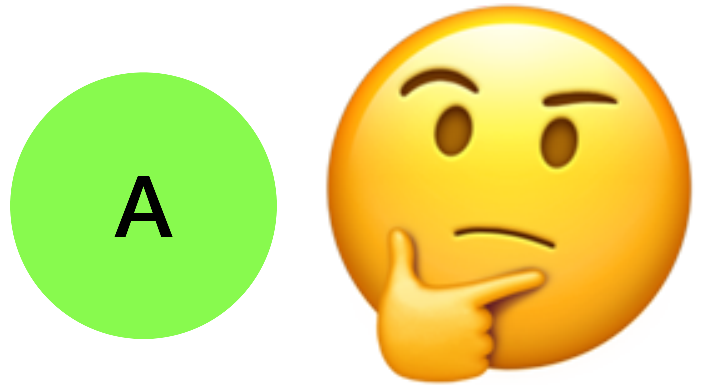
A	A
B	B
C	Z

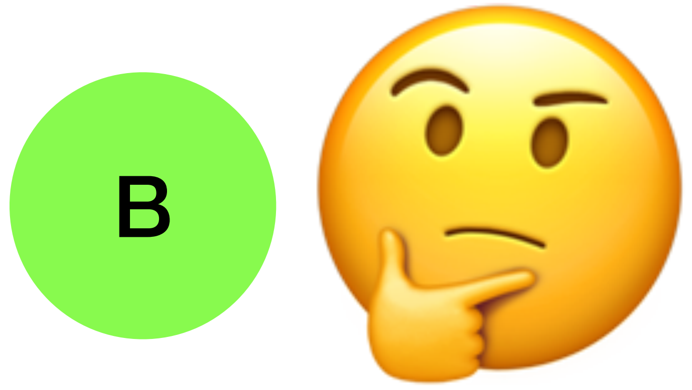
CCE: No player wants to make a coarse deviation

CE: No player wants to make a swap deviation

Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate





Coarse Deviations

A	Z
B	Z
C	Z

Swap Deviations

A	A
B	B
C	Z

CCE: No player wants to make a coarse deviation

CE: No player wants to make a swap deviation

NE: No player wants to make a swap deviation AND profile must be independent

Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate

Coarse Deviations

A	Z
B	Z
C	Z

Swap Deviations

A	A
B	B
C	Z

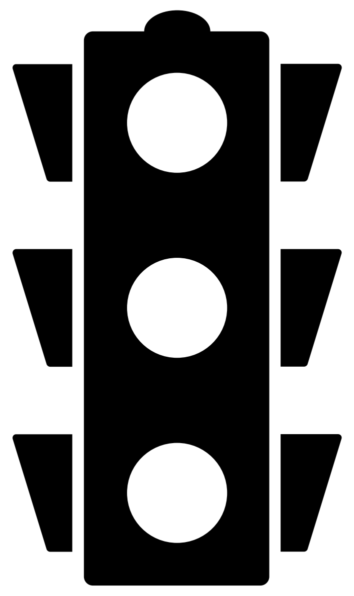
CCE: No player wants to make a coarse deviation

CE: No player wants to make a swap deviation

NE: No player wants to make a swap deviation AND profile must be independent

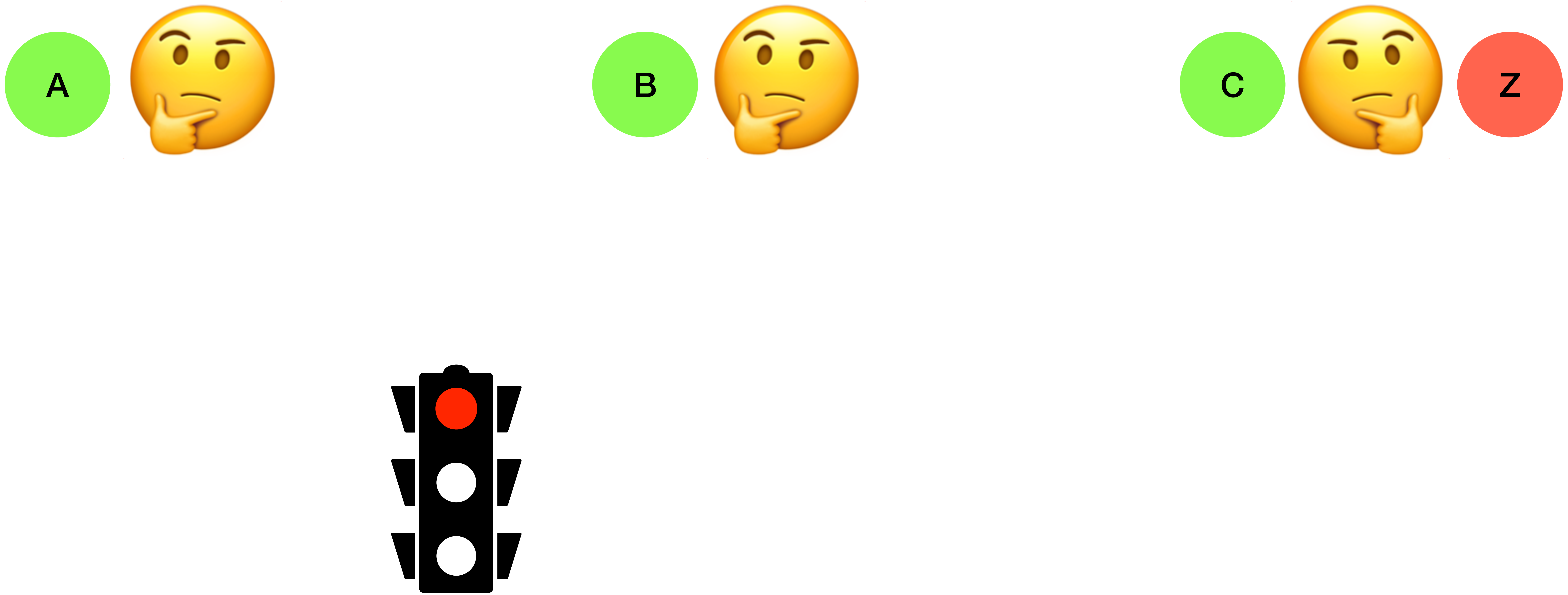
Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate



Learning in Games

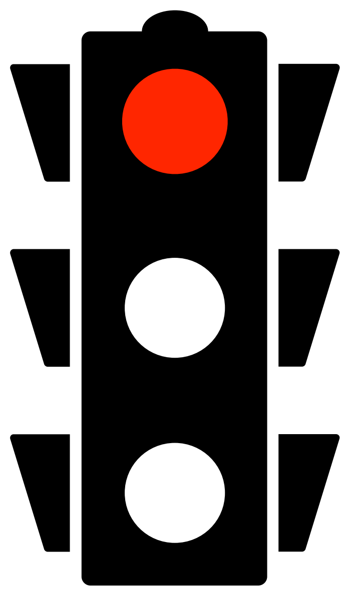
Game Equilibrium: Strategy profile where no player wants to deviate



Learning in Games

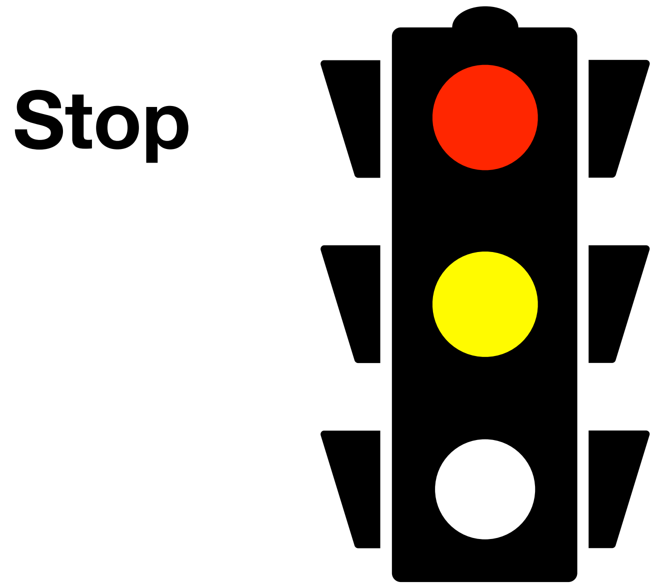
Game Equilibrium: Strategy profile where no player wants to deviate

Stop



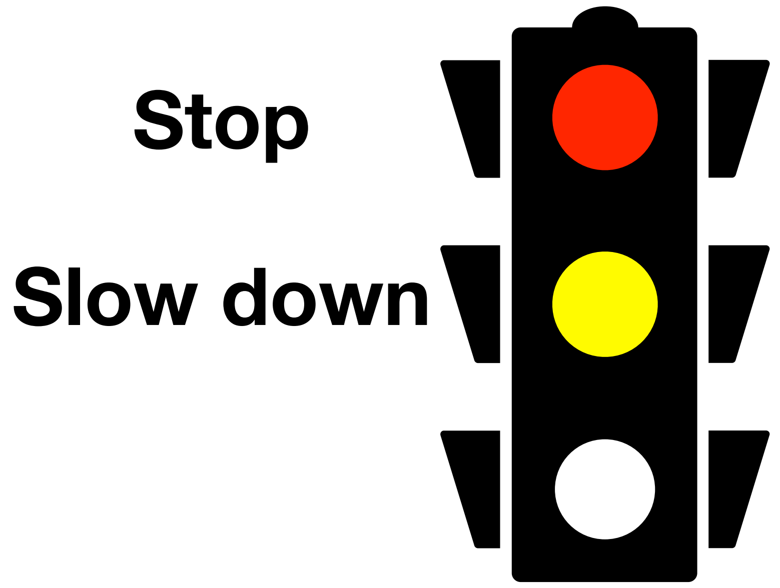
Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate



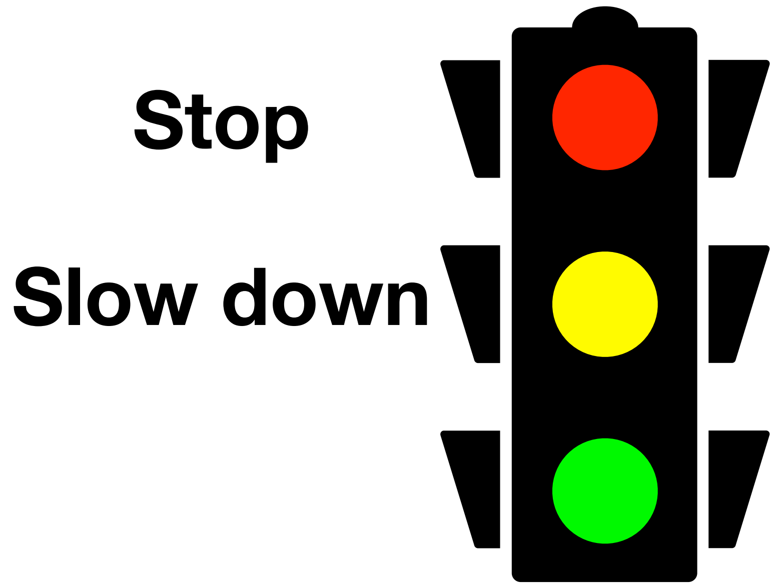
Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate



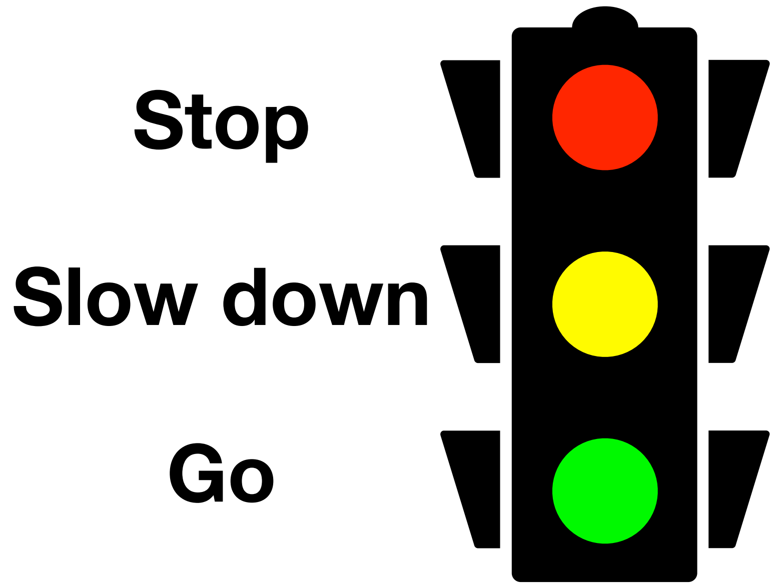
Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate



Learning in Games

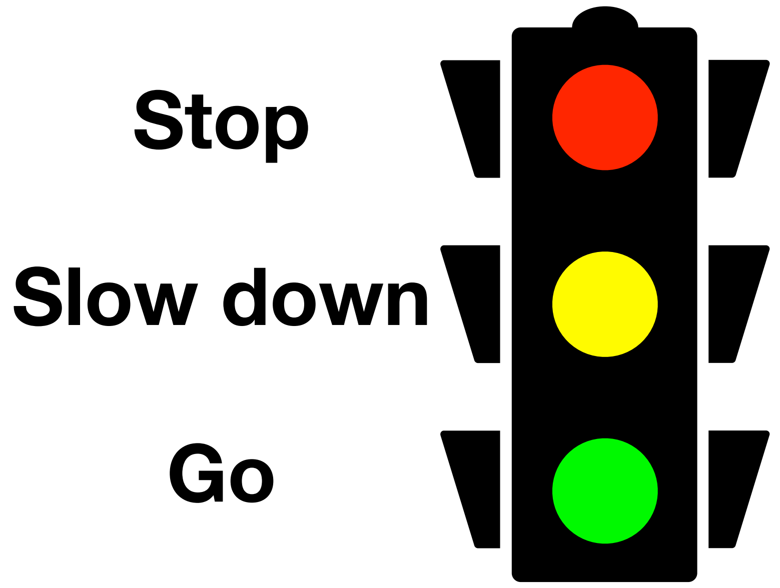
Game Equilibrium: Strategy profile where no player wants to deviate



Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate

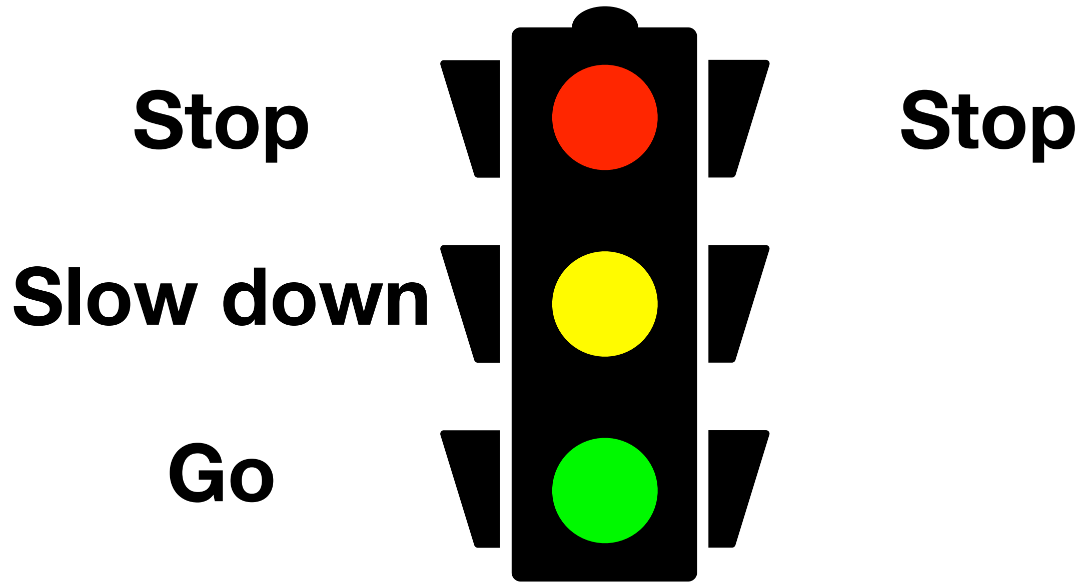
Swap Deviation



Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate

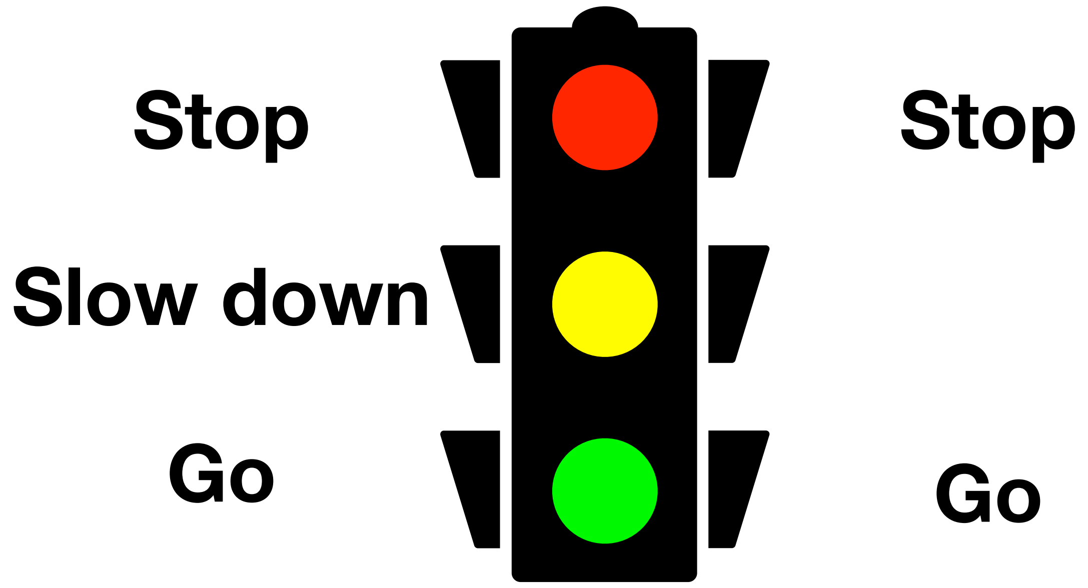
Swap Deviation



Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate

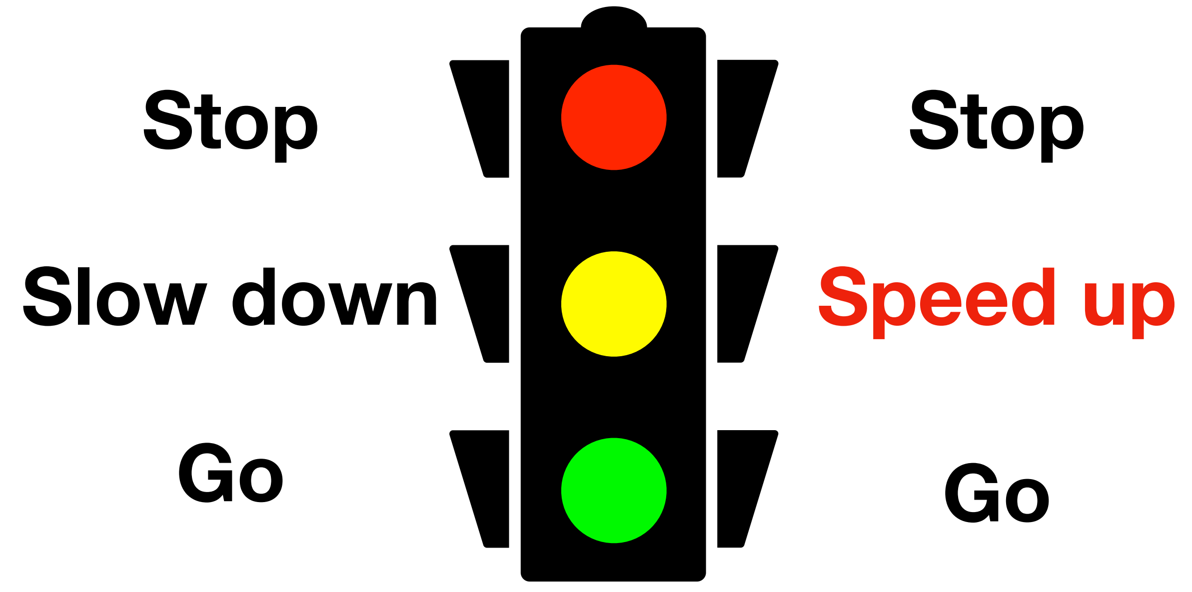
Swap Deviation



Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate

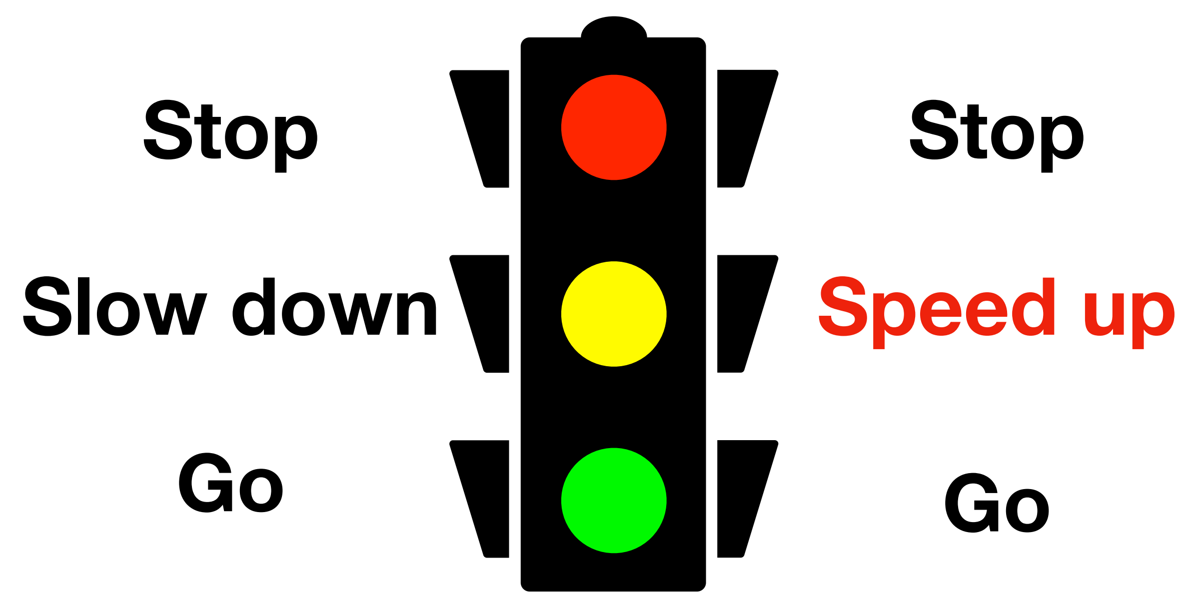
Swap Deviation



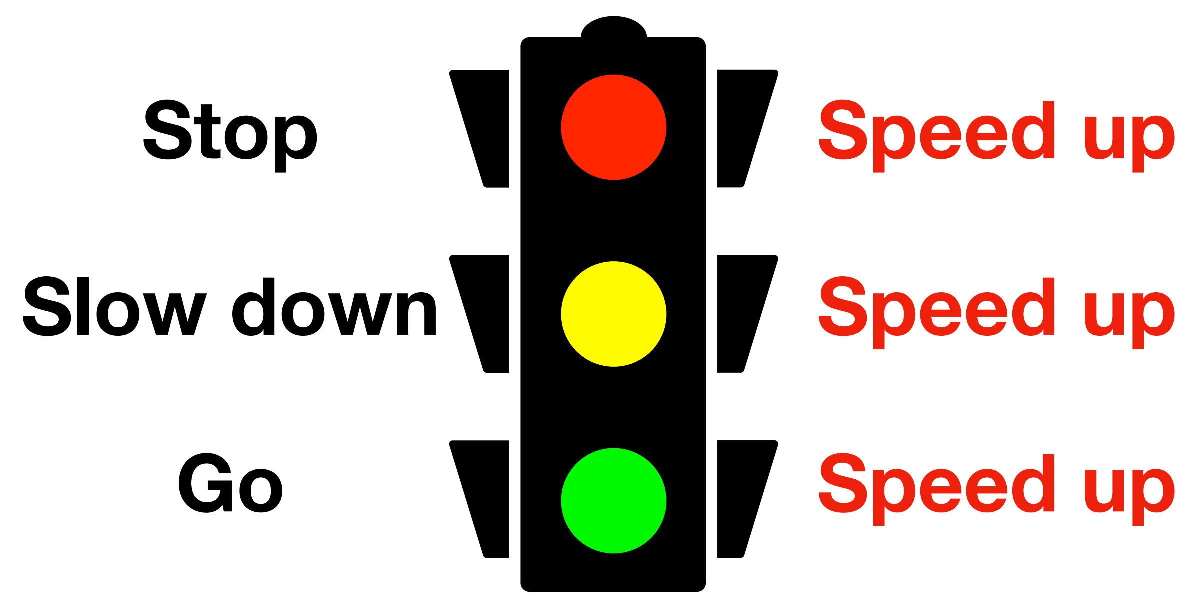
Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate

Swap Deviation



Coarse Deviation



So why not Swap Regret?

So why not Swap Regret?

Multiplicative Weight Updates: ϵ-External-Regret for $T = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

So why not Swap Regret?

Multiplicative Weight Updates: ϵ-External-Regret for $T = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

Blum-Mansour MWU [BM07]**:** ϵ-Swap-Regret for $T = \Omega\left(\frac{N \log N}{\epsilon^2}\right)$

So why not Swap Regret?

Multiplicative Weight Updates: ϵ-External-Regret for $T = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

Blum-Mansour MWU [BM07]**:** ϵ-Swap-Regret for $T = \Omega\left(\frac{N \log N}{\epsilon^2}\right)$

Question: can we improve for large N?

So why not Swap Regret?

Multiplicative Weight Updates: ϵ-External-Regret for $T = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

Blum-Mansour MWU [BM07]**:** ϵ-Swap-Regret for $T = \Omega\left(\frac{N \log N}{\epsilon^2}\right)$

Question: can we improve for large N? **YES** [DD🐟G23]

Main Results

Main Results

- **Tree-Swap:** ϵ-Swap-Regret for $T = \log(N)^{\tilde{\Omega}(1/\epsilon)}$

Main Results

- **Tree-Swap:** ϵ-Swap-Regret for $T = \log(N)^{\tilde{\Omega}(1/\epsilon)}$
- **Tree-Swap:** ϵ-Swap-Regret for $T = \text{Littlestone Dimension}(\mathcal{F})^{\tilde{\Omega}(1/\epsilon)}$

Main Results

- **Tree-Swap:** ϵ-Swap-Regret for $T = \log(N)^{\tilde{\Omega}(1/\epsilon)}$
- **Tree-Swap:** ϵ-Swap-Regret for $T = \text{Littlestone Dimension}(\mathcal{F})^{\tilde{\Omega}(1/\epsilon)}$
- **Bandit-Tree-Swap:** ϵ-Swap-Regret for $T = N\log(N)^{\tilde{\Omega}(1/\epsilon)}$

Main Results

- **Tree-Swap:** ϵ-Swap-Regret for $T = \log(N)^{\tilde{\Omega}(1/\epsilon)}$
- **Tree-Swap:** ϵ-Swap-Regret for $T = \text{Littlestone Dimension}(\mathscr{F})^{\tilde{\Omega}(1/\epsilon)}$
- **Bandit-Tree-Swap:** ϵ-Swap-Regret for $T = N\log(N)^{\tilde{\Omega}(1/\epsilon)}$

- **Communication Complexity of ϵ-CE:** $\log(N)^{\tilde{\Omega}(1/\epsilon)}$

Main Results

- **Tree-Swap:** ϵ-Swap-Regret for $T = \log(N)^{\tilde{\Omega}(1/\epsilon)}$
- **Tree-Swap:** ϵ-Swap-Regret for $T = \text{Littlestone Dimension}(\mathcal{F})^{\tilde{\Omega}(1/\epsilon)}$
- **Bandit-Tree-Swap:** ϵ-Swap-Regret for $T = N\log(N)^{\tilde{\Omega}(1/\epsilon)}$

- **Communication Complexity of ϵ-CE:** $\log(N)^{\tilde{\Omega}(1/\epsilon)}$
- **Query Complexity of ϵ-CE:** $N\log(N)^{\tilde{\Omega}(1/\epsilon)}$

Main Results

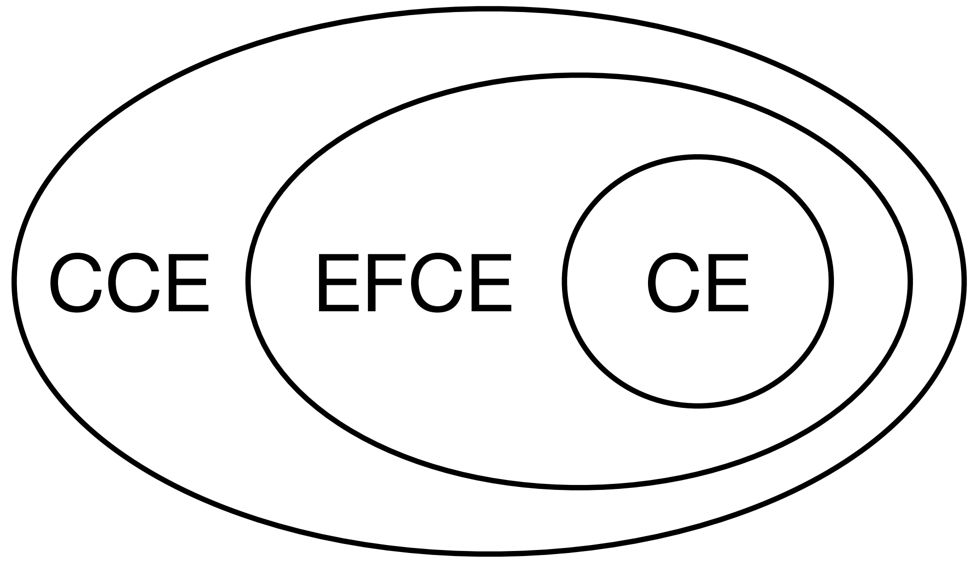
- **Tree-Swap:** ϵ-Swap-Regret for $T = \log(N)^{\tilde{\Omega}(1/\epsilon)}$
- **Tree-Swap:** ϵ-Swap-Regret for $T = \text{Littlestone Dimension}(\mathcal{F})^{\tilde{\Omega}(1/\epsilon)}$
- **Bandit-Tree-Swap:** ϵ-Swap-Regret for $T = N\log(N)^{\tilde{\Omega}(1/\epsilon)}$

- **Communication Complexity of ϵ-CE:** $\log(N)^{\tilde{\Omega}(1/\epsilon)}$
- **Query Complexity of ϵ-CE:** $N\log(N)^{\tilde{\Omega}(1/\epsilon)}$
- **Extensive Form ϵ-CE:** poly(m, I, A) for $\epsilon = \Omega(1)$

Main Results

- **Tree-Swap:** ϵ-Swap-Regret for $T = \log(N)^{\tilde{\Omega}(1/\epsilon)}$
- **Tree-Swap:** ϵ-Swap-Regret for $T = \text{Littlestone Dimension}(\mathcal{F})^{\tilde{\Omega}(1/\epsilon)}$
- **Bandit-Tree-Swap:** ϵ-Swap-Regret for $T = N\log(N)^{\tilde{\Omega}(1/\epsilon)}$

- **Communication Complexity of ϵ-CE:** $\log(N)^{\tilde{\Omega}(1/\epsilon)}$
- **Query Complexity of ϵ-CE:** $N\log(N)^{\tilde{\Omega}(1/\epsilon)}$
- **Extensive Form ϵ-CE:** poly(m, I, A) for $\epsilon = \Omega(1)$



Main Results

Main Results

Lower Bounds

Main Results

Lower Bounds

- **Tree-Adversary:** Force learner to incur Swap Regret $\geq \epsilon$ for $T = O(1) \cdot \min \left\{ \exp(\epsilon^{-1/6}), \frac{N}{\text{poly}\log(N) \cdot \epsilon^2} \right\}$

Main Results

Lower Bounds

- **Tree-Adversary:** Force learner to incur Swap Regret $\geq \epsilon$ for $T = O(1) \cdot \min \left\{ \exp(\epsilon^{-1/6}), \frac{N}{\operatorname{poly}\log(N) \cdot \epsilon^2} \right\}$
- Oblivious

Main Results

Lower Bounds

- **Tree-Adversary:** Force learner to incur Swap Regret $\geq \epsilon$ for $T = O(1) \cdot \min \left\{ \exp(\epsilon^{-1/6}), \frac{N}{\operatorname{poly}\log(N) \cdot \epsilon^2} \right\}$
- Oblivious
- Point functions: Littlestone Dimension 1

Main Results

Lower Bounds

- **Tree-Adversary:** Force learner to incur Swap Regret $\geq \epsilon$ for $T = O(1) \cdot \min \left\{ \exp(\epsilon^{-1/6}), \frac{N}{\text{poly}\log(N) \cdot \epsilon^2} \right\}$
- Oblivious
- Point functions: Littlestone Dimension 1
- **Slide-Adversary:** Force learner to incur Swap Regret $\geq \epsilon$ for $T = \exp(O(1) \cdot \epsilon^{-1/3})$

Algorithms

External Regret

External Regret

Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity)

External Regret

Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity)

$$
\begin{array}{r}
u^{(1)} \\
u^{(2)} \\
\vdots \\
+u^{(t-1)} \\
\hline
\end{array}
$$

External Regret

Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity)

$$
\begin{array}{r}
u^{(1)} \\
u^{(2)} \\
\vdots \\
+u^{(t-1)} \\
\hline
U^{(t-1)}
\end{array}
$$

External Regret

Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity)

$$
\begin{array}{r}
u^{(1)} \\
u^{(2)} \\
\vdots \\
+u^{(t-1)} \\
\hline
U^{(t-1)}
\end{array} = [60, 120, 90]
$$

External Regret

Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity)

$$\begin{array}{r} u^{(1)} \\ u^{(2)} \\ \vdots \\ +u^{(t-1)} \\ \hline U^{(t-1)} \end{array} = [60,120,90]$$

$$x^{(t)} = \text{Softmax}(\eta U^{(t-1)})$$

External Regret

Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity)

$$\begin{array}{r} u^{(1)} \\ u^{(2)} \\ \vdots \\ +u^{(t-1)} \\ \hline U^{(t-1)} \end{array} = [60,120,90]$$

$$x^{(t)} = \text{Softmax}(\eta U^{(t-1)})$$
$$= [.002,.95,.048]$$

External Regret

Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity)

External Regret

Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity)

Rewards	Actions

External Regret

Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity)

Rewards	Actions
$U^{(0)}$	

External Regret

Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity)

Rewards	Actions
$U^{(0)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$

External Regret

Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity)

Rewards	Actions
$U^{(0)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(0)} + u^{(1)} = U^{(1)}$	

External Regret

Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity)

Rewards	Actions
$U^{(0)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(0)} + u^{(1)} = U^{(1)}$	$x^{(2)} = \text{Softmax}(\eta U^{(1)})$

External Regret

Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity)

Rewards	Actions
$U^{(0)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(0)} + u^{(1)} = U^{(1)}$	$x^{(2)} = \text{Softmax}(\eta U^{(1)})$
$U^{(1)} + u^{(2)} = U^{(2)}$	

External Regret

Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity)

Rewards	Actions
$U^{(0)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(0)} + u^{(1)} = U^{(1)}$	$x^{(2)} = \text{Softmax}(\eta U^{(1)})$
$U^{(1)} + u^{(2)} = U^{(2)}$	$x^{(3)} = \text{Softmax}(\eta U^{(2)})$

External Regret

Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity)

Rewards	Actions
$U^{(0)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(0)} + u^{(1)} = U^{(1)}$	$x^{(2)} = \text{Softmax}(\eta U^{(1)})$
$U^{(1)} + u^{(2)} = U^{(2)}$	$x^{(3)} = \text{Softmax}(\eta U^{(2)})$
$U^{(2)} + u^{(3)} = U^{(3)}$	

External Regret

Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity)

Rewards	Actions
$U^{(0)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(0)} + u^{(1)} = U^{(1)}$	$x^{(2)} = \text{Softmax}(\eta U^{(1)})$
$U^{(1)} + u^{(2)} = U^{(2)}$	$x^{(3)} = \text{Softmax}(\eta U^{(2)})$
$U^{(2)} + u^{(3)} = U^{(3)}$	$x^{(4)} = \text{Softmax}(\eta U^{(3)})$

External Regret

Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity)

Rewards	Actions
$U^{(0)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(0)} + u^{(1)} = U^{(1)}$	$x^{(2)} = \text{Softmax}(\eta U^{(1)})$
$U^{(1)} + u^{(2)} = U^{(2)}$	$x^{(3)} = \text{Softmax}(\eta U^{(2)})$
$U^{(2)} + u^{(3)} = U^{(3)}$	$x^{(4)} = \text{Softmax}(\eta U^{(3)})$
$U^{(3)} + u^{(4)} = U^{(4)}$	

External Regret

Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity)

Rewards	Actions
$U^{(0)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(0)} + u^{(1)} = U^{(1)}$	$x^{(2)} = \text{Softmax}(\eta U^{(1)})$
$U^{(1)} + u^{(2)} = U^{(2)}$	$x^{(3)} = \text{Softmax}(\eta U^{(2)})$
$U^{(2)} + u^{(3)} = U^{(3)}$	$x^{(4)} = \text{Softmax}(\eta U^{(3)})$
$U^{(3)} + u^{(4)} = U^{(4)}$	$x^{(5)} = \text{Softmax}(\eta U^{(4)})$

External Regret

Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity)

Rewards	Actions

External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity)

Rewards	Actions

External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**

Rewards	Actions

External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**

Rewards	Actions
$U^{(0)}$	

External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**

Rewards	Actions
$U^{(0)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$

External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**

Rewards	Actions
$U^{(0)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(0)} + u^{(1)} = U^{(1)}$	

External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**

Rewards	Actions
$U^{(0)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(0)} + u^{(1)} = U^{(1)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$

External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**

Rewards	Actions
$U^{(0)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(0)} + u^{(1)} = U^{(1)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(1)} + u^{(2)} = U^{(2)}$	

External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**

Rewards	Actions
$U^{(0)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(0)} + u^{(1)} = U^{(1)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(1)} + u^{(2)} = U^{(2)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$

External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**

Rewards	Actions
$U^{(0)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(0)} + u^{(1)} = U^{(1)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(1)} + u^{(2)} = U^{(2)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(2)} + u^{(3)} = U^{(3)}$	

External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**

Rewards	Actions
$U^{(0)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(0)} + u^{(1)} = U^{(1)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(1)} + u^{(2)} = U^{(2)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(2)} + u^{(3)} = U^{(3)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$

External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**

Rewards	Actions
$U^{(0)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(0)} + u^{(1)} = U^{(1)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(1)} + u^{(2)} = U^{(2)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(2)} + u^{(3)} = U^{(3)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(3)} + u^{(4)} = U^{(4)}$	

External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**

Rewards	Actions
$U^{(0)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(0)} + u^{(1)} = U^{(1)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(1)} + u^{(2)} = U^{(2)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(2)} + u^{(3)} = U^{(3)}$	$x^{(1)} = \text{Softmax}(\eta U^{(0)})$
$U^{(3)} + u^{(4)} = U^{(4)}$	$x^{(5)} = \text{Softmax}(\eta U^{(4)})$

External Regret

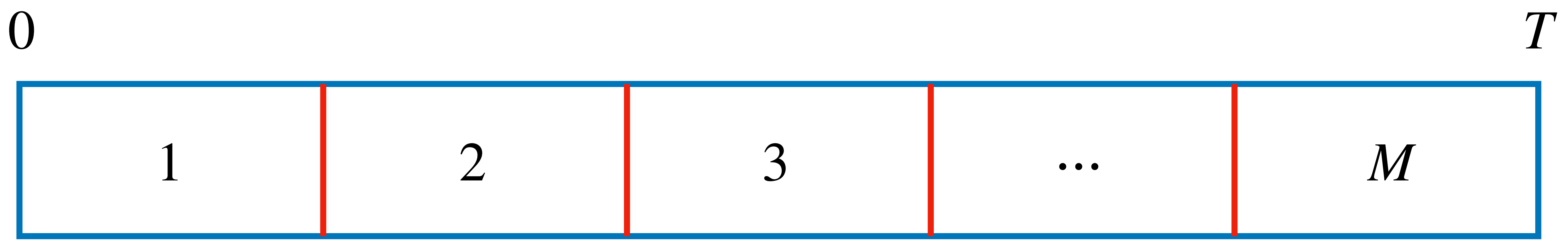
Lazy **Multiplicative Weight Updates:** Look at the history, do the best thing (with some regularity) **every few rounds**

External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**

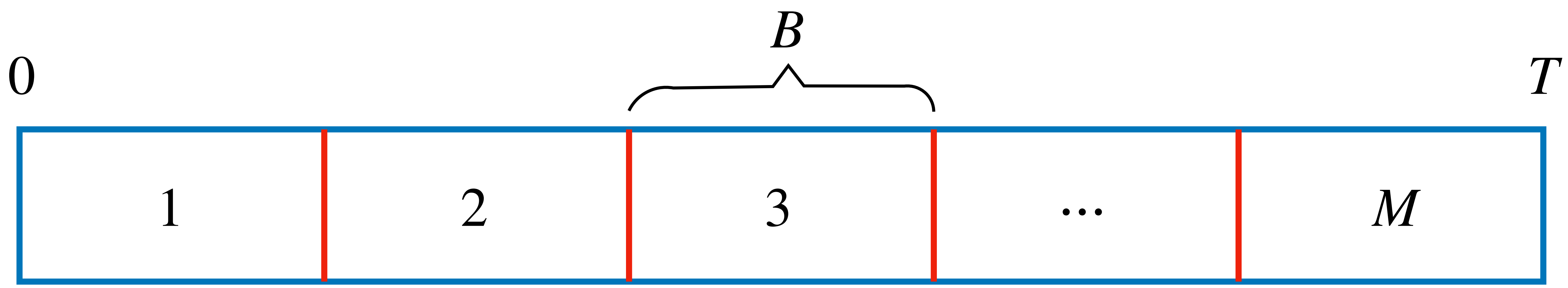
External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



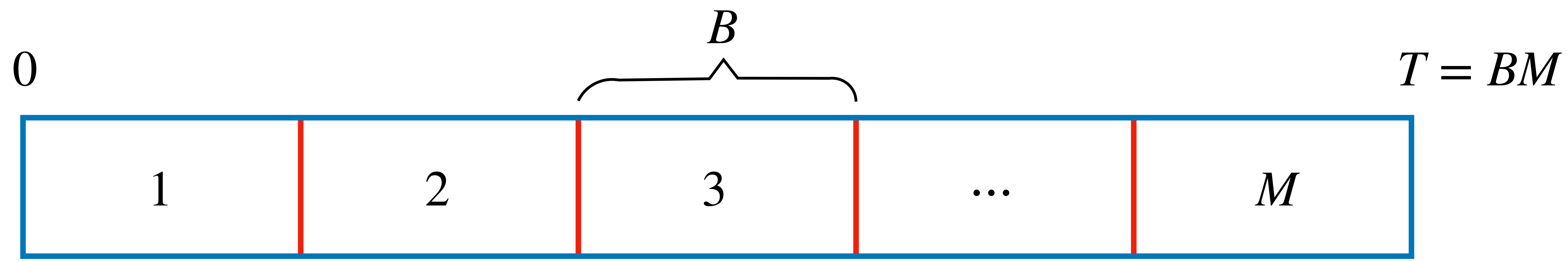
External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



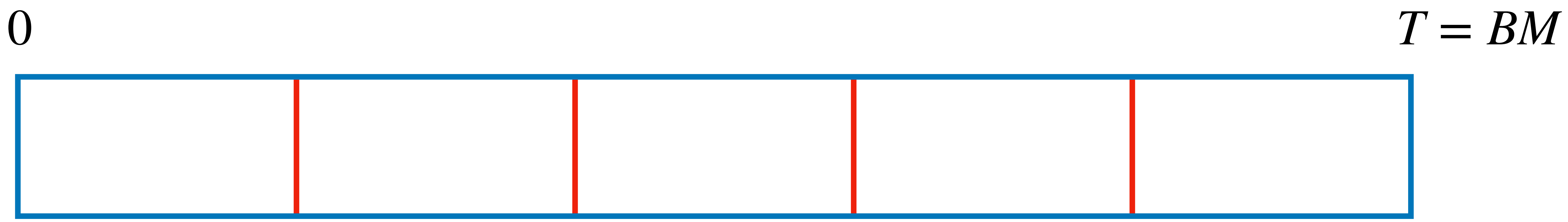
External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



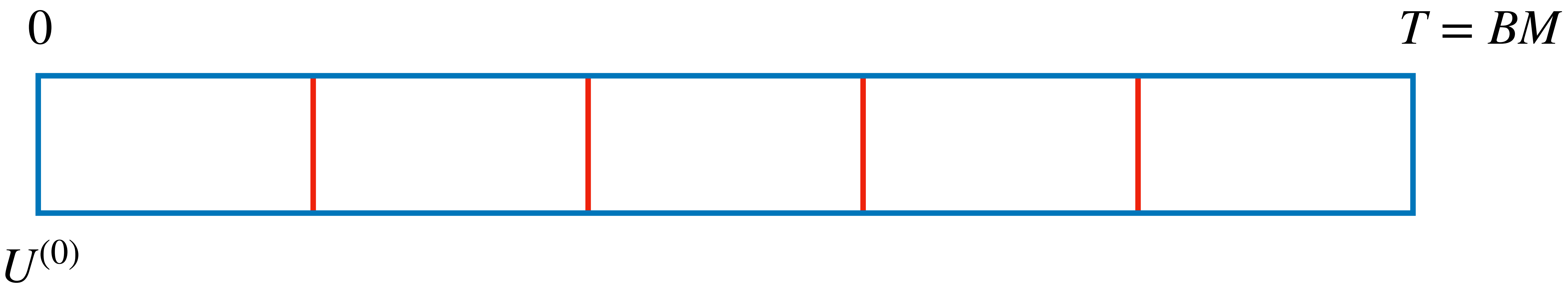
External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



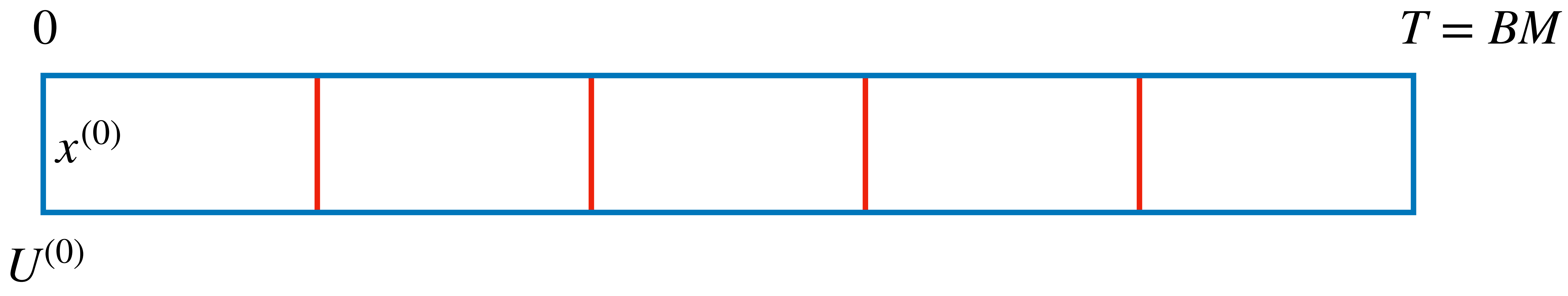
External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



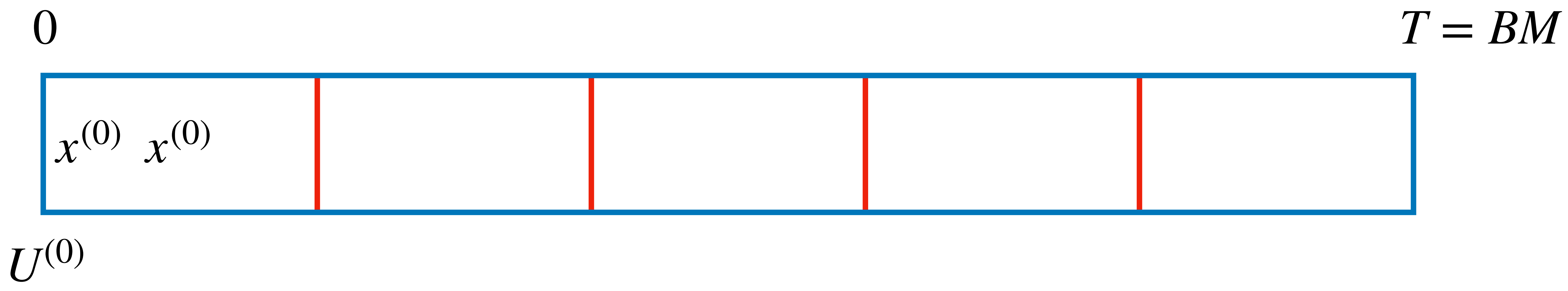
External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



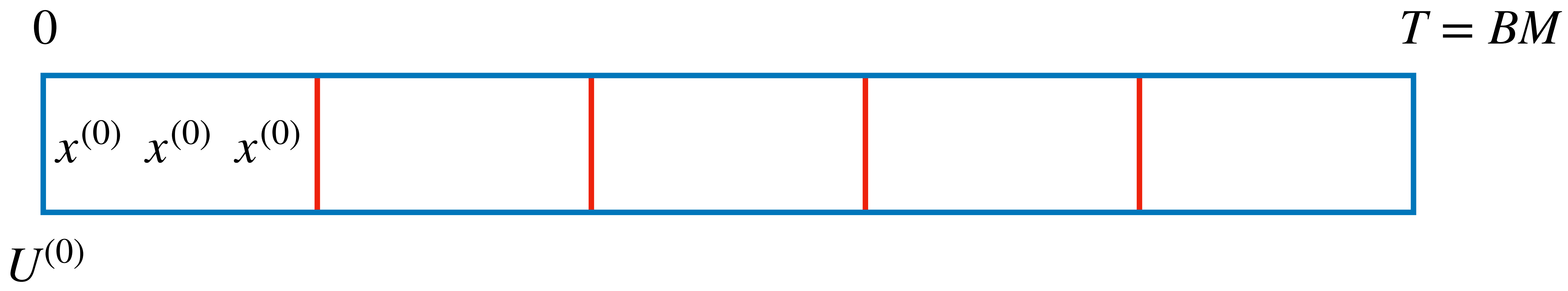
External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



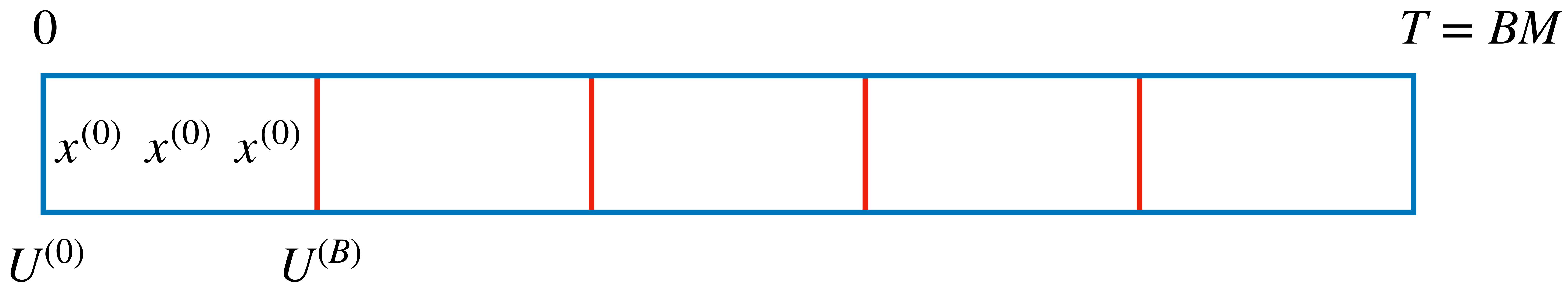
External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



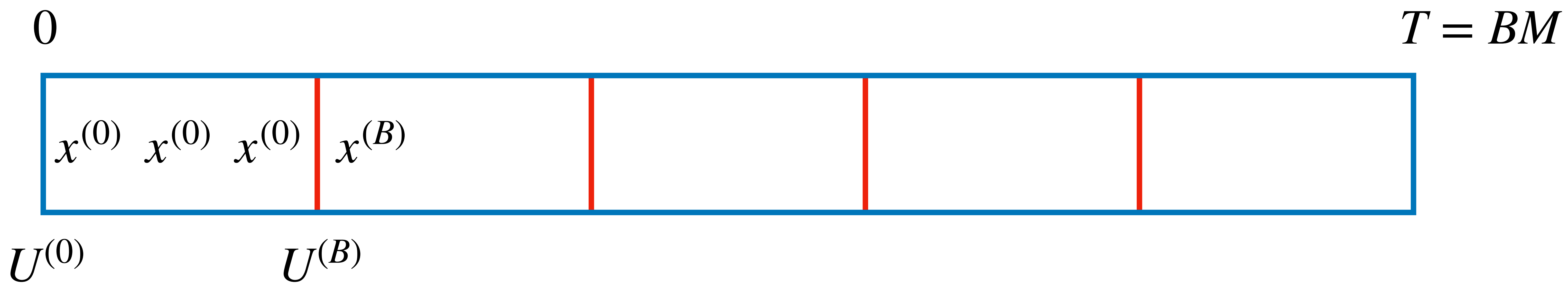
External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



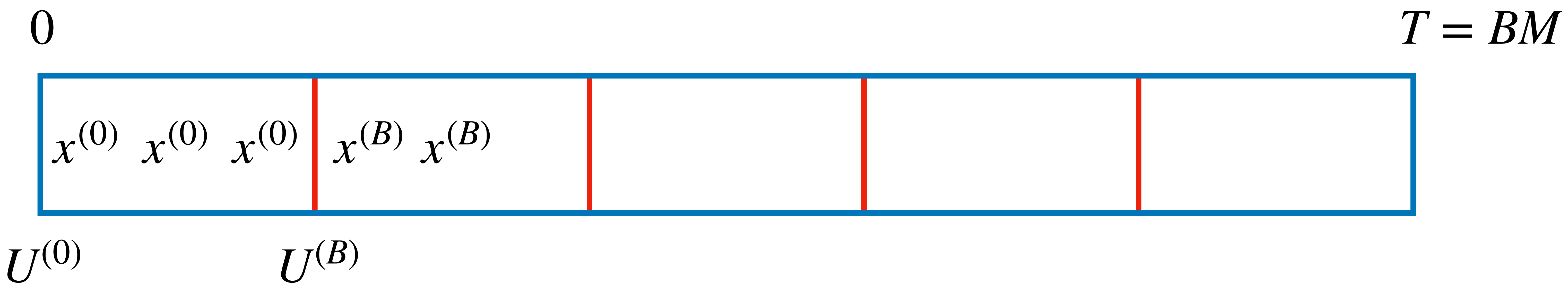
External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



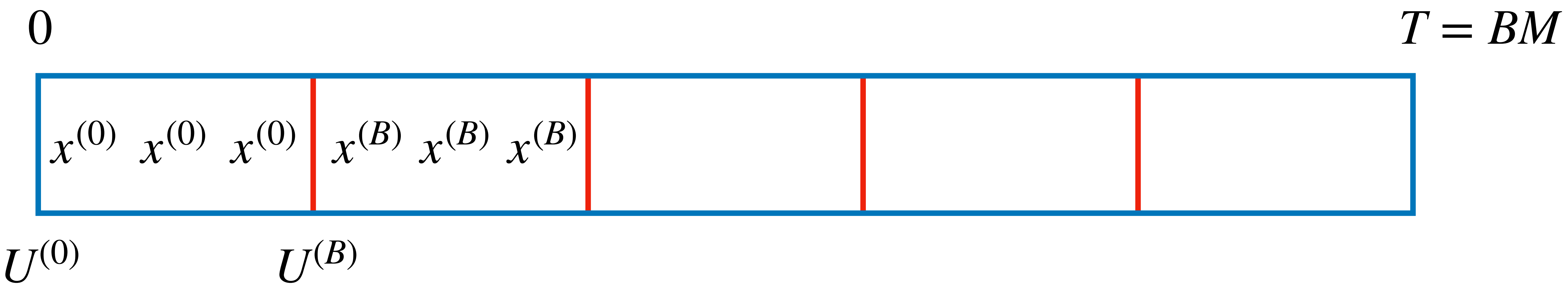
External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



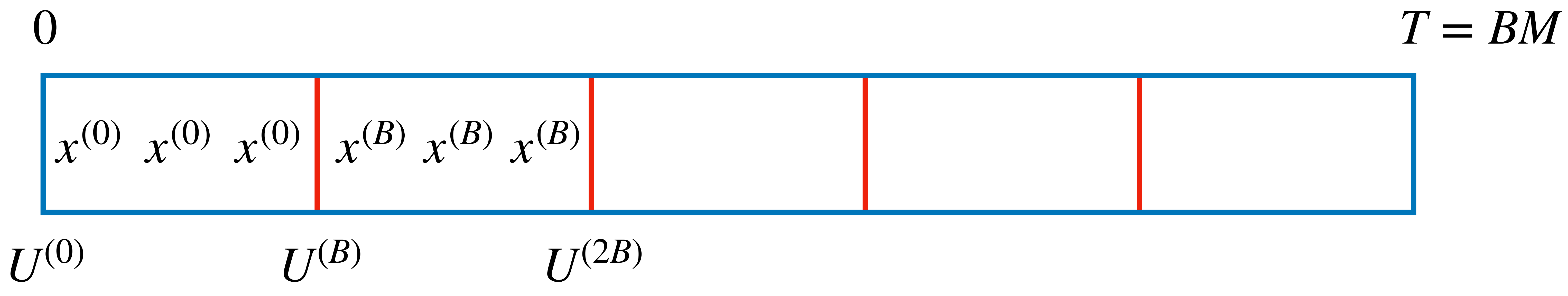
External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



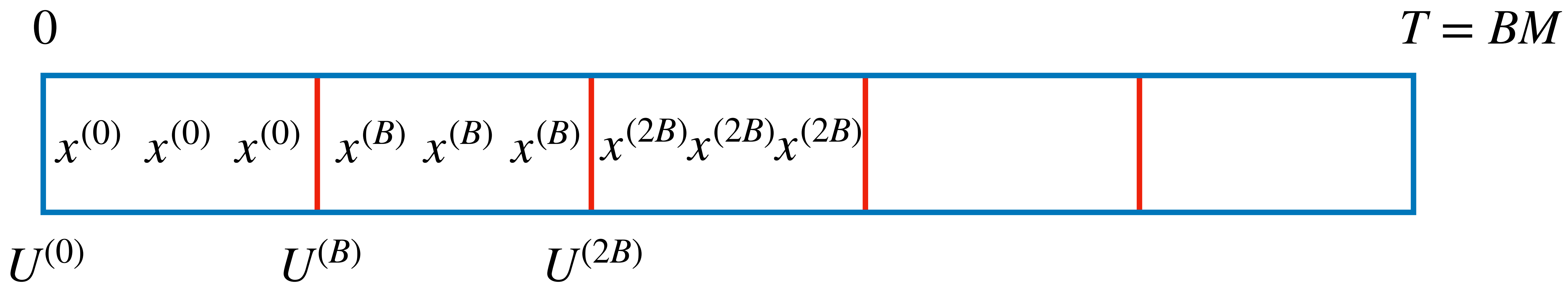
External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



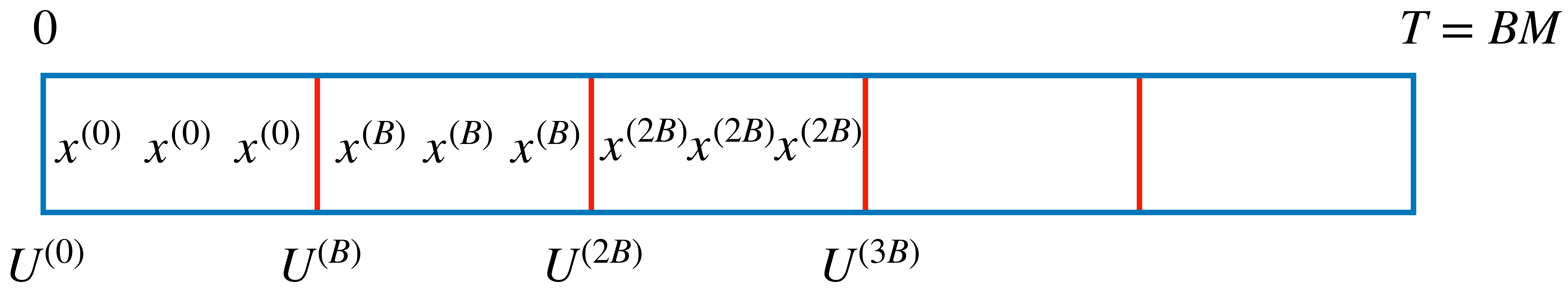
External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



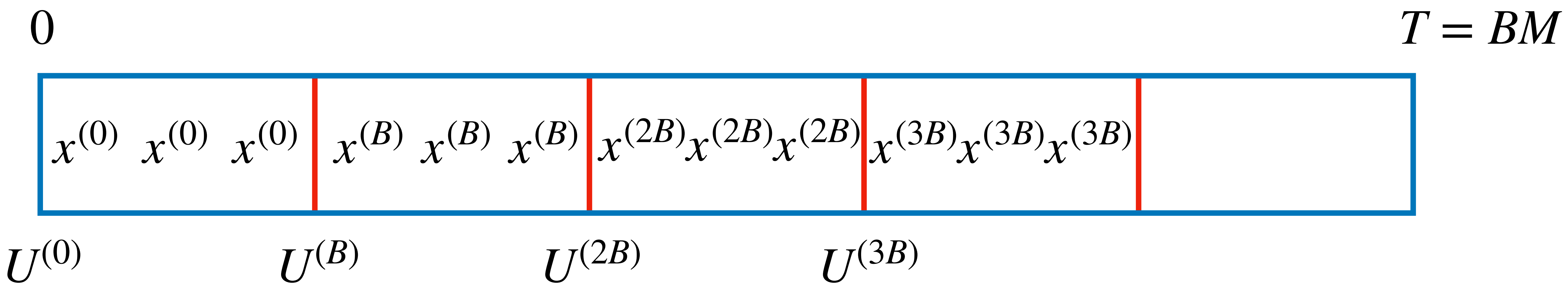
External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



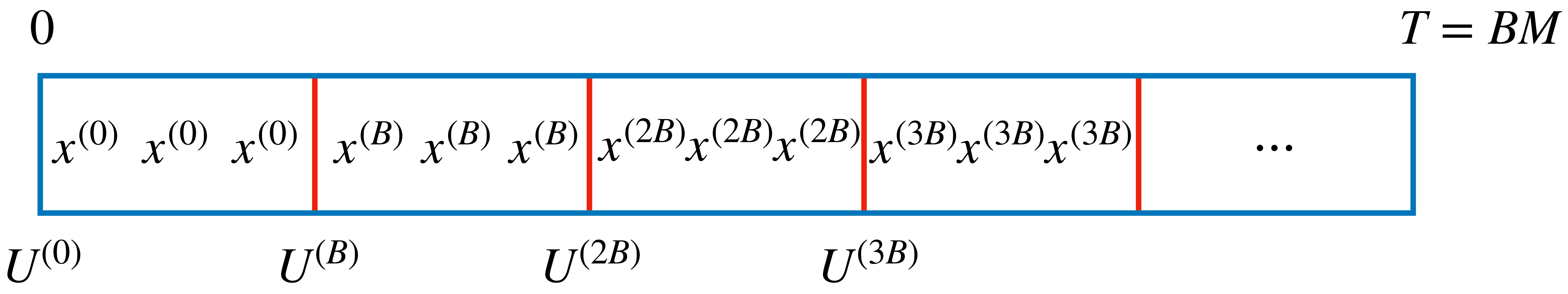
External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



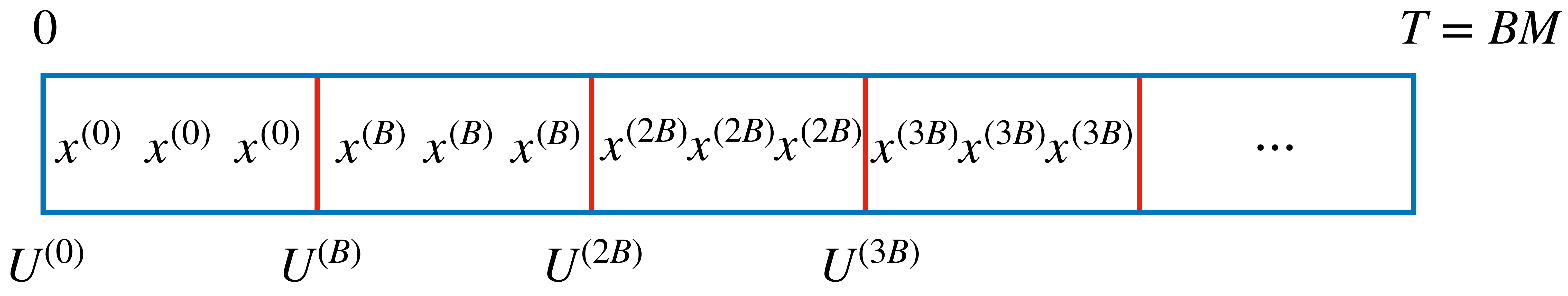
External Regret

Lazy **Multiplicative Weight Updates:** Look at the history, do the best thing (with some regularity) **every few rounds**



External Regret

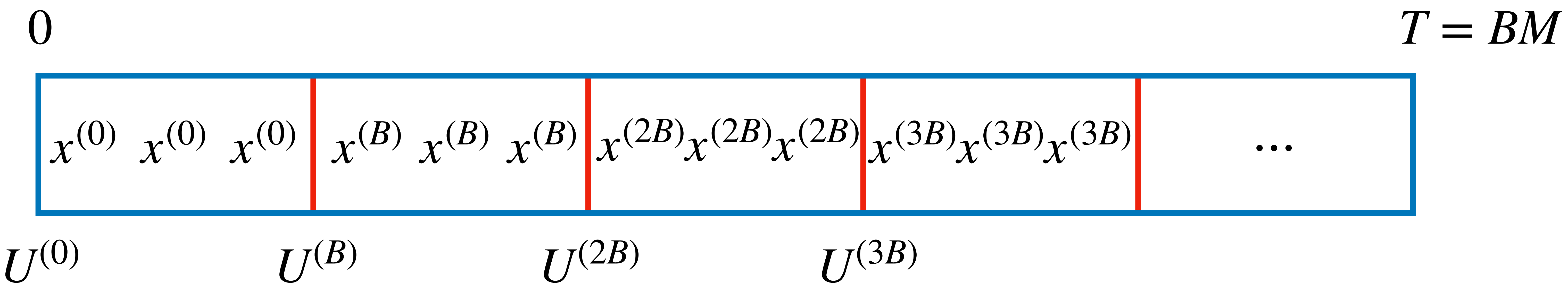
Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



MWU: ϵ-external-regret for $T = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**

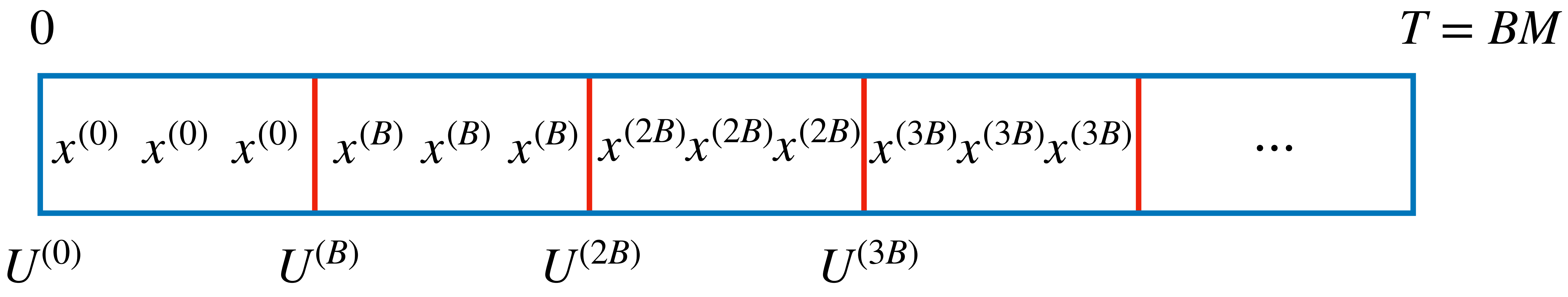


MWU: ϵ-external-regret for $T = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

Lazy MWU: ϵ-external-regret for $M = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



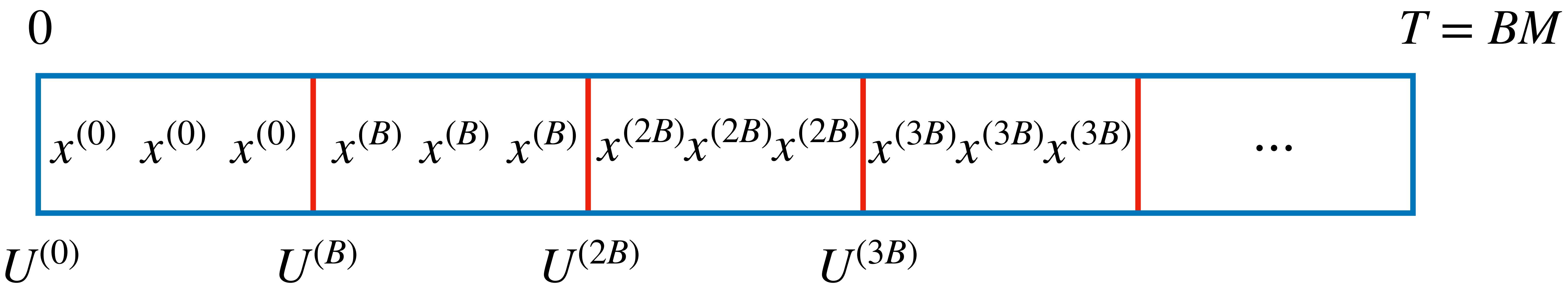
MWU: ϵ-external-regret for $T = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

Lazy MWU: ϵ-external-regret for $M = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

Why?

External Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



MWU: ϵ-external-regret for $T = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

Lazy MWU: ϵ-external-regret for $M = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

Why?

Fewer Actions

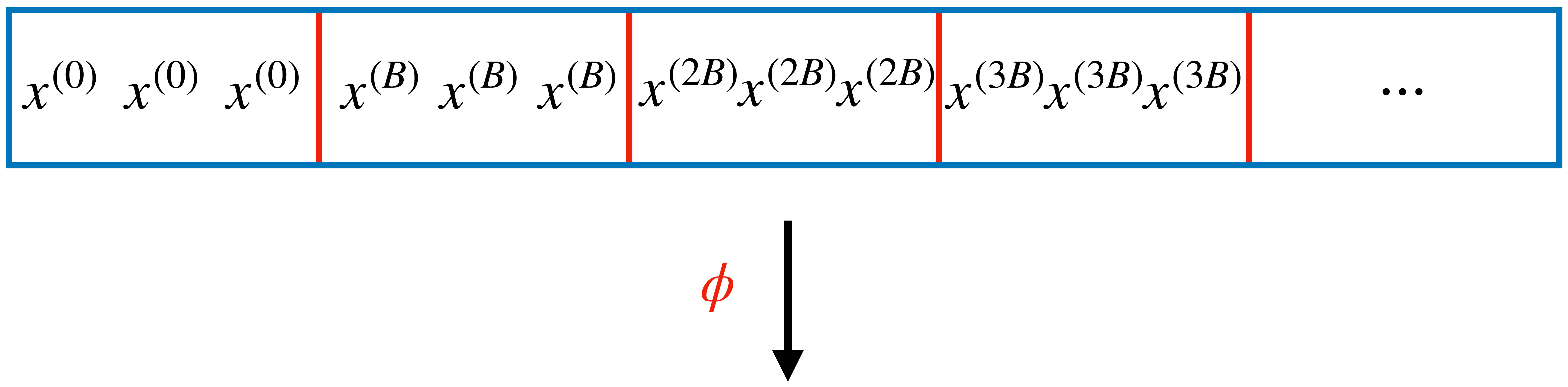
Swap Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**

$x^{(0)}$ $x^{(0)}$ $x^{(0)}$	$x^{(B)}$ $x^{(B)}$ $x^{(B)}$	$x^{(2B)}$ $x^{(2B)}$ $x^{(2B)}$	$x^{(3B)}$ $x^{(3B)}$ $x^{(3B)}$	$\dots$

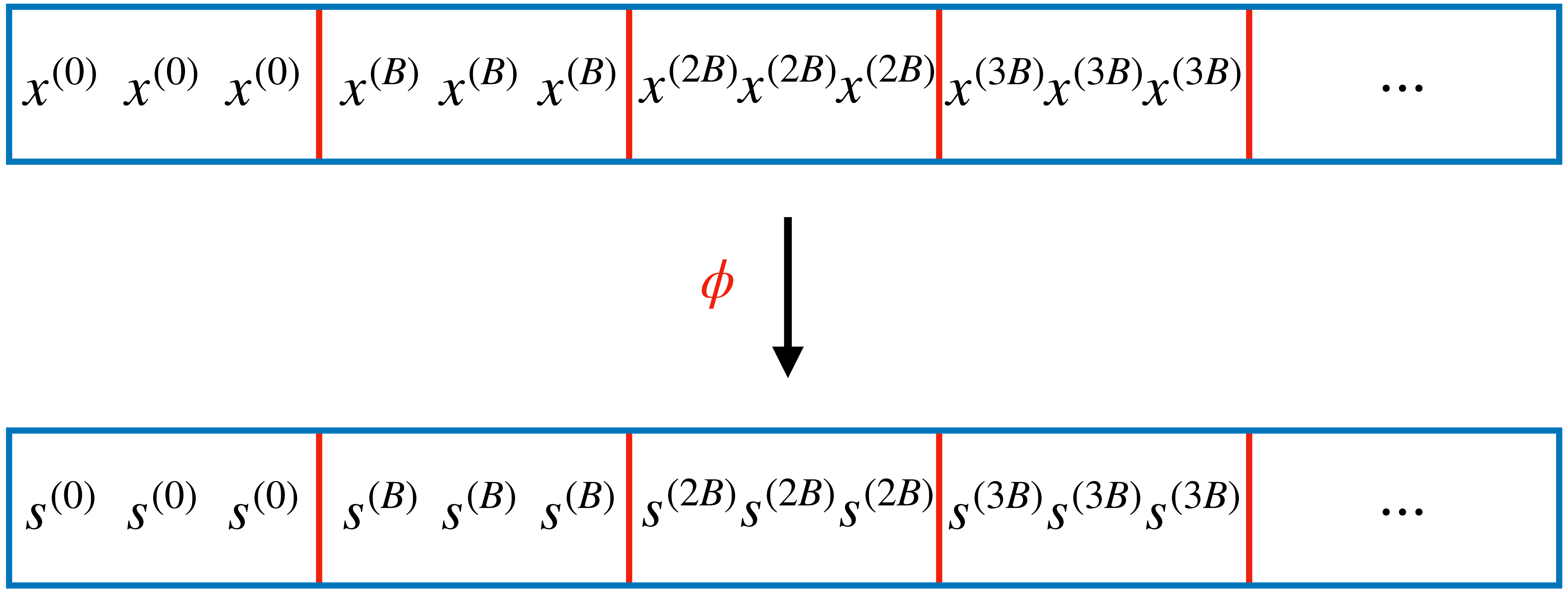
Swap Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



Swap Regret

Lazy Multiplicative Weight Updates: Look at the history, do the best thing (with some regularity) **every few rounds**



Tree-Swap

Tree-Swap

$$T = M^d$$

Tree-Swap

$$T = M^d$$

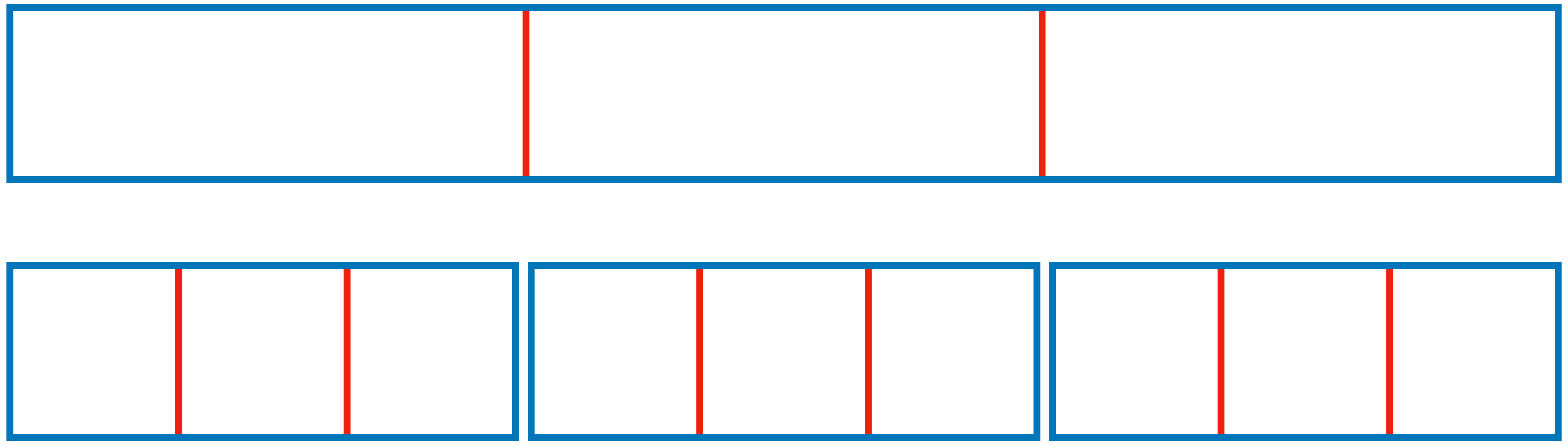
1 Lazy-MWU instance: $[0, M^d]$

Tree-Swap

$$T = M^d$$

1 Lazy-MWU instance: $[0, M^d]$

M Lazy-MWU instances: $[0, M^{d-1}], [M^{d-1}, 2M^{d-1}], \cdots$



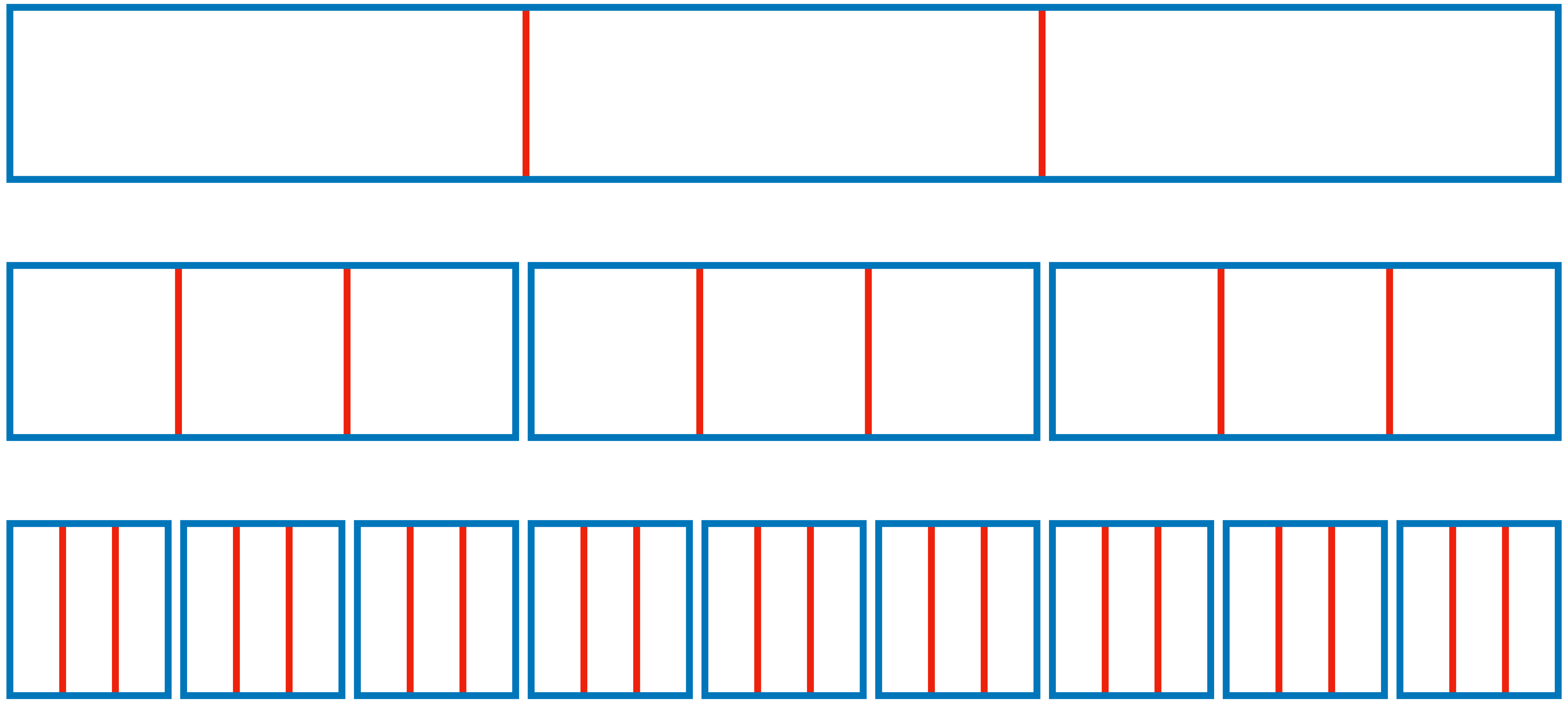
Tree-Swap

$T = M^d$

1 Lazy-MWU instance: $[0, M^d]$

M Lazy-MWU instances: $[0, M^{d-1}], [M^{d-1}, 2M^{d-1}], \cdots$

M^2 Lazy-MWU instances: $[0, M^{d-2}], [M^{d-2}, 2M^{d-2}], \cdots$



Tree-Swap

$$T = M^d$$

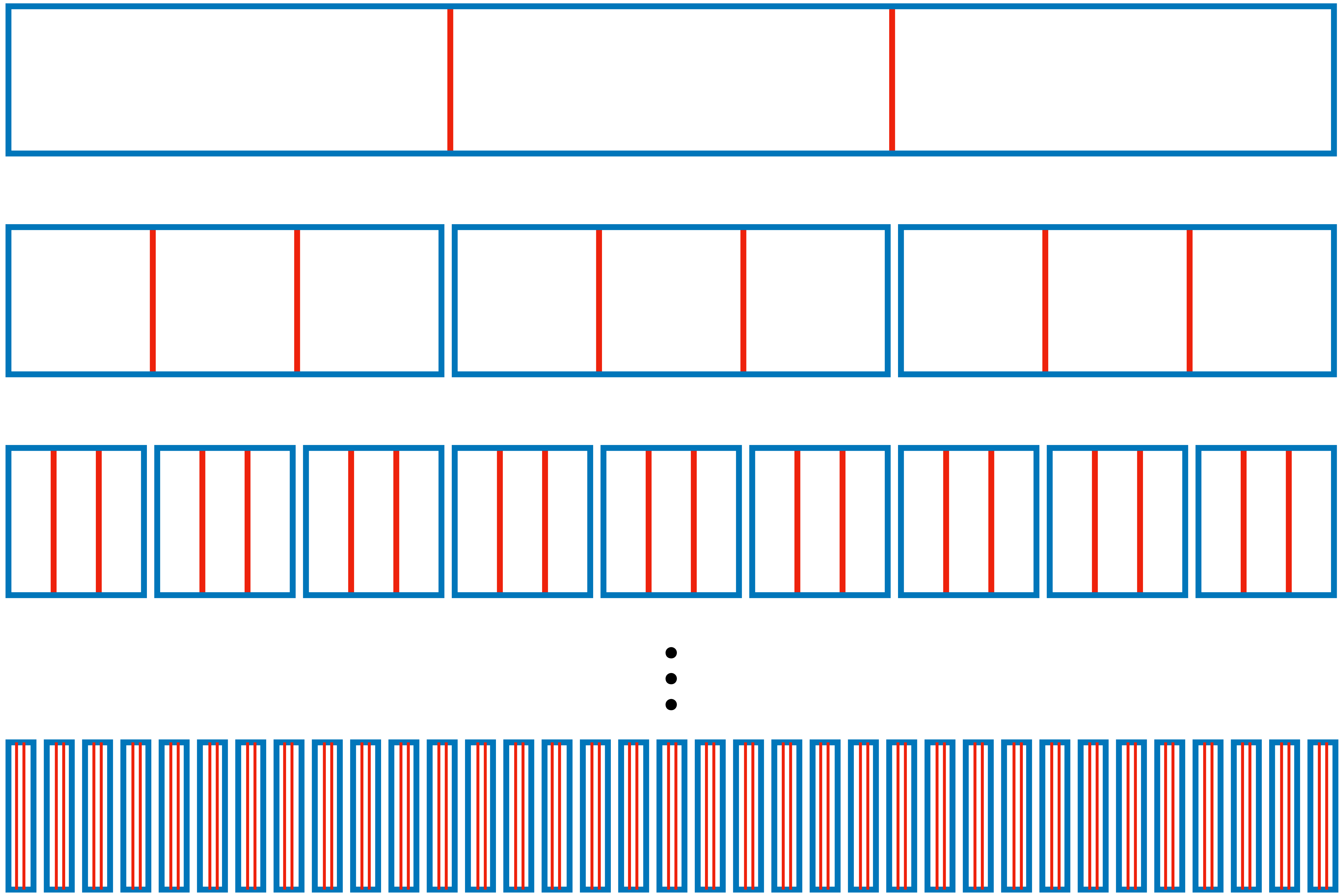
1 Lazy-MWU instance: $[0, M^d]$

M Lazy-MWU instances: $[0, M^{d-1}], [M^{d-1}, 2M^{d-1}], \cdots$

M^2 Lazy-MWU instances: $[0, M^{d-2}], [M^{d-2}, 2M^{d-2}], \cdots$

$\vdots$

M^{d-1} Lazy-MWU instances: $[0, M], [M, 2M], \cdots$



Tree-Swap

$T = M^d$

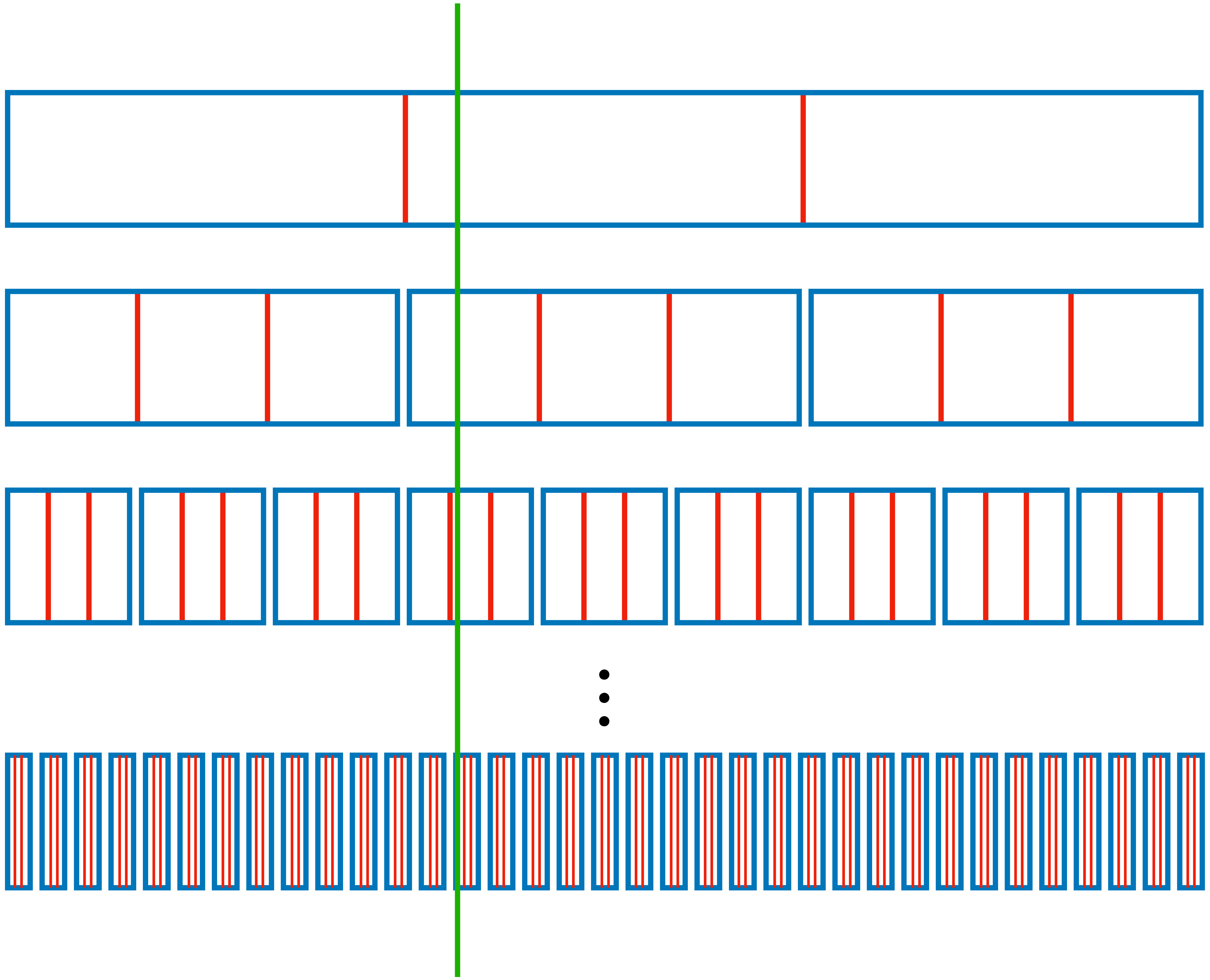
1 Lazy-MWU instance: $[0, M^d]$

M Lazy-MWU instances: $[0, M^{d-1}], [M^{d-1}, 2M^{d-1}], \cdots$

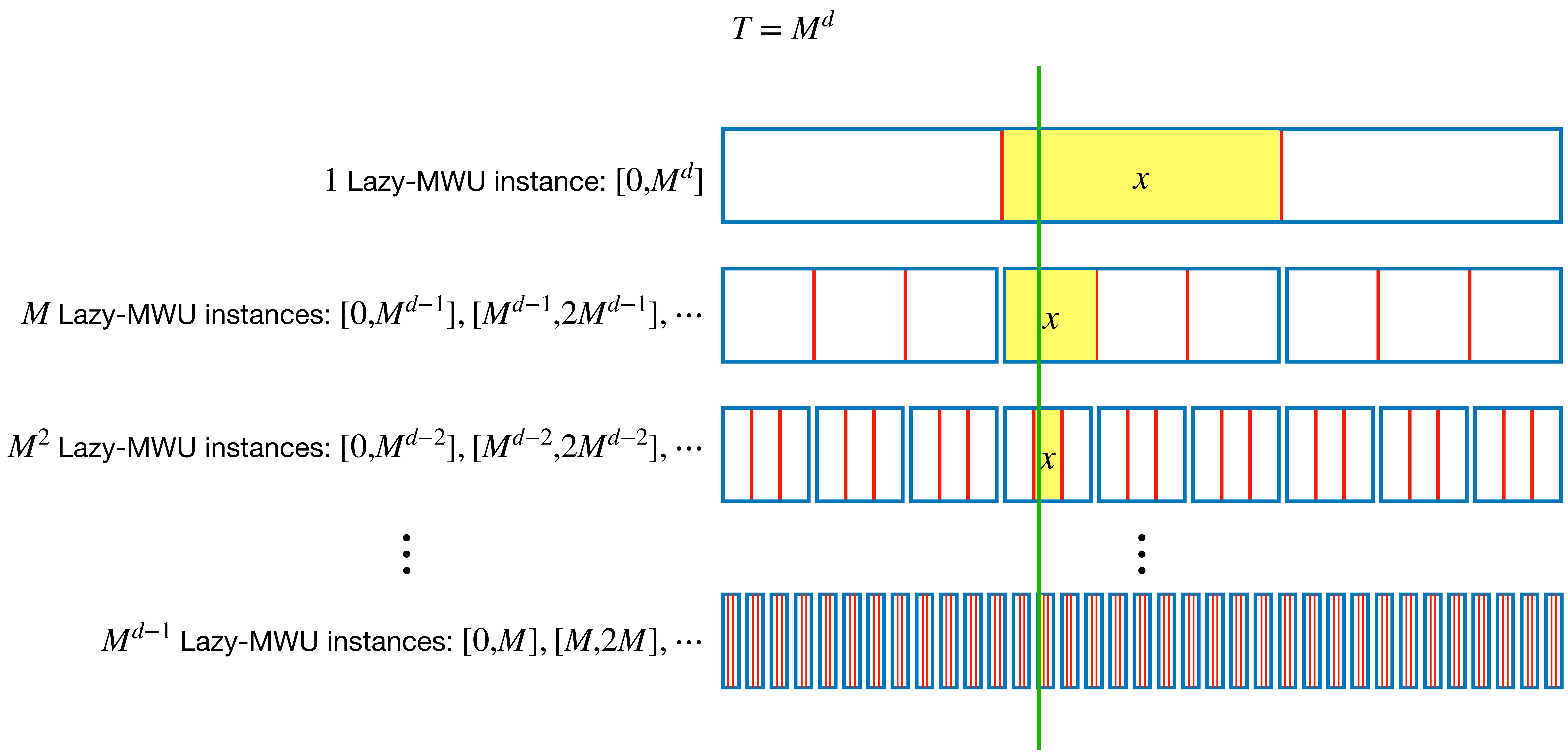
M^2 Lazy-MWU instances: $[0, M^{d-2}], [M^{d-2}, 2M^{d-2}], \cdots$

$\vdots$

M^{d-1} Lazy-MWU instances: $[0, M], [M, 2M], \cdots$



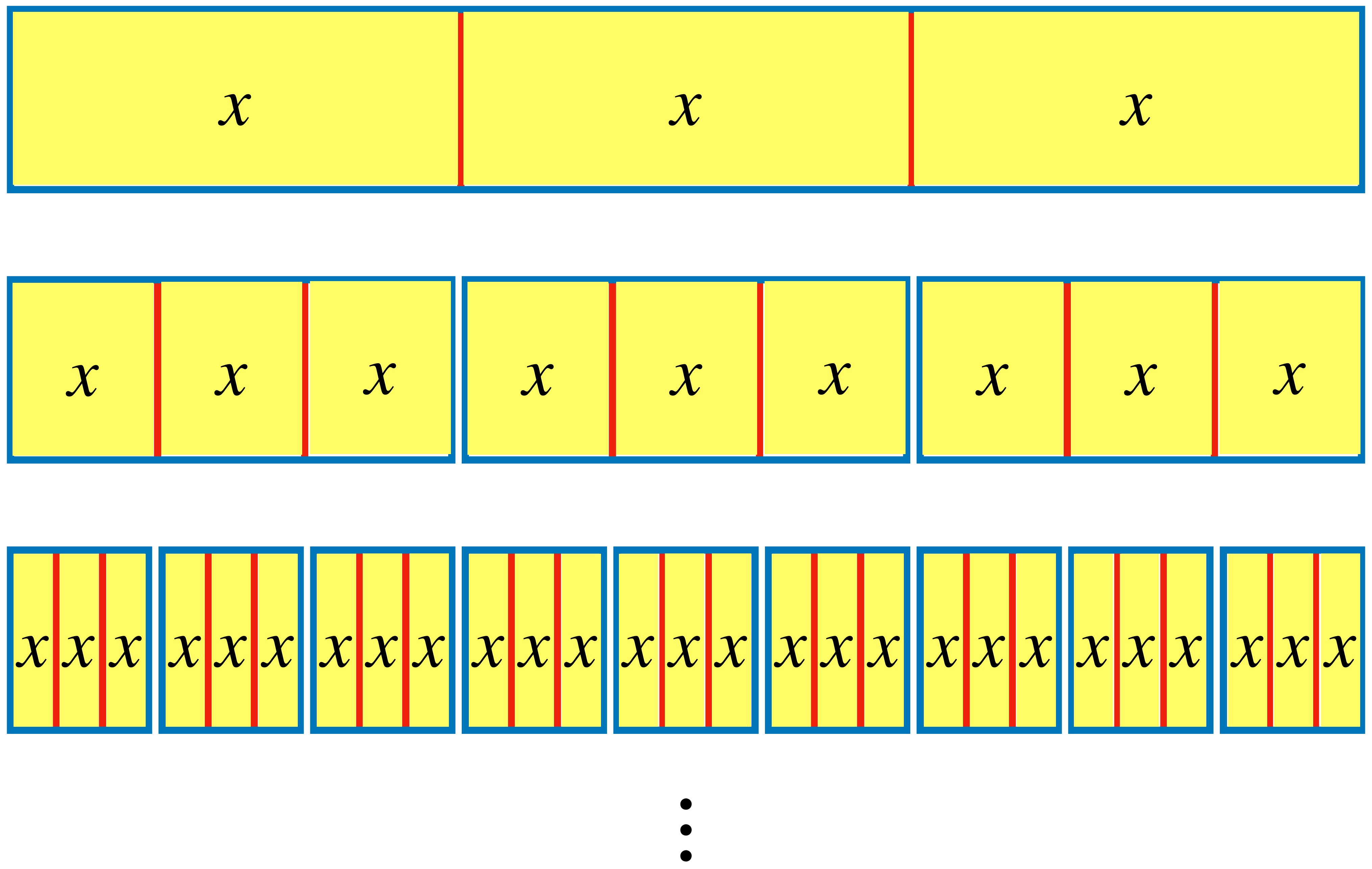
Tree-Swap



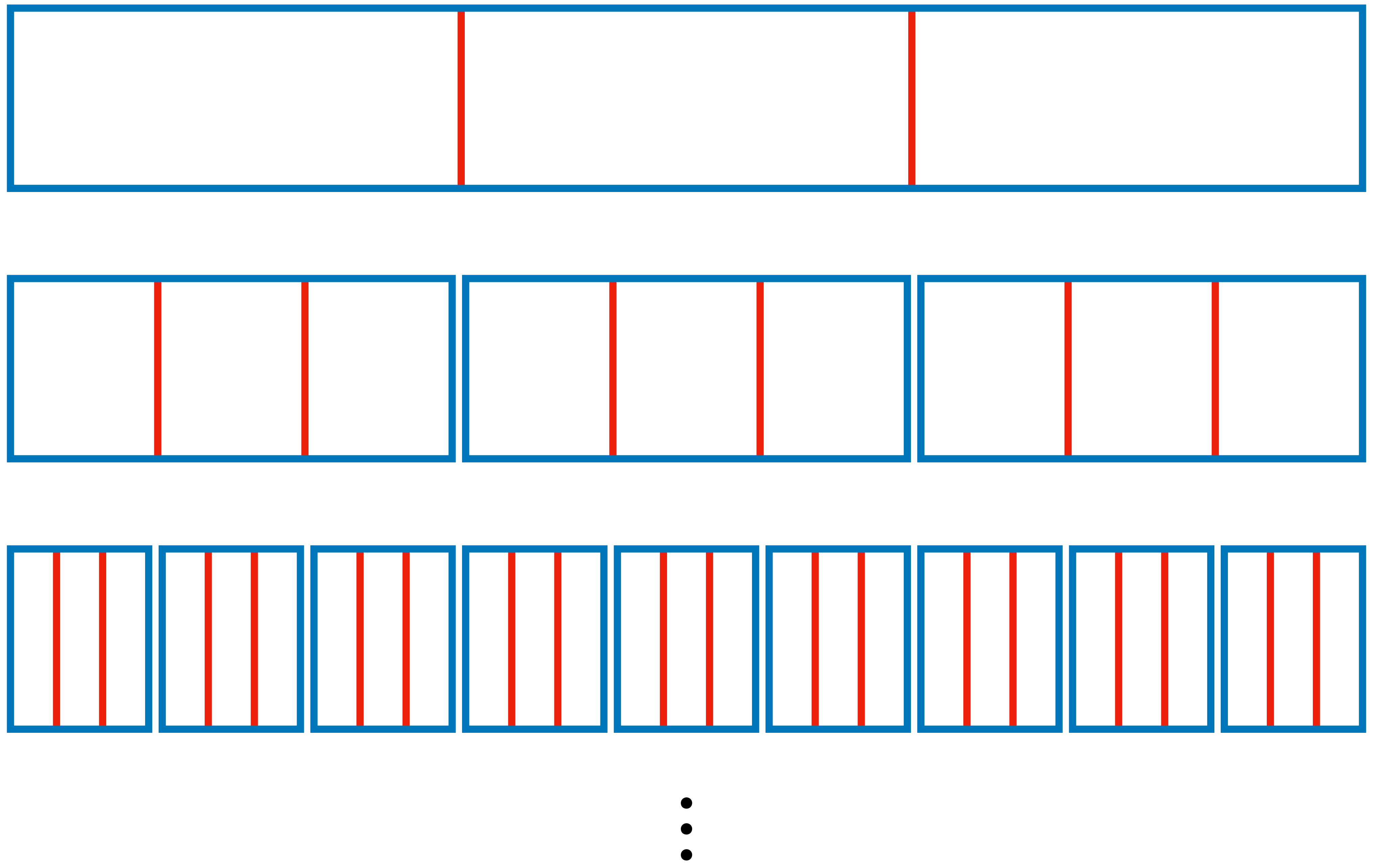
$T = M^d$
1 Lazy-MWU instance: $[0, M^d]$
M Lazy-MWU instances: $[0, M^{d-1}], [M^{d-1}, 2M^{d-1}], \cdots$
M^2 Lazy-MWU instances: $[0, M^{d-2}], [M^{d-2}, 2M^{d-2}], \cdots$
M^{d-1} Lazy-MWU instances: $[0, M], [M, 2M], \cdots$
x
x
x

Tree-Swap Swap Regret

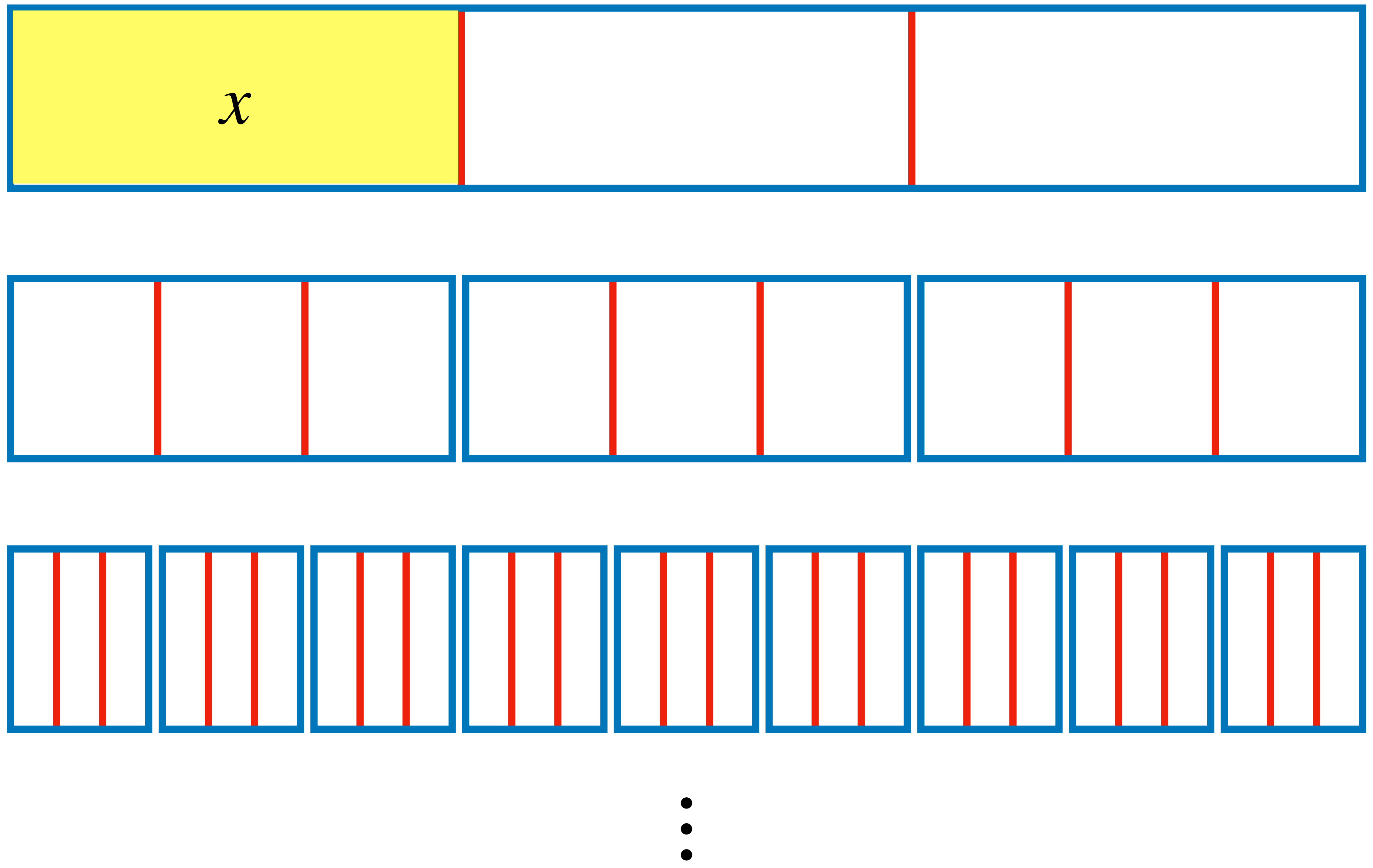
Tree-Swap Swap Regret



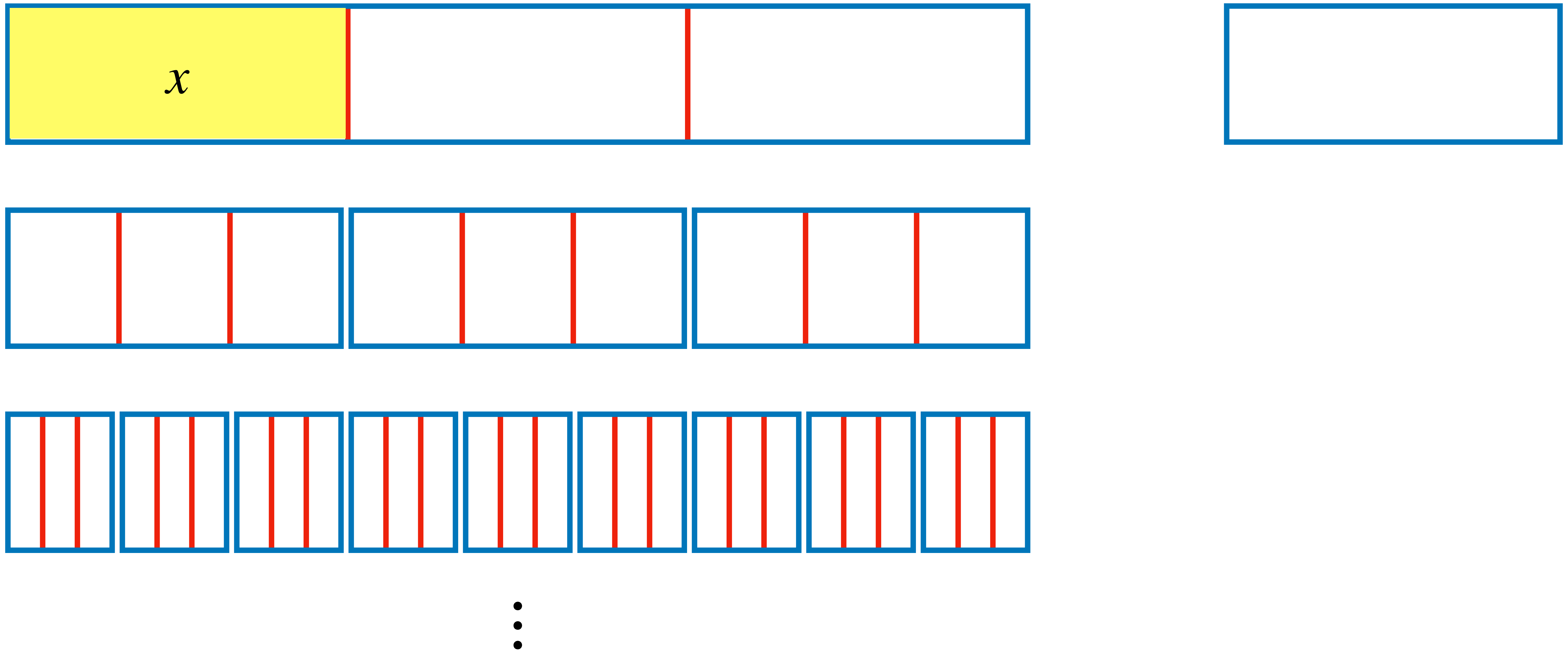
Tree-Swap Swap Regret



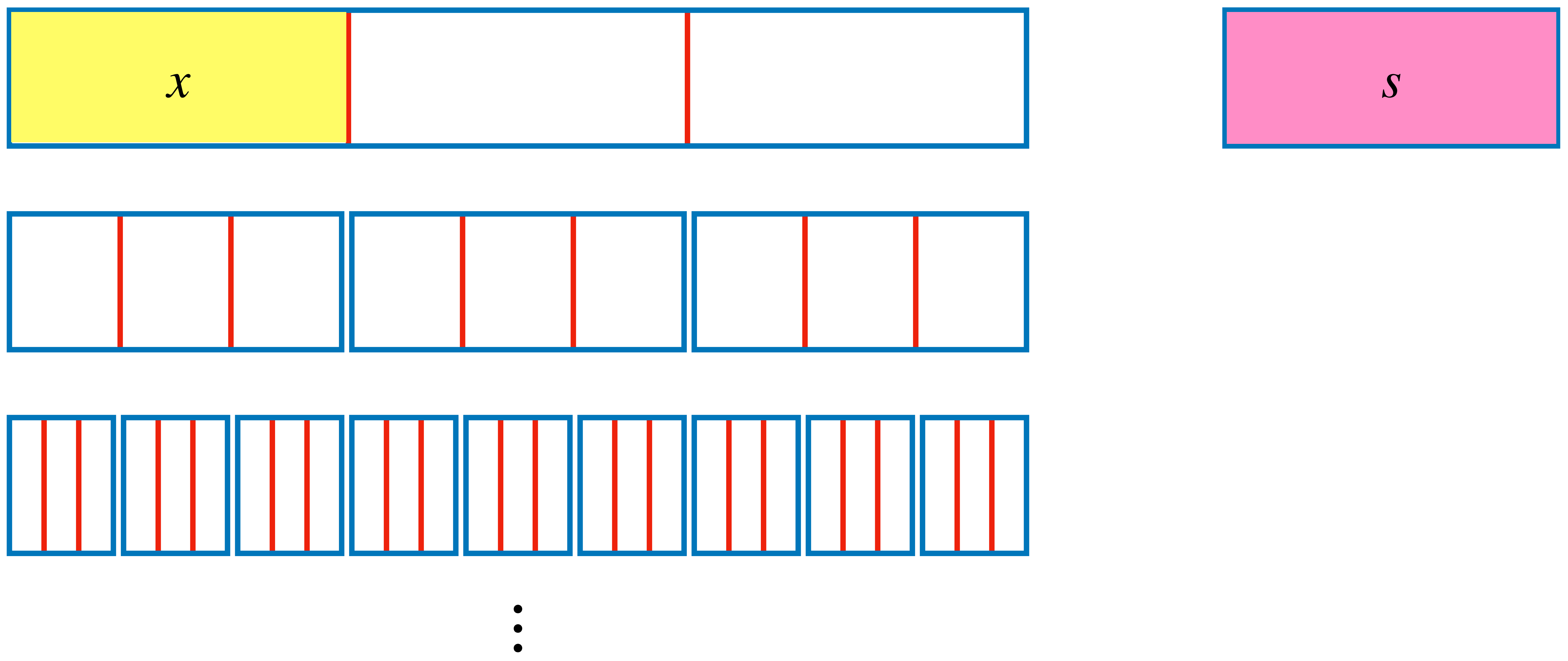
Tree-Swap Swap Regret



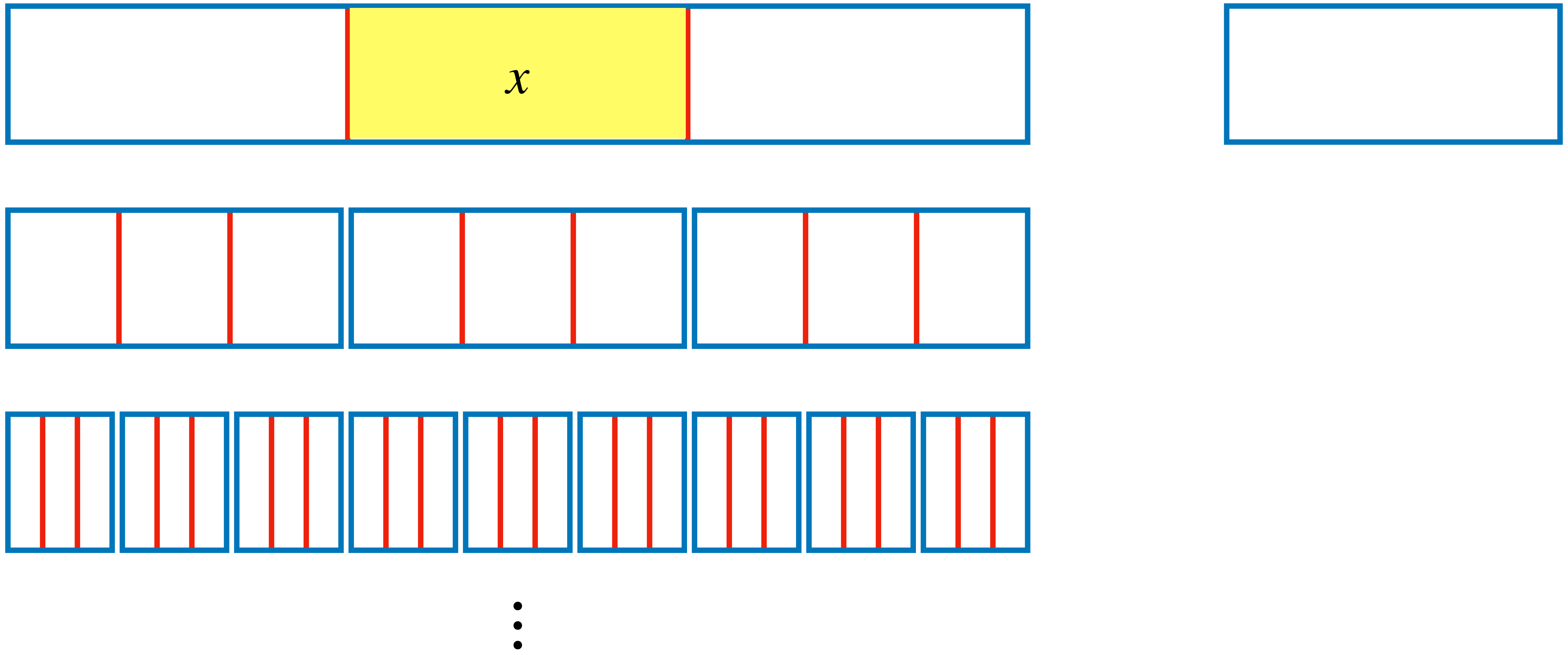
Tree-Swap Swap Regret



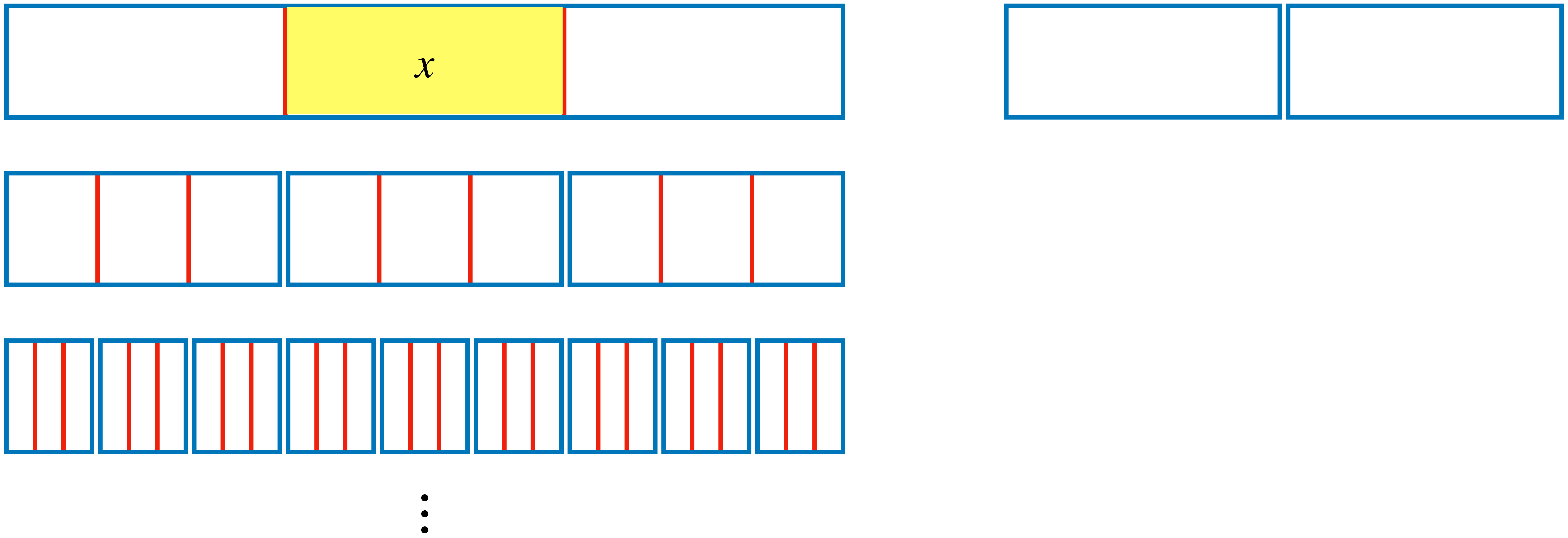
Tree-Swap Swap Regret



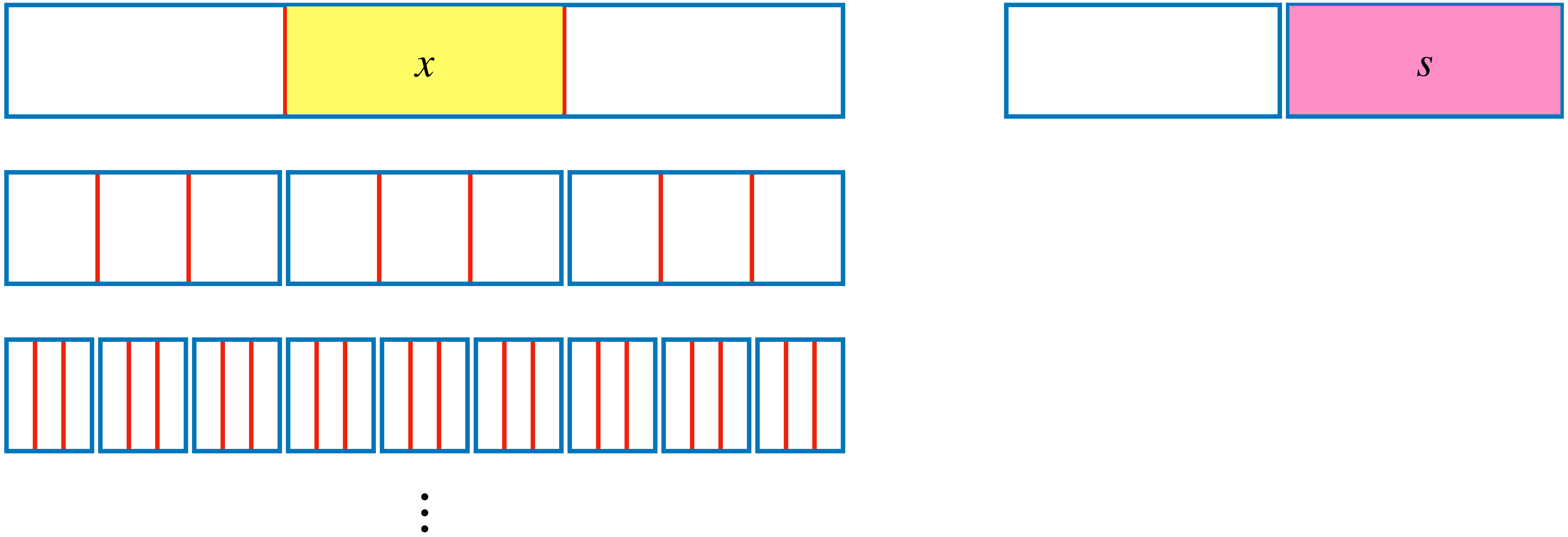
Tree-Swap Swap Regret



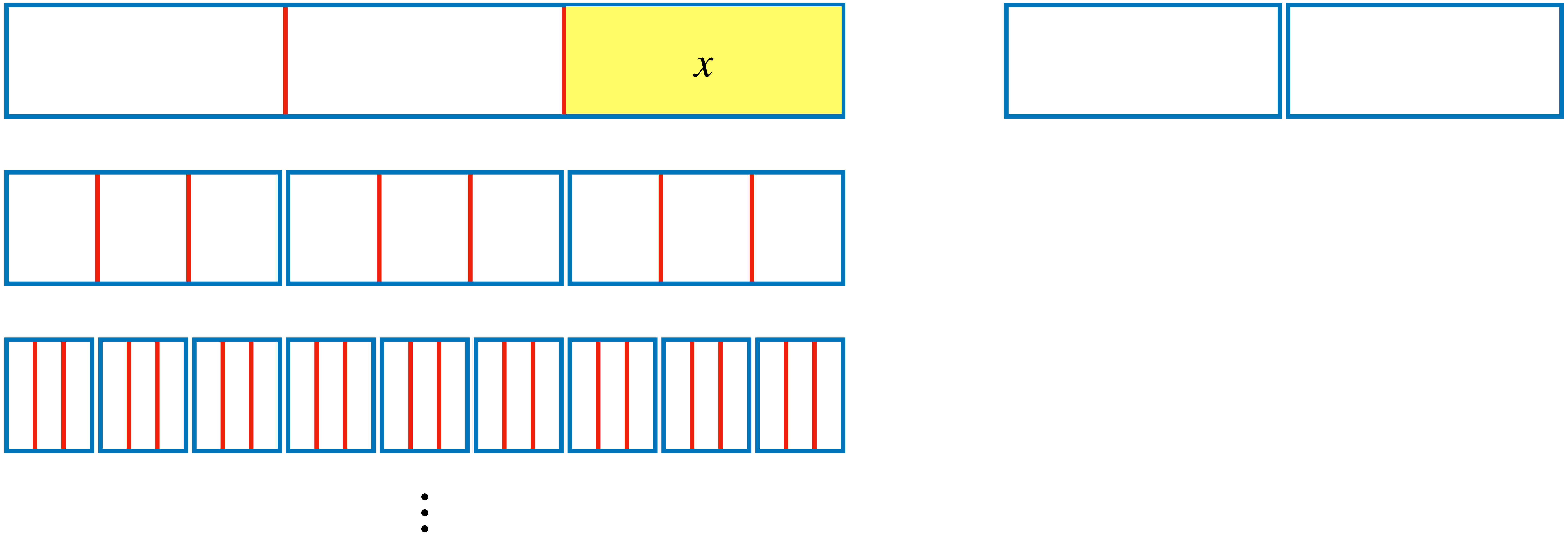
Tree-Swap Swap Regret



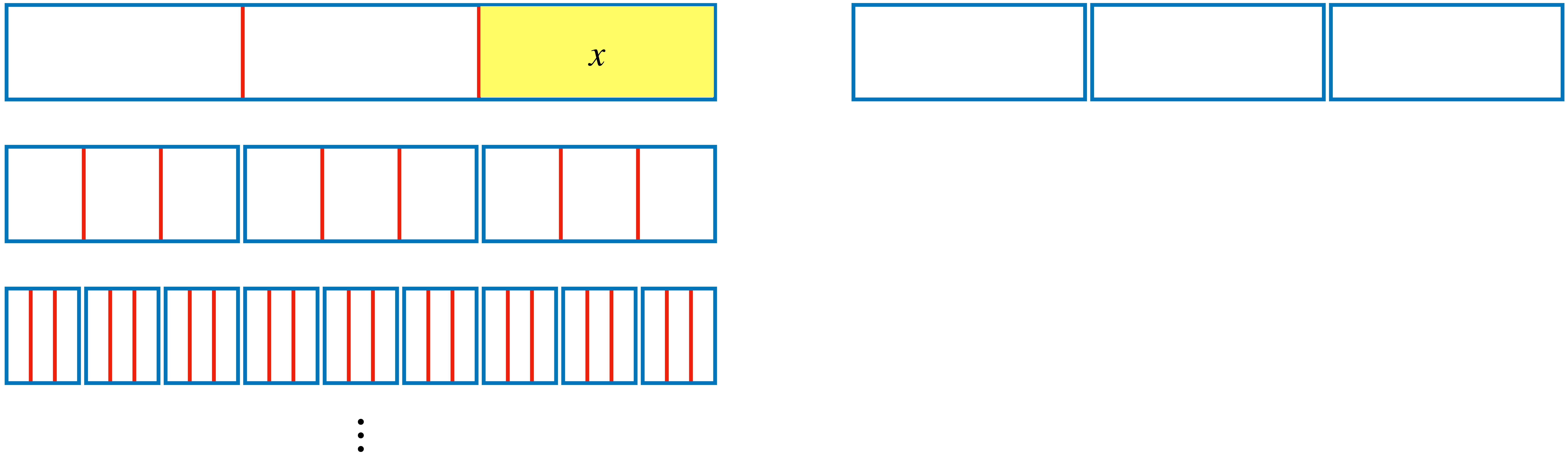
Tree-Swap Swap Regret



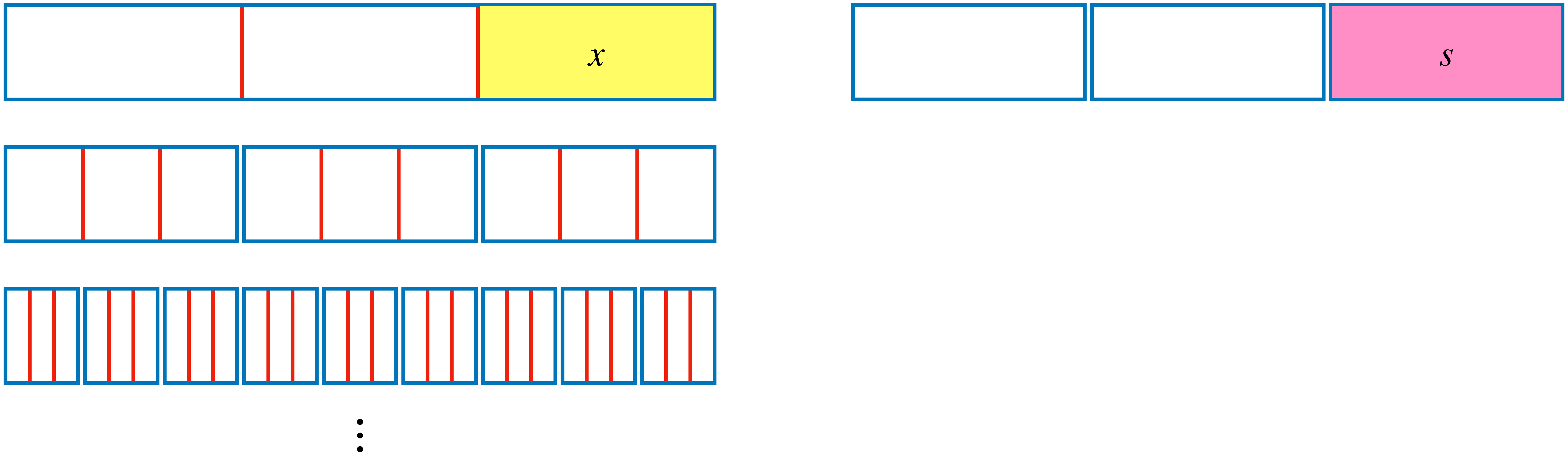
Tree-Swap Swap Regret



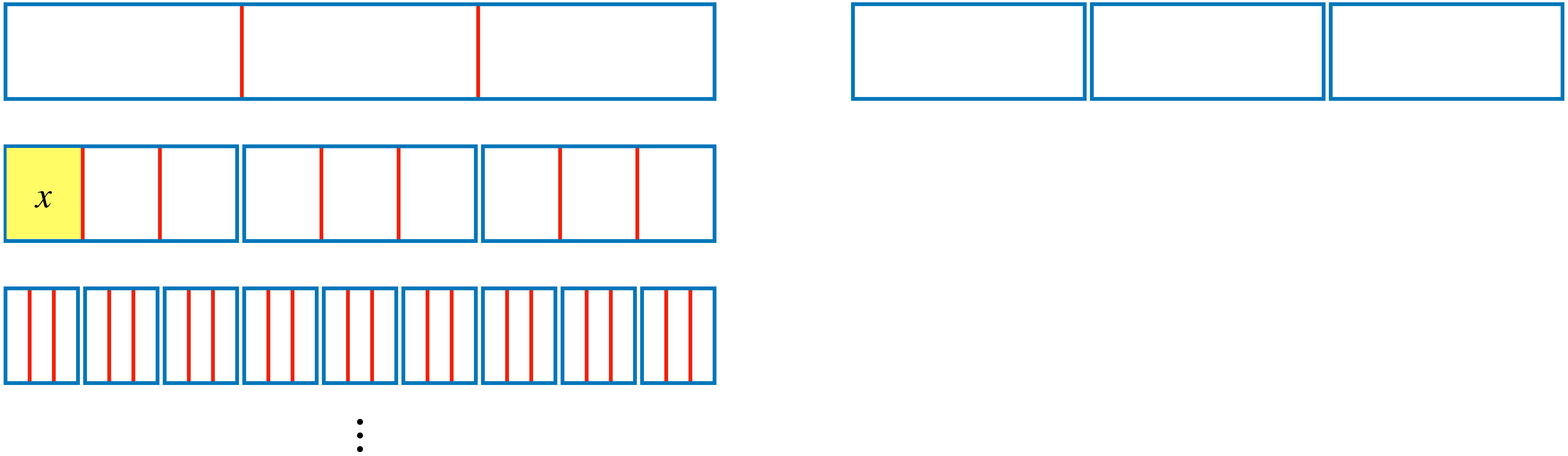
Tree-Swap Swap Regret



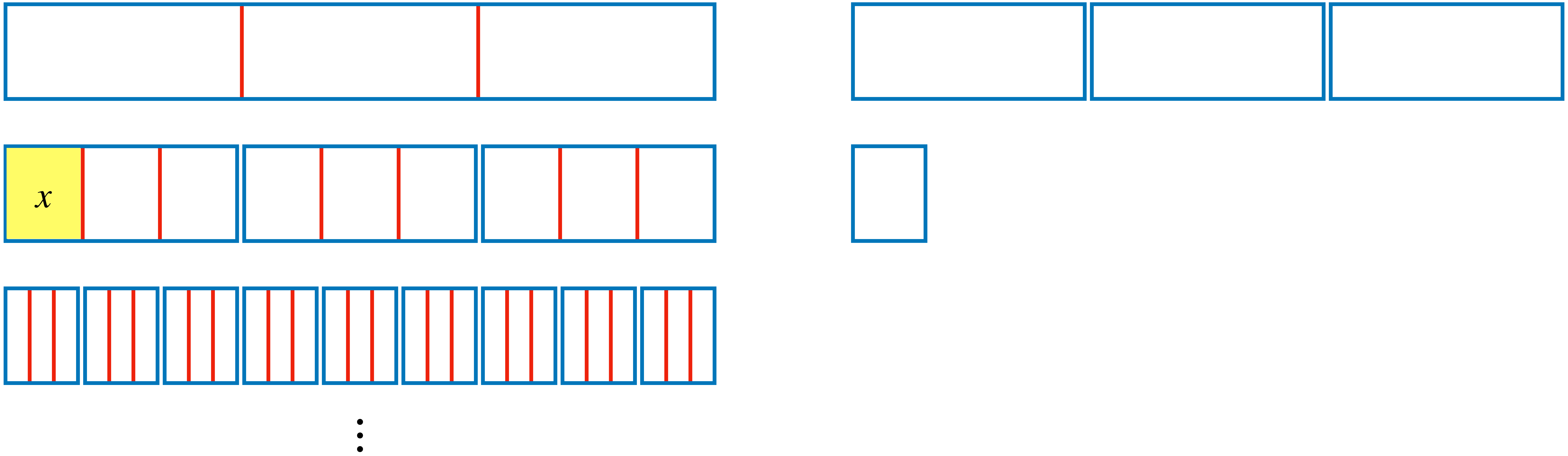
Tree-Swap Swap Regret



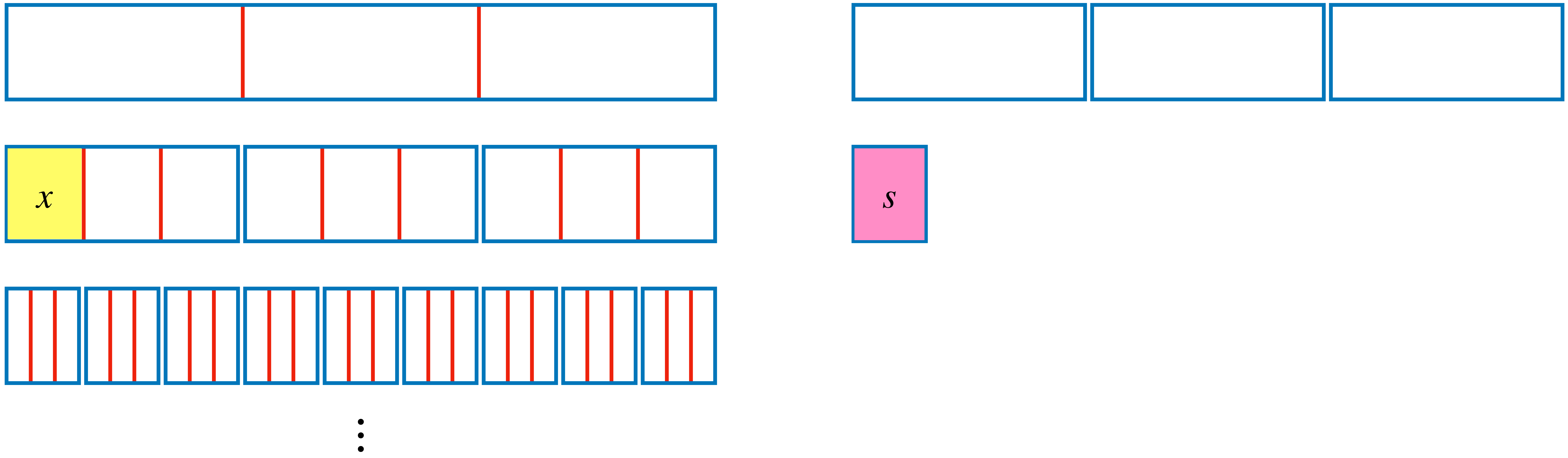
Tree-Swap Swap Regret



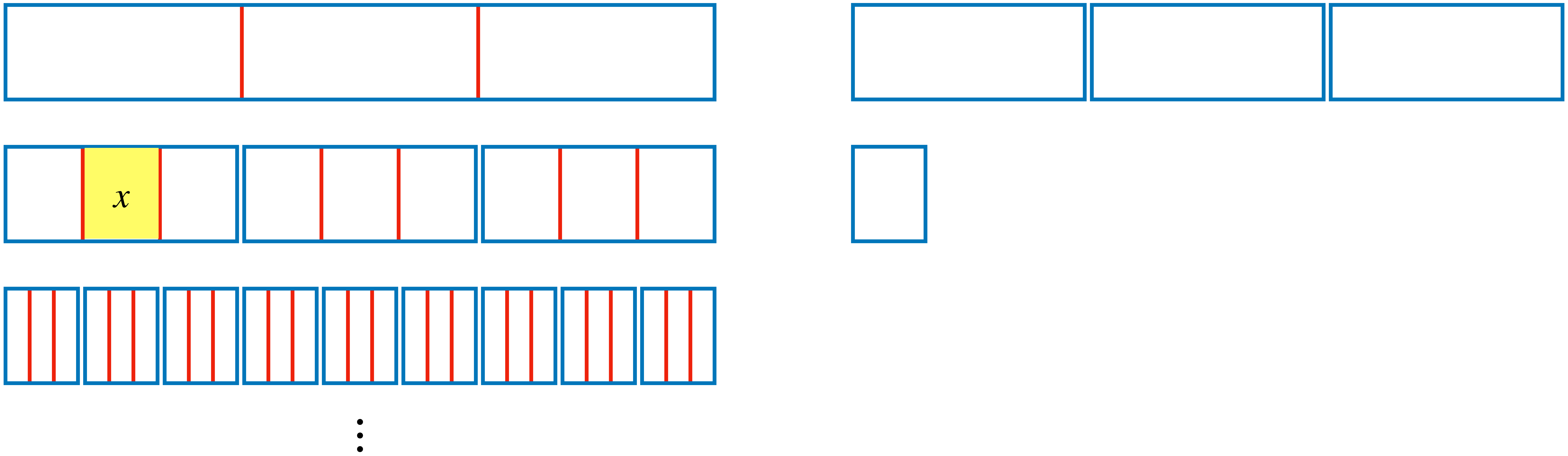
Tree-Swap Swap Regret



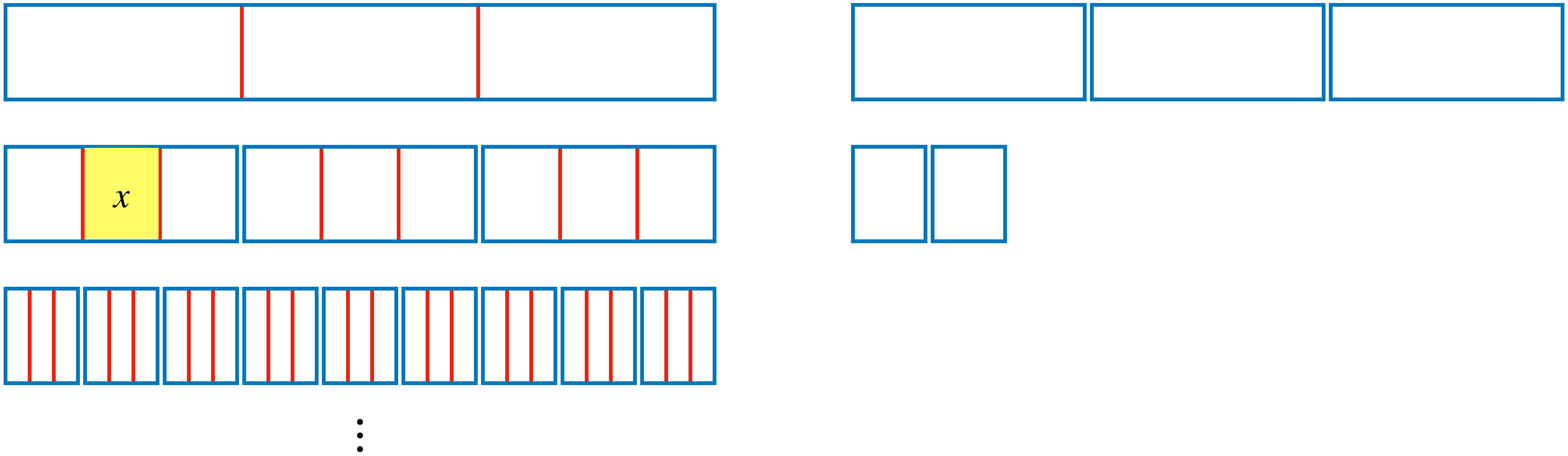
Tree-Swap Swap Regret



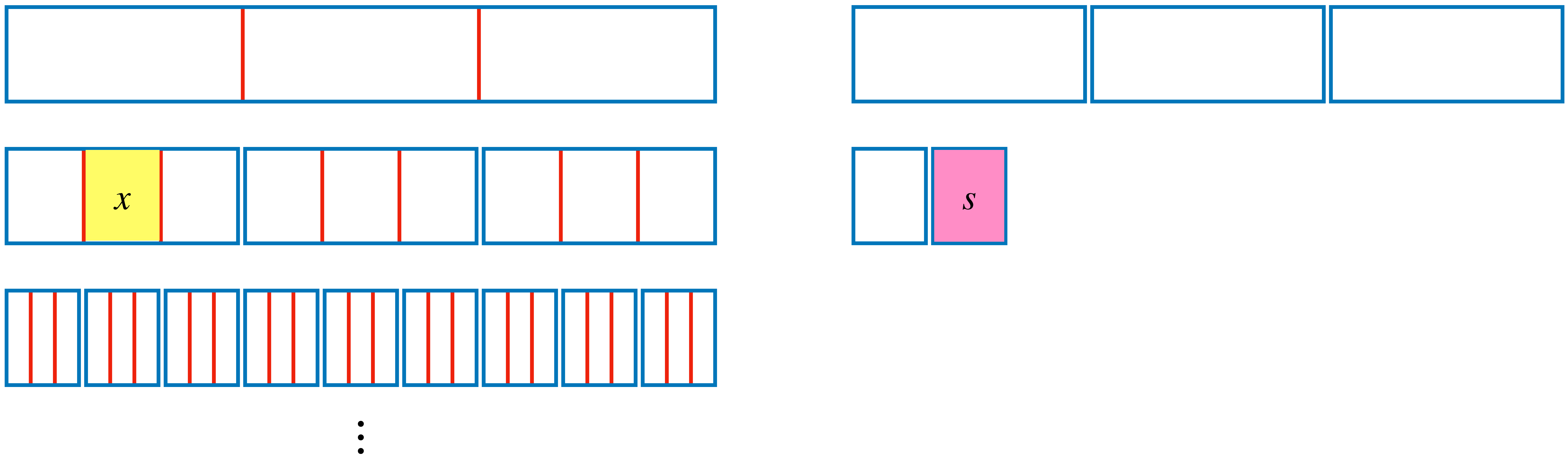
Tree-Swap Swap Regret



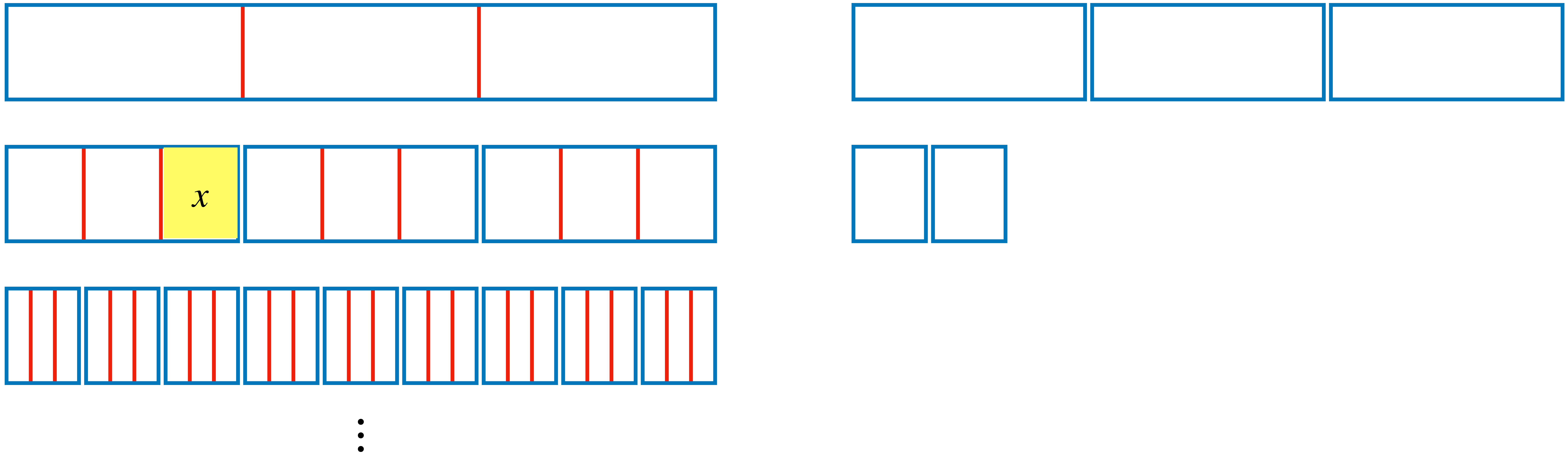
Tree-Swap Swap Regret



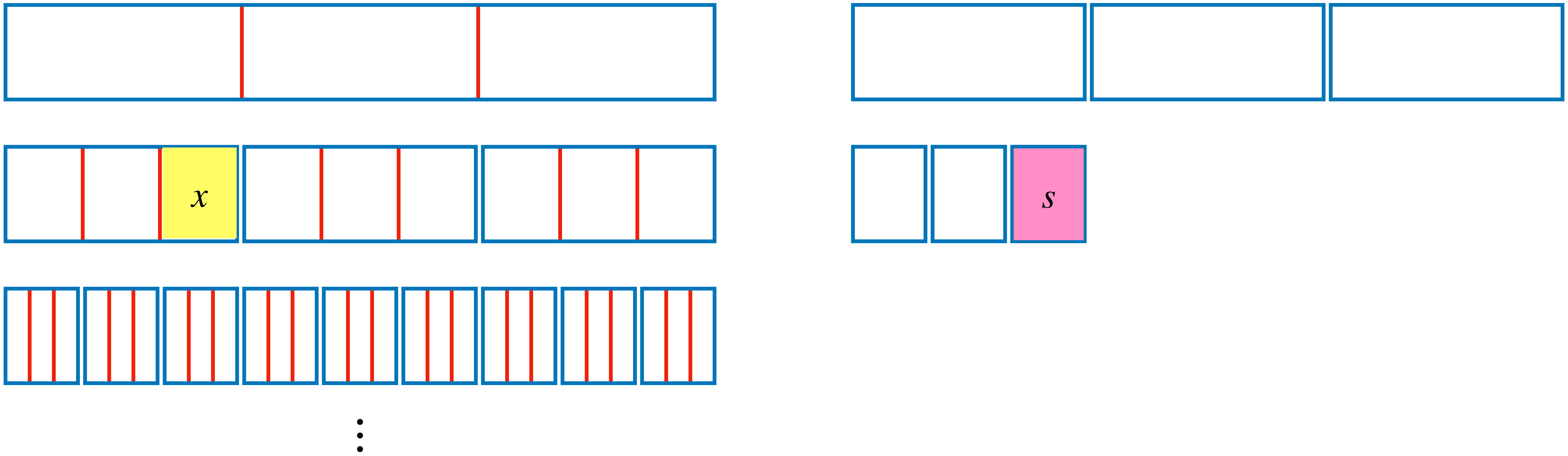
Tree-Swap Swap Regret



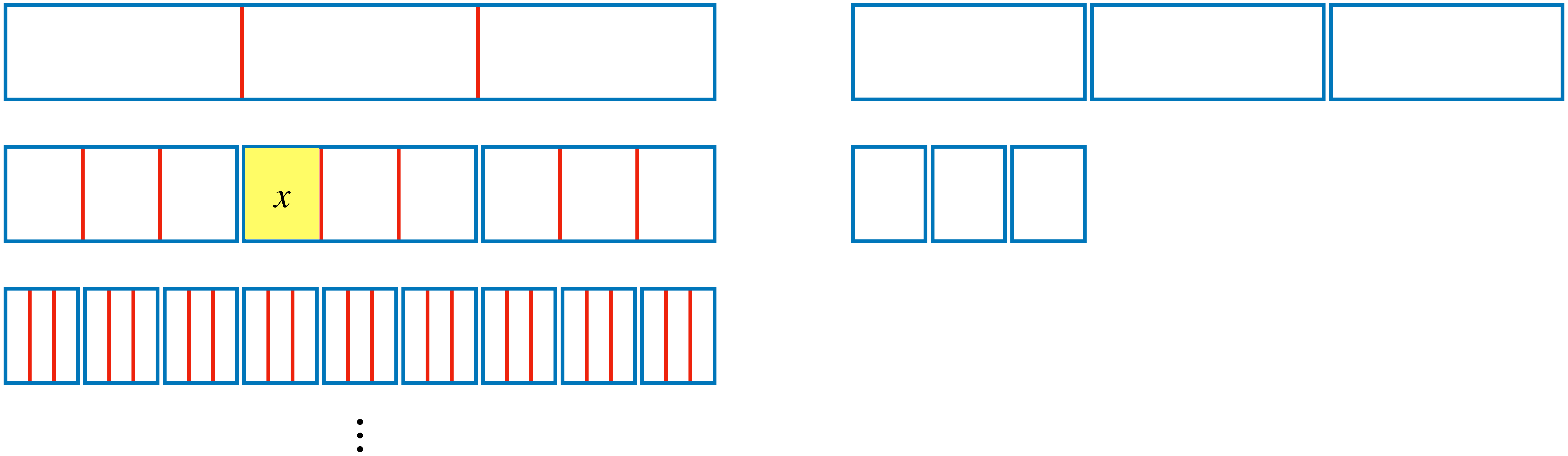
Tree-Swap Swap Regret



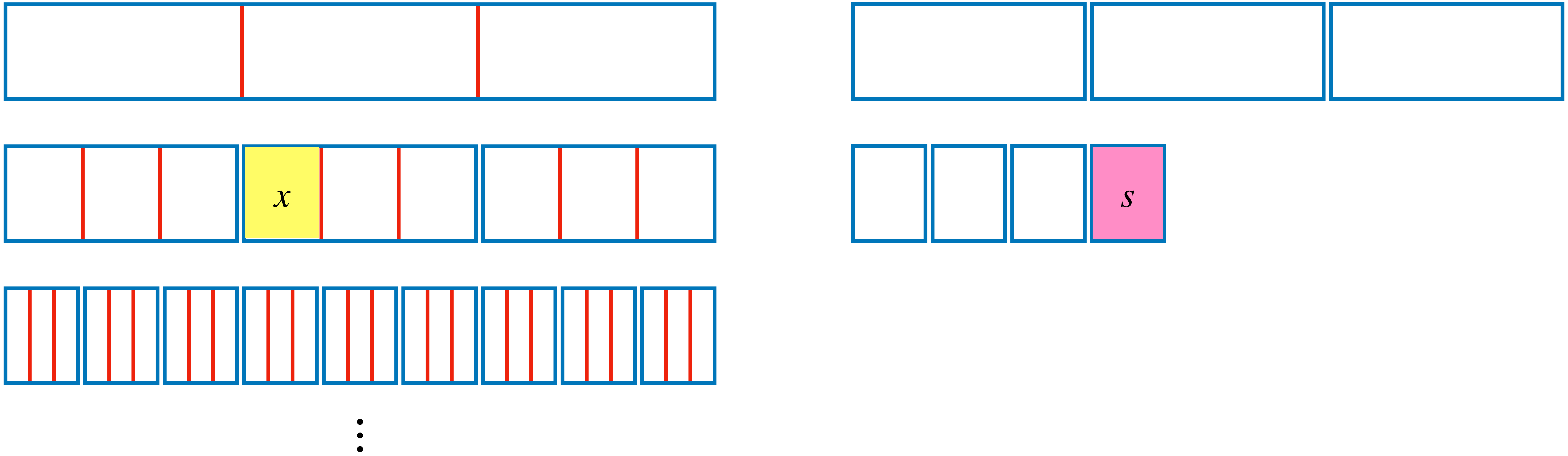
Tree-Swap Swap Regret



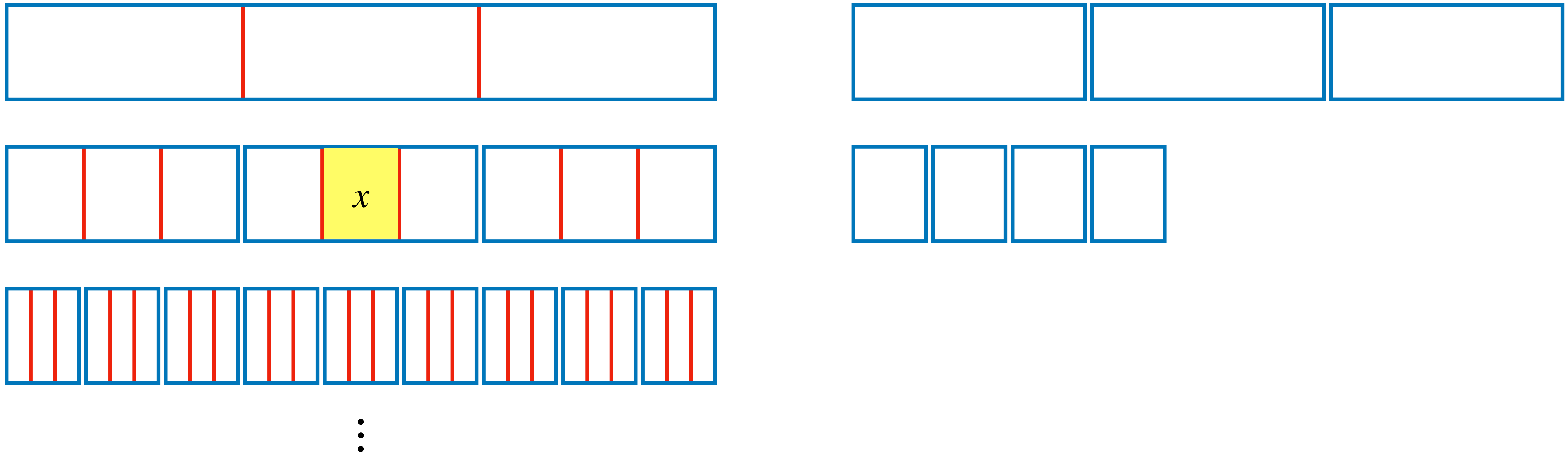
Tree-Swap Swap Regret



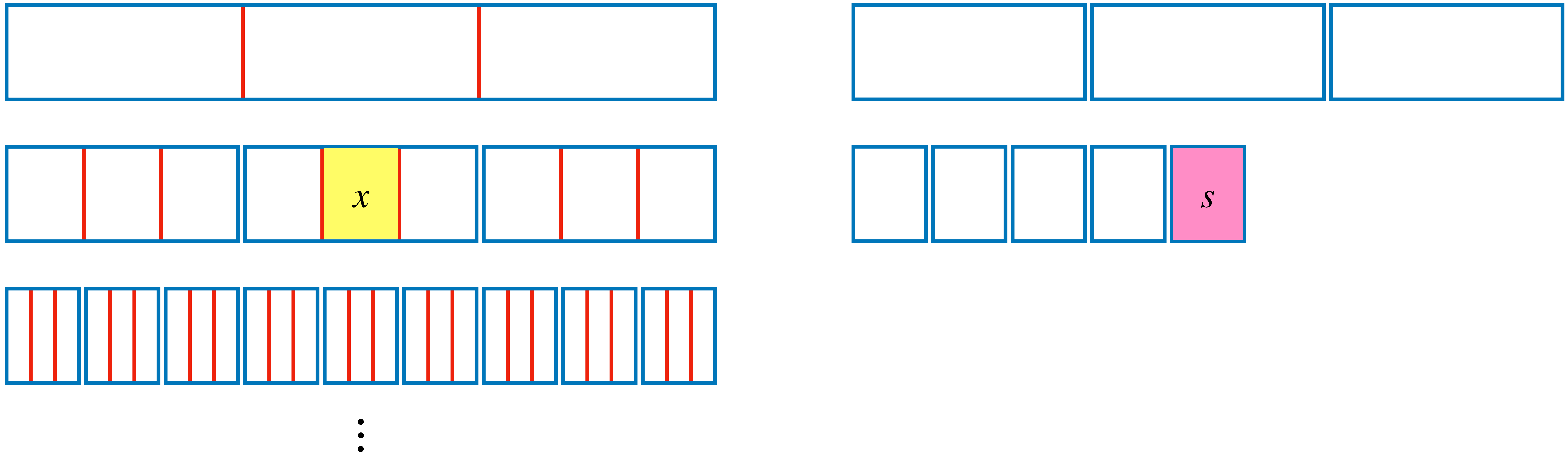
Tree-Swap Swap Regret



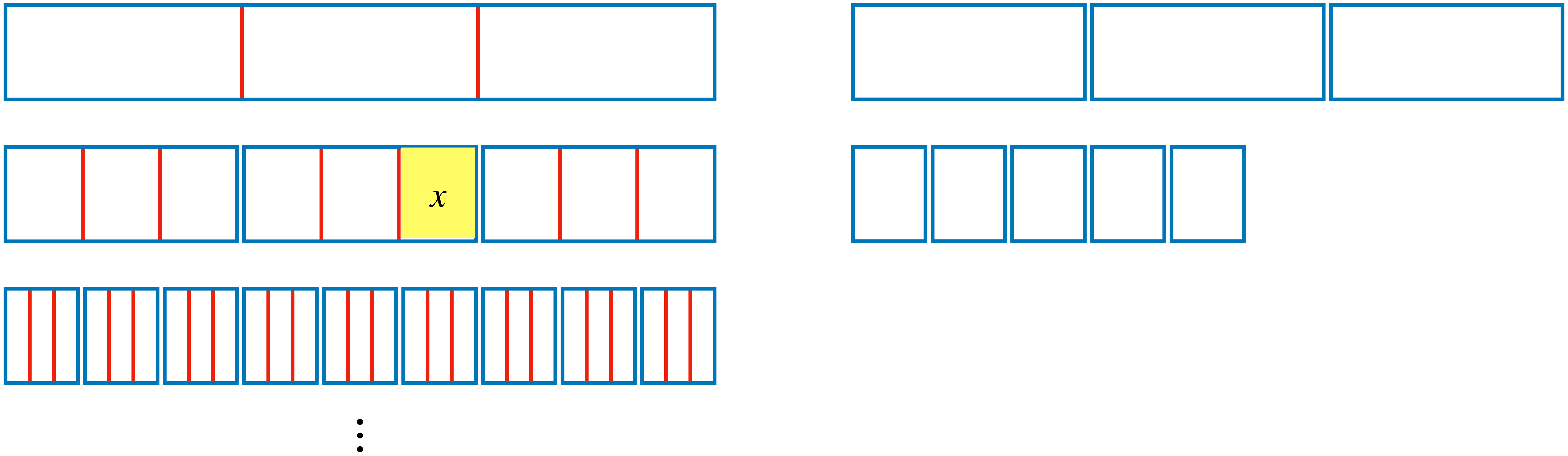
Tree-Swap Swap Regret



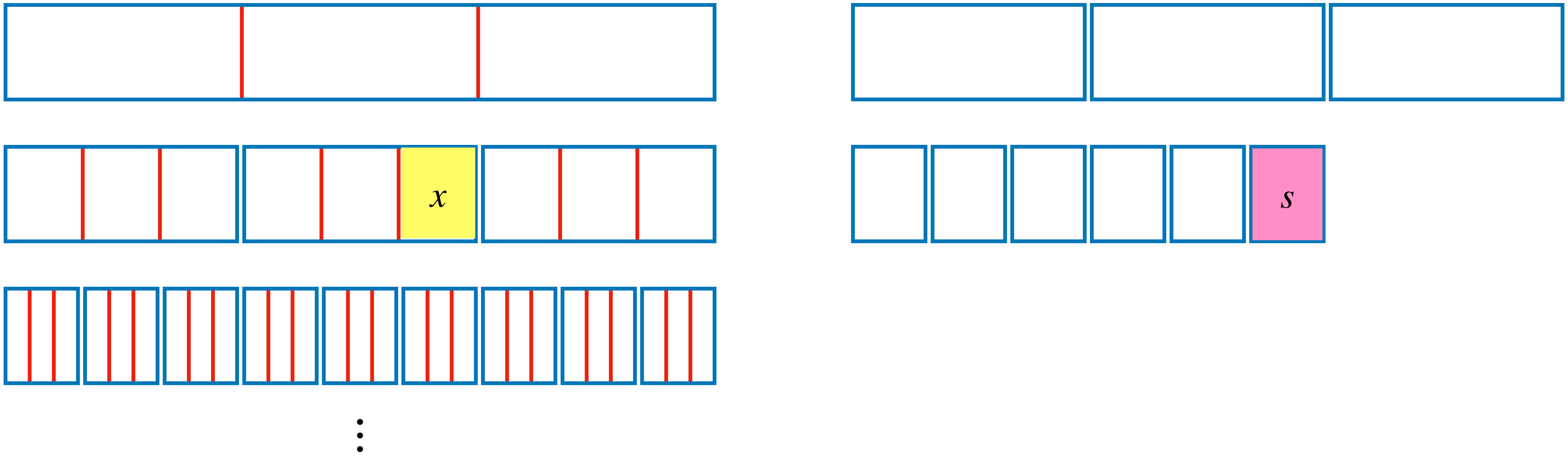
Tree-Swap Swap Regret



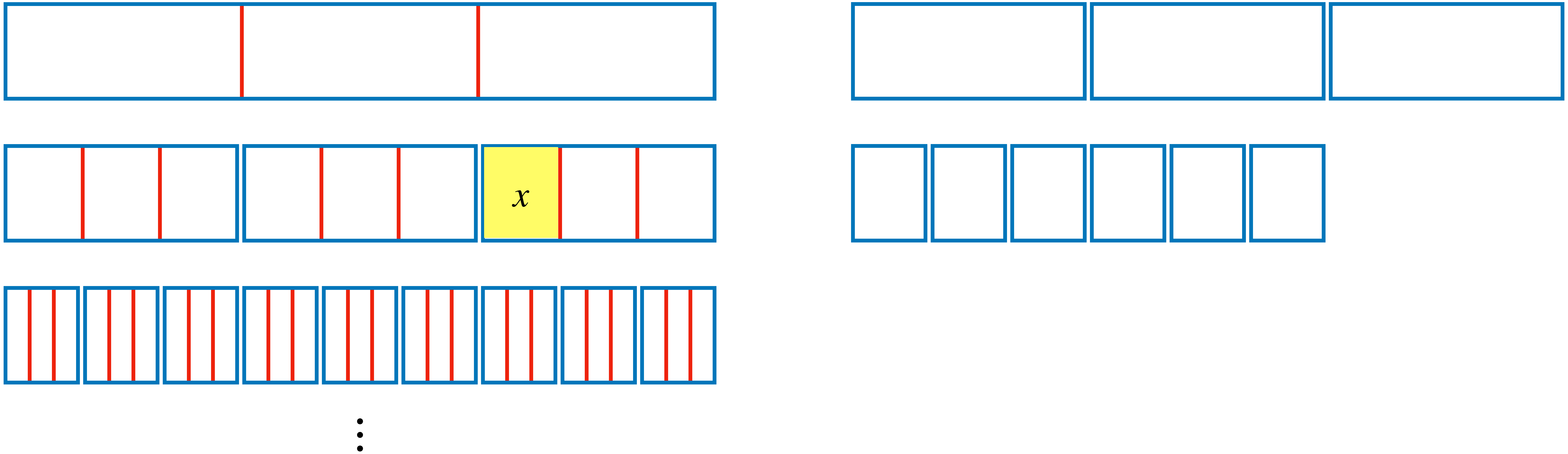
Tree-Swap Swap Regret



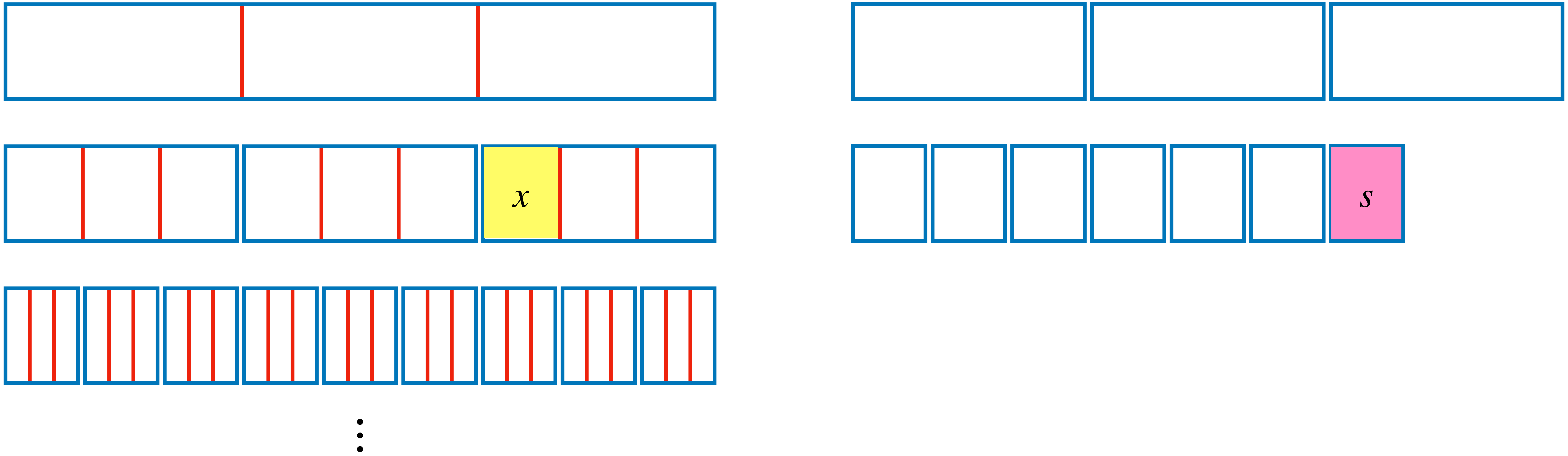
Tree-Swap Swap Regret



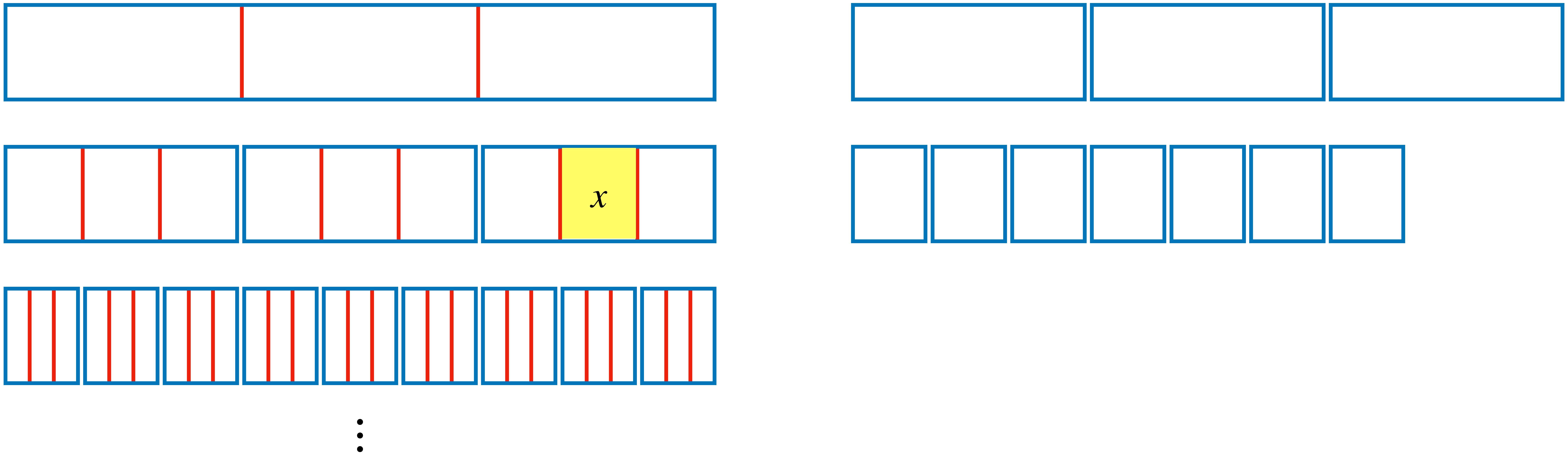
Tree-Swap Swap Regret



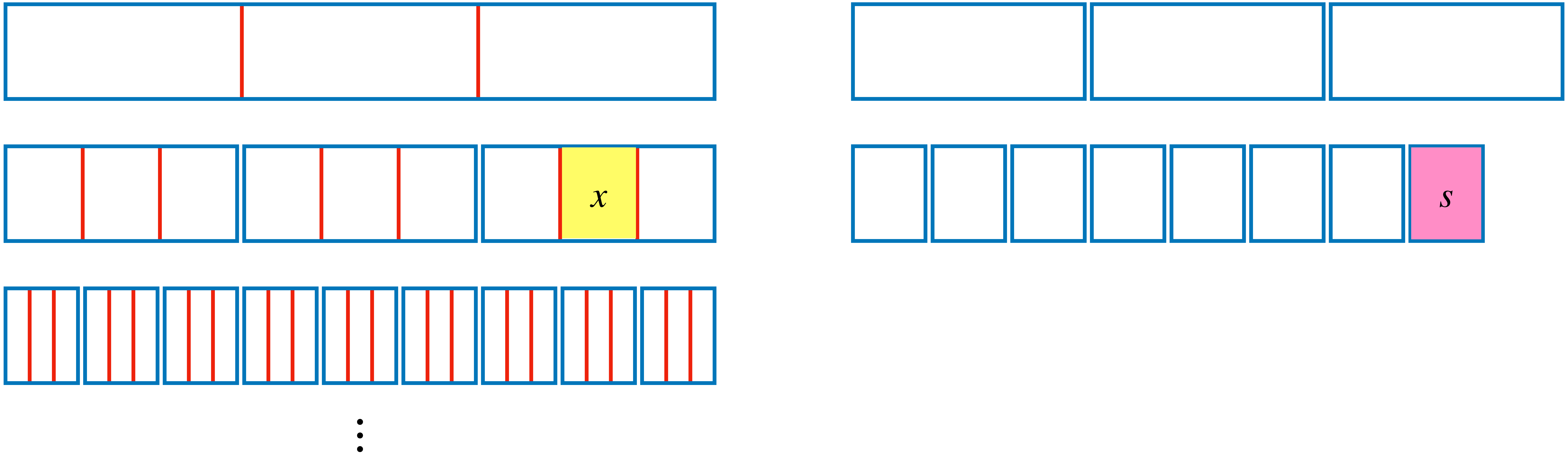
Tree-Swap Swap Regret



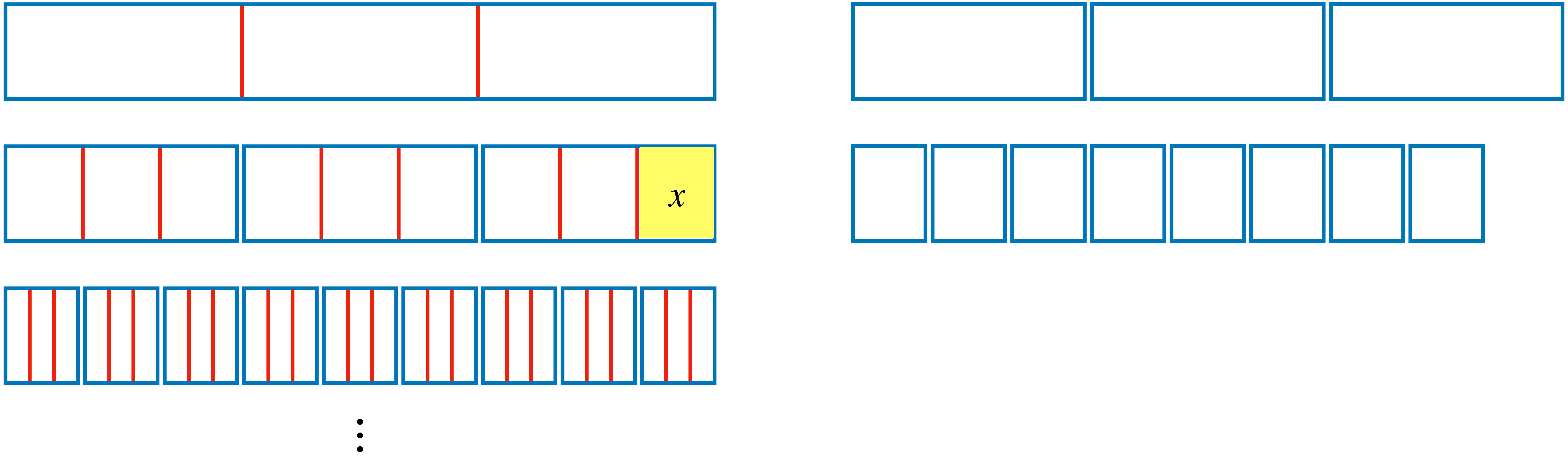
Tree-Swap Swap Regret



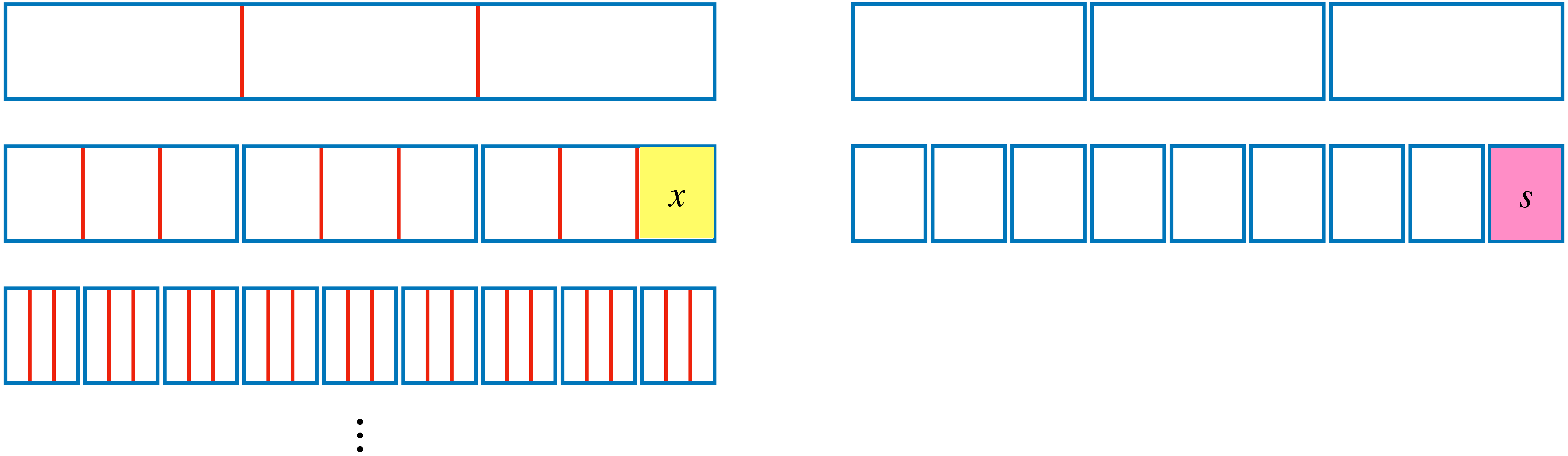
Tree-Swap Swap Regret



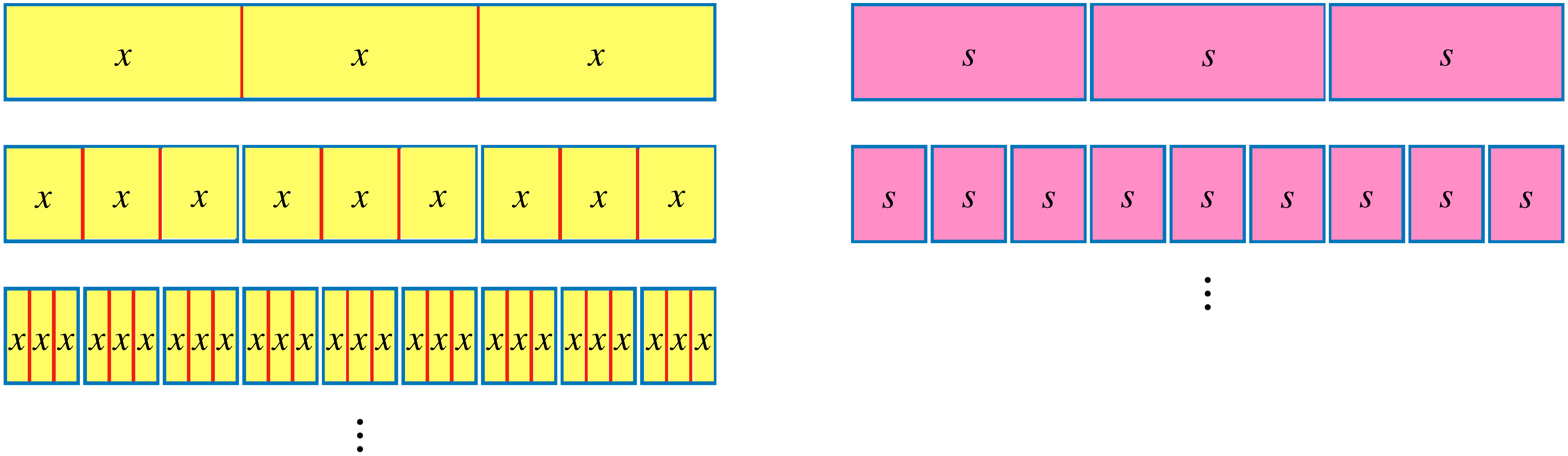
Tree-Swap Swap Regret



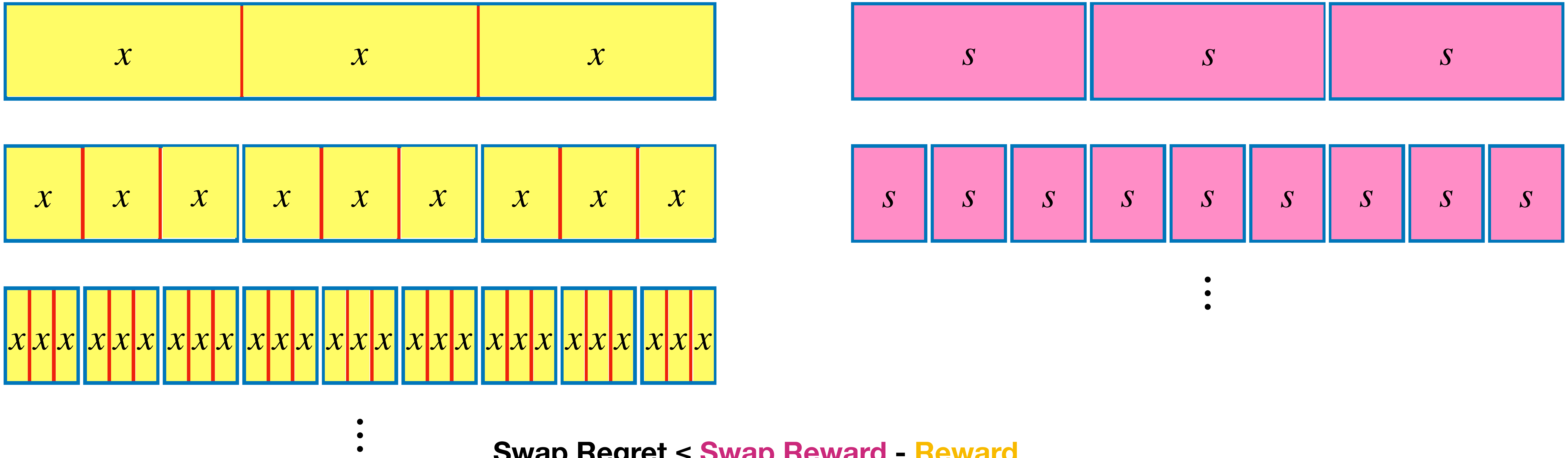
Tree-Swap Swap Regret



Tree-Swap Swap Regret

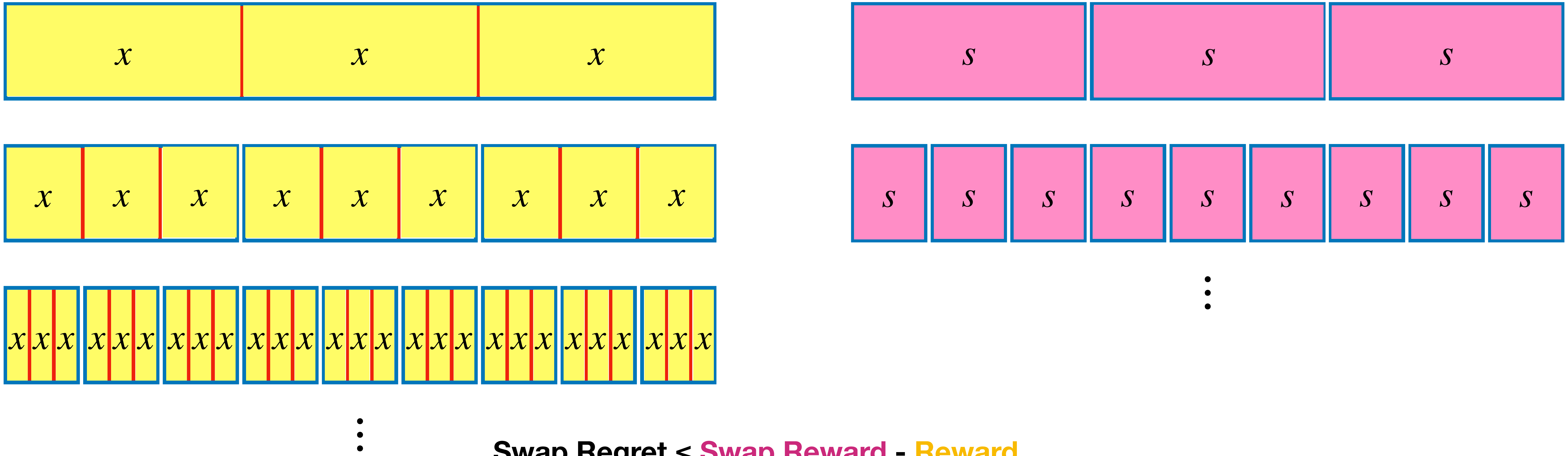


Tree-Swap Swap Regret



Swap Regret ≤ Swap Reward - Reward

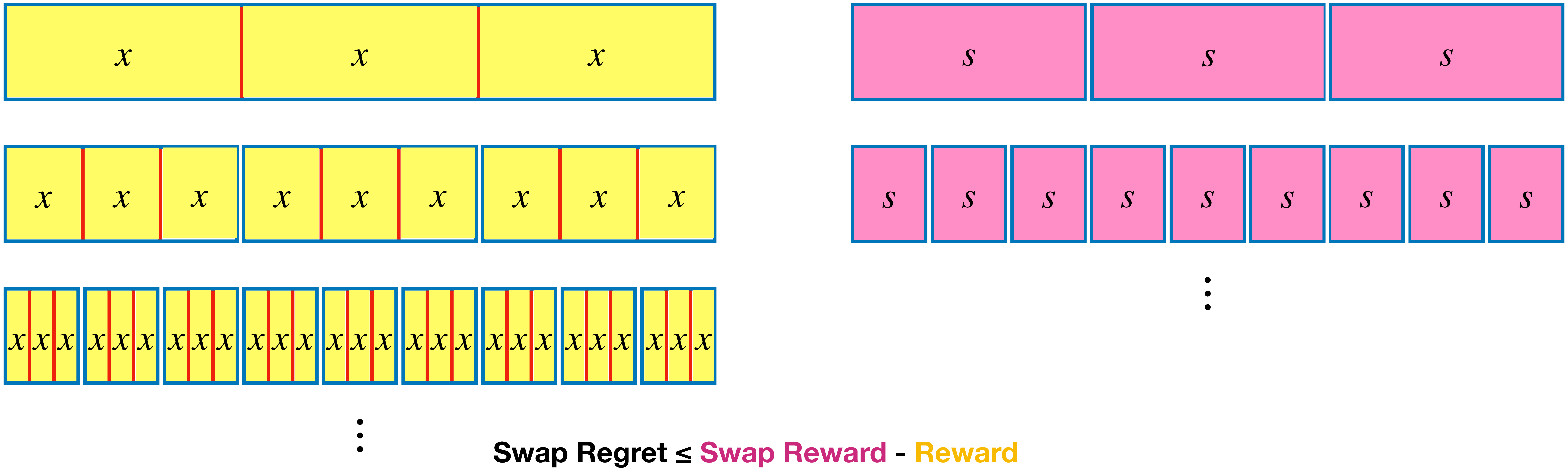
Tree-Swap Swap Regret



Swap Regret ≤ Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

Tree-Swap Swap Regret

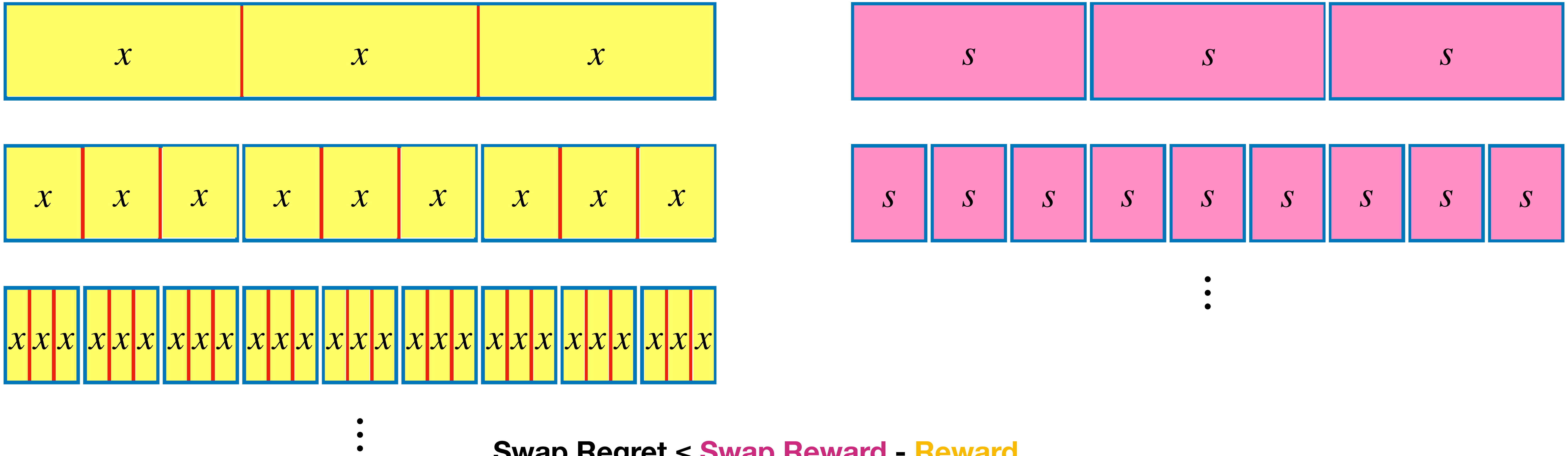


Swap Regret ≤ Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(S_1 - R_2 + S_2 - R_3 + S_3 - R_4 + \cdots)$$

Tree-Swap Swap Regret

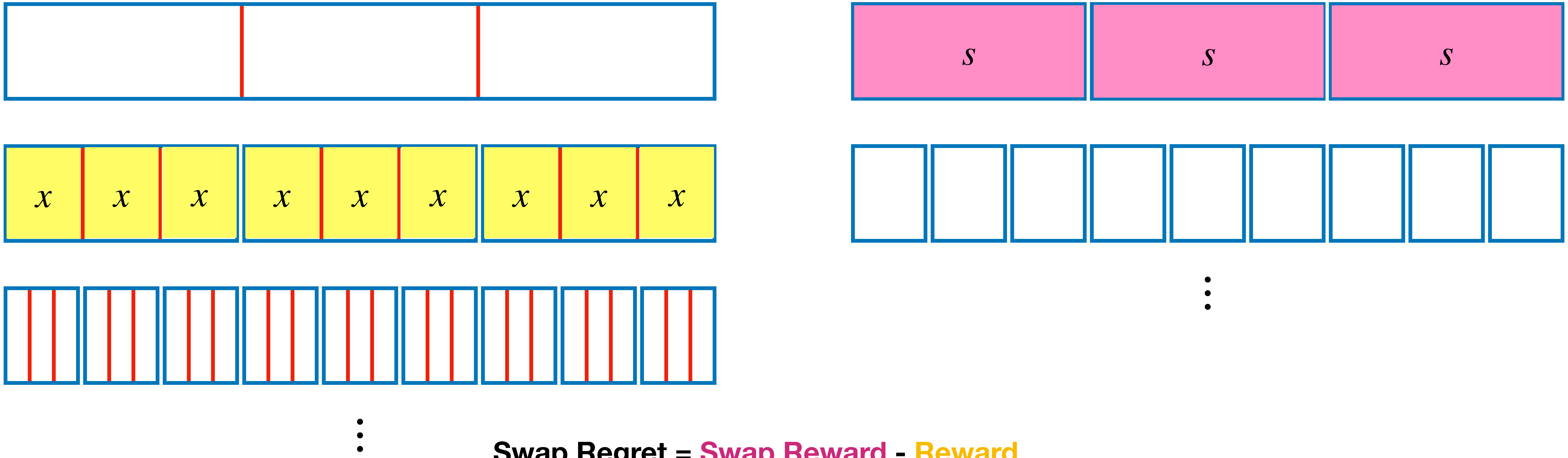


Swap Regret ≤ Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(S_1 - R_2 + S_2 - R_3 + S_3 - R_4 + \cdots)$$

Tree-Swap Swap Regret

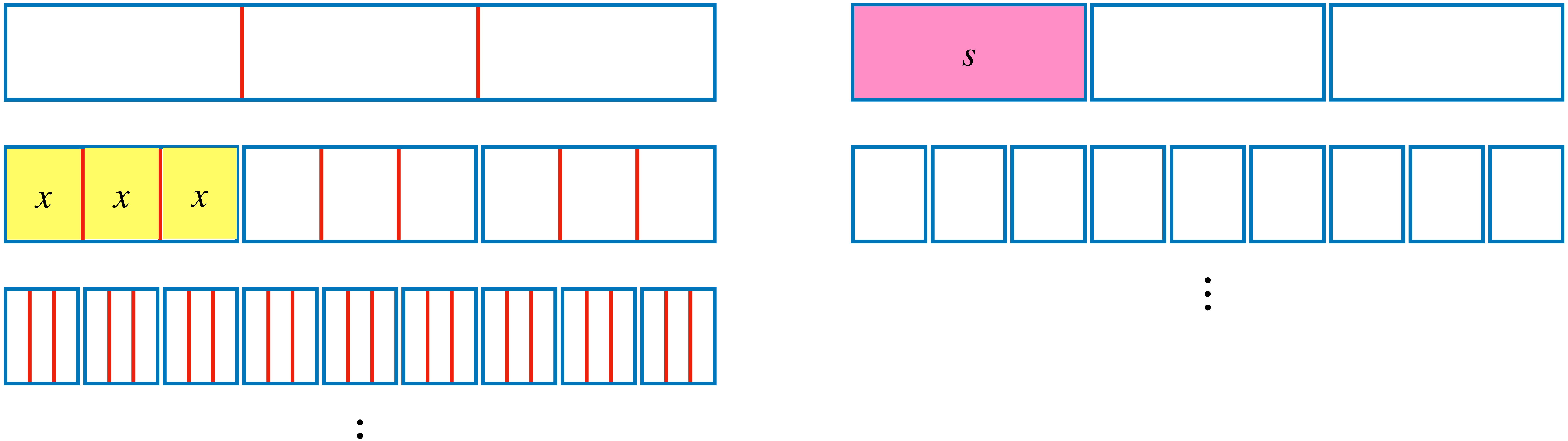


Swap Regret = Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(\boxed{S_1 - R_2} + S_2 - R_3 + S_3 - R_4 + \cdots)$$

Tree-Swap Swap Regret

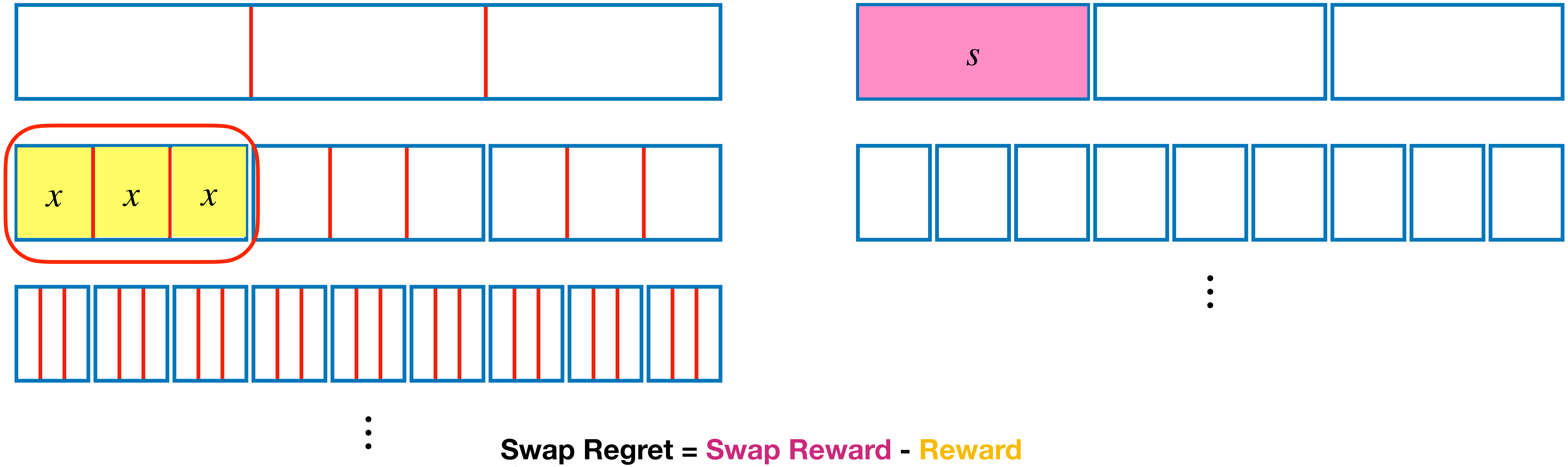


Swap Regret = Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(\boxed{S_1 - R_2} + S_2 - R_3 + S_3 - R_4 + \cdots)$$

Tree-Swap Swap Regret

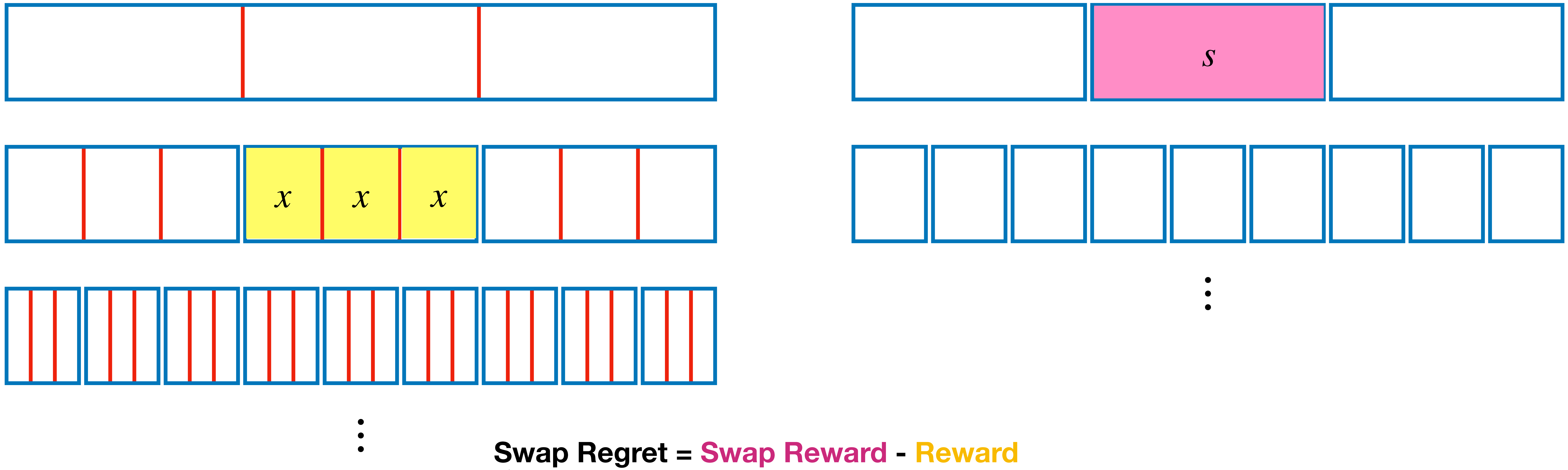


Swap Regret = Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(S_1 - R_2 + S_2 - R_3 + S_3 - R_4 + \cdots)$$

Tree-Swap Swap Regret

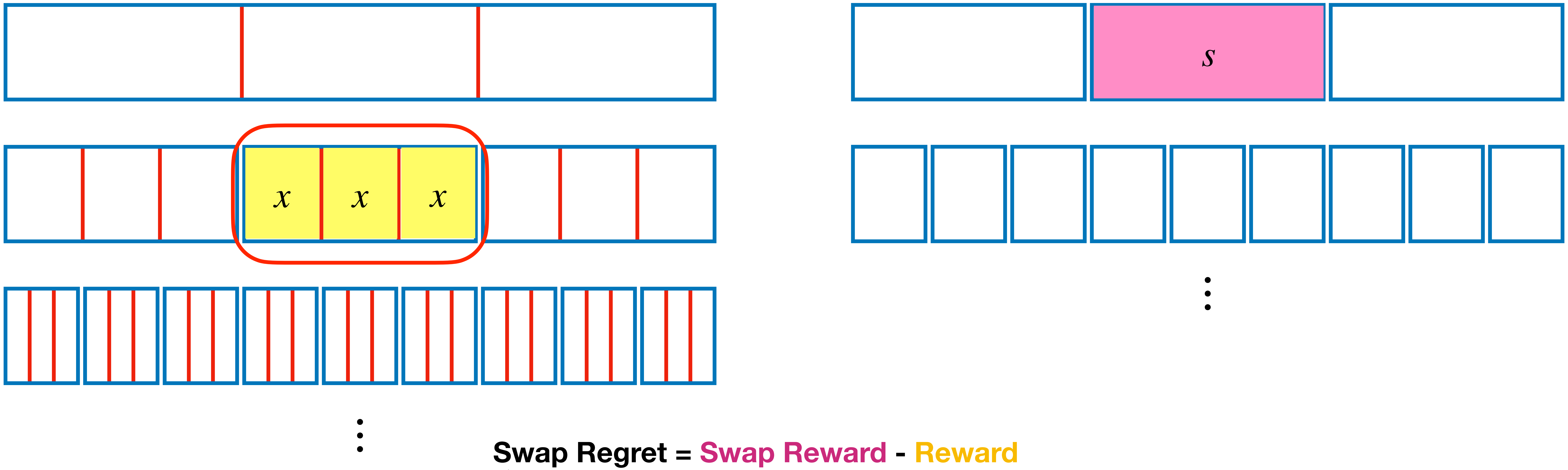


Swap Regret = Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(S_1 - R_2 + S_2 - R_3 + S_3 - R_4 + \cdots)$$

Tree-Swap Swap Regret

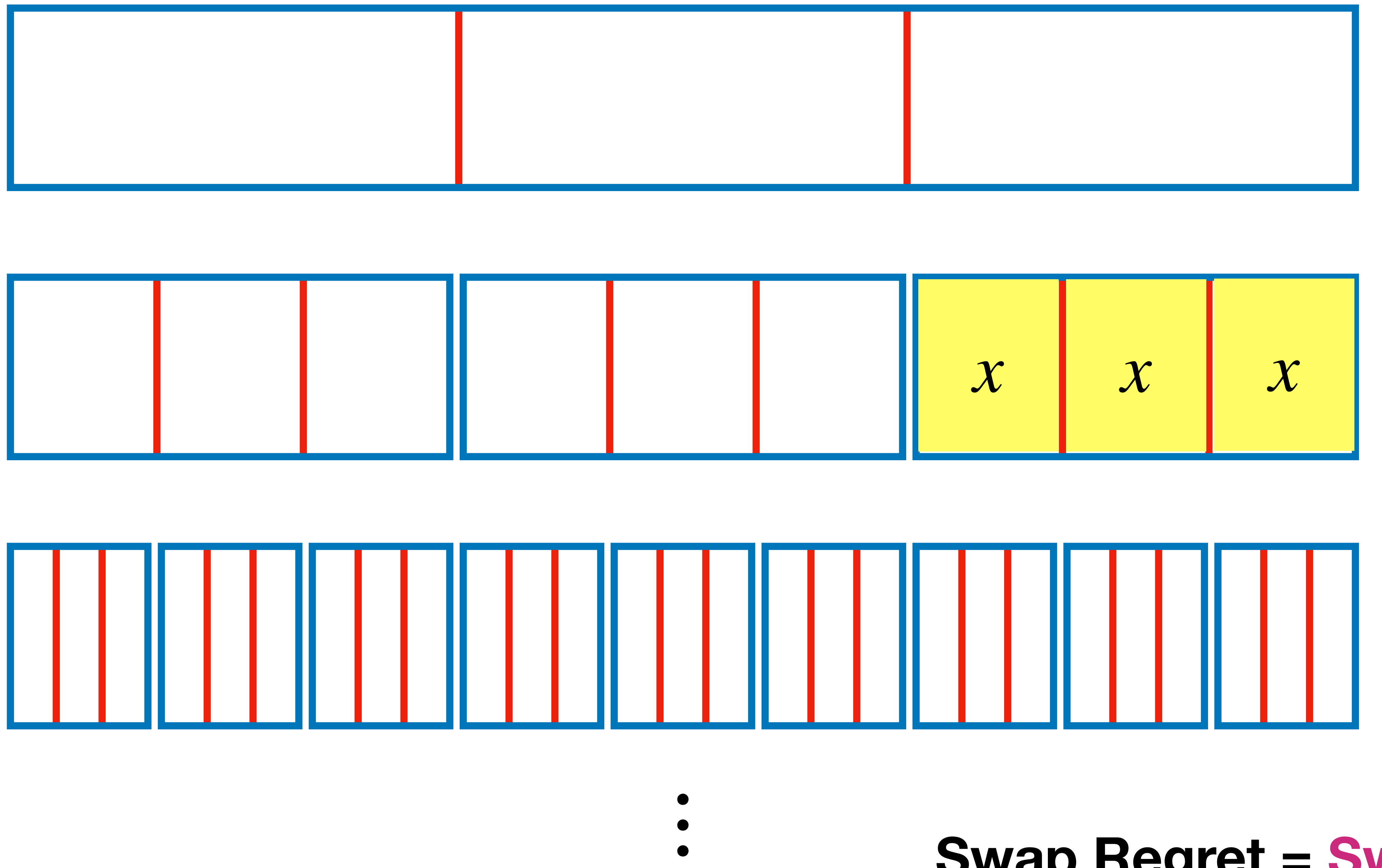


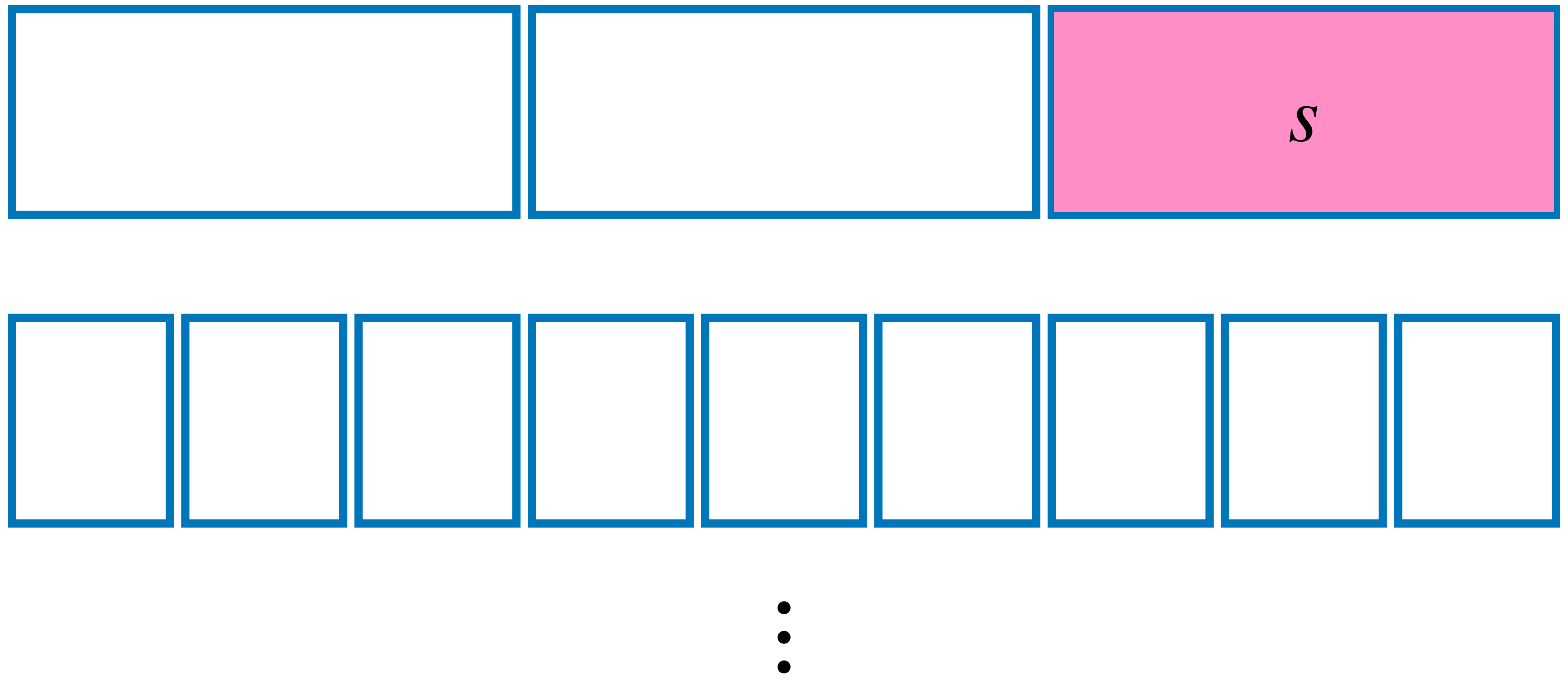
Swap Regret = Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(S_1 - R_2 + S_2 - R_3 + S_3 - R_4 + \cdots)$$

Tree-Swap Swap Regret



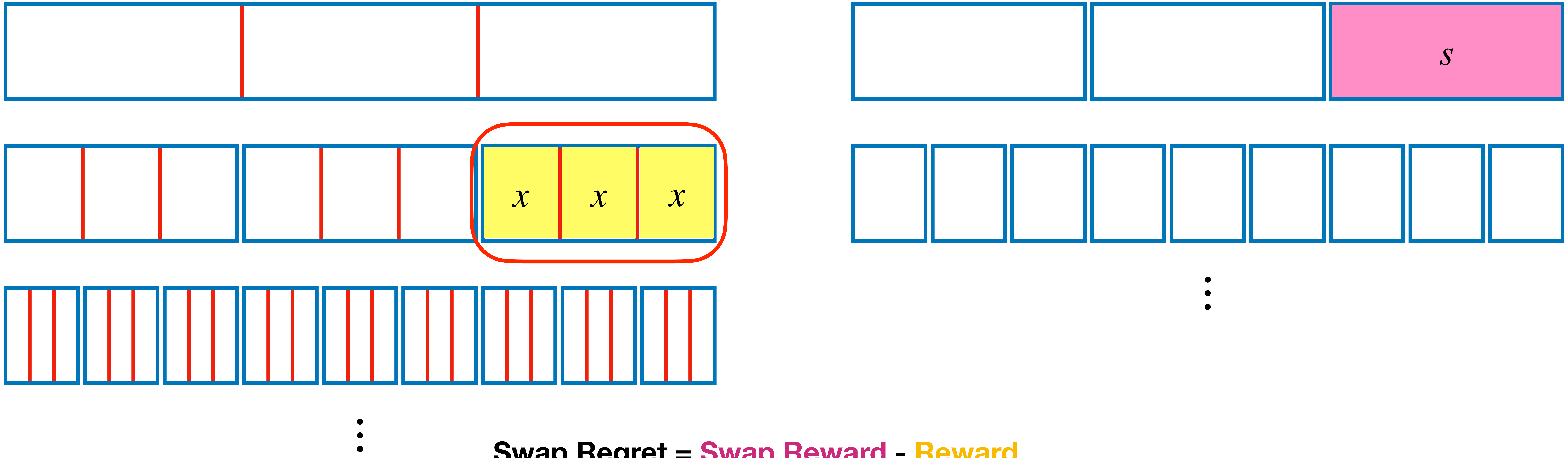


Swap Regret = Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(S_1 - R_2 + S_2 - R_3 + S_3 - R_4 + \cdots)$$

Tree-Swap Swap Regret

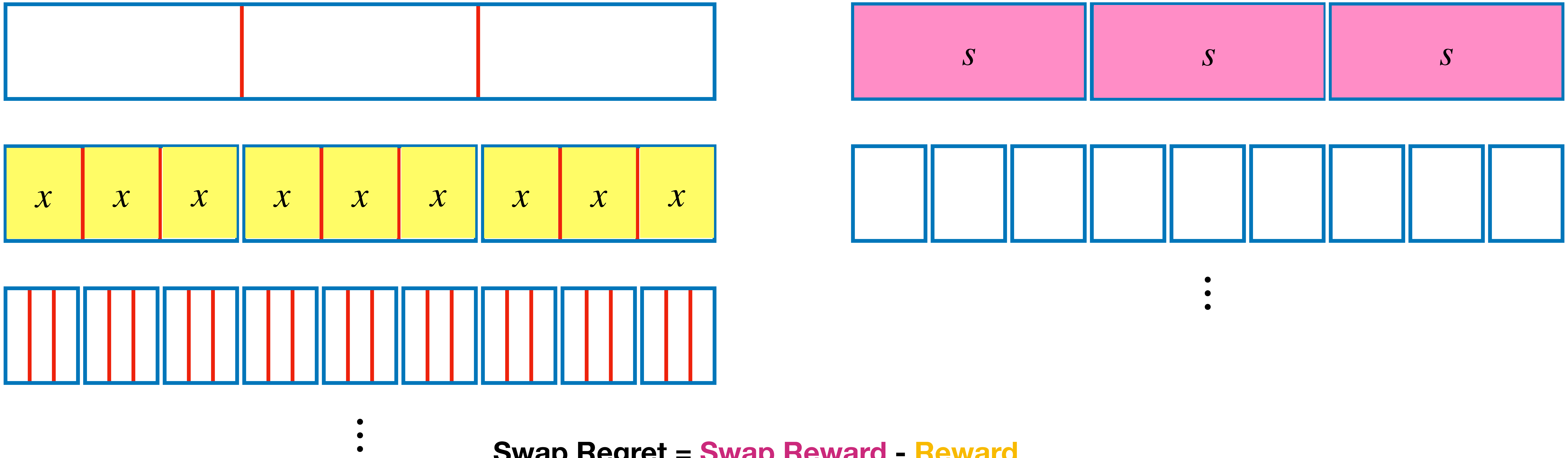


Swap Regret = Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(S_1 - R_2 + S_2 - R_3 + S_3 - R_4 + \cdots)$$

Tree-Swap Swap Regret

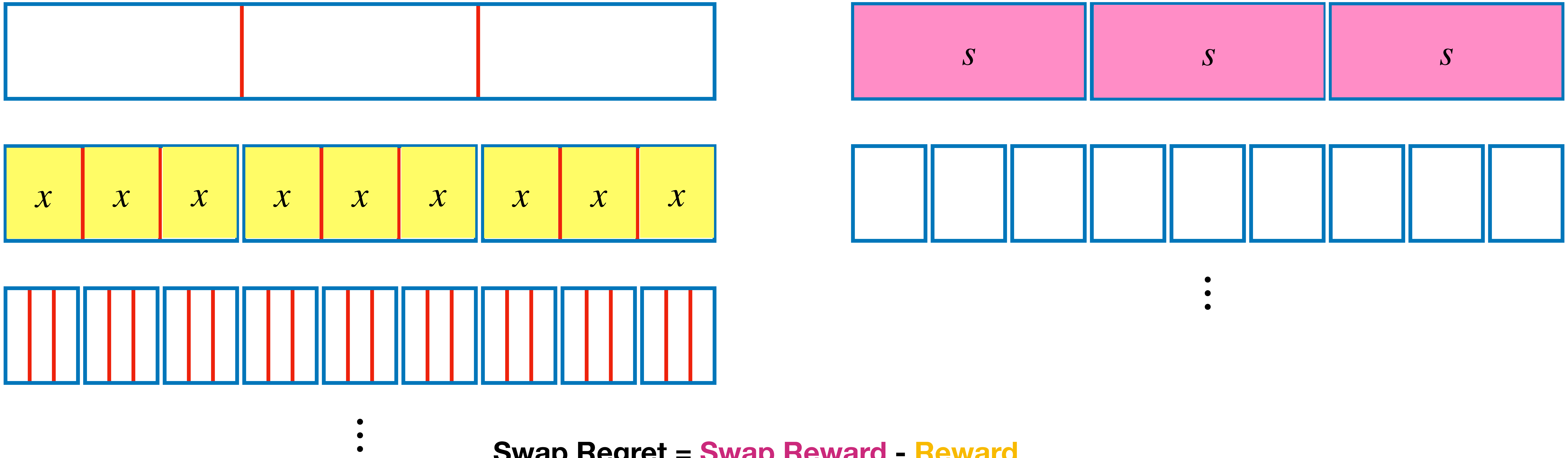


Swap Regret = Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(\boxed{S_1 - R_2} + S_2 - R_3 + S_3 - R_4 + \cdots)$$

Tree-Swap Swap Regret



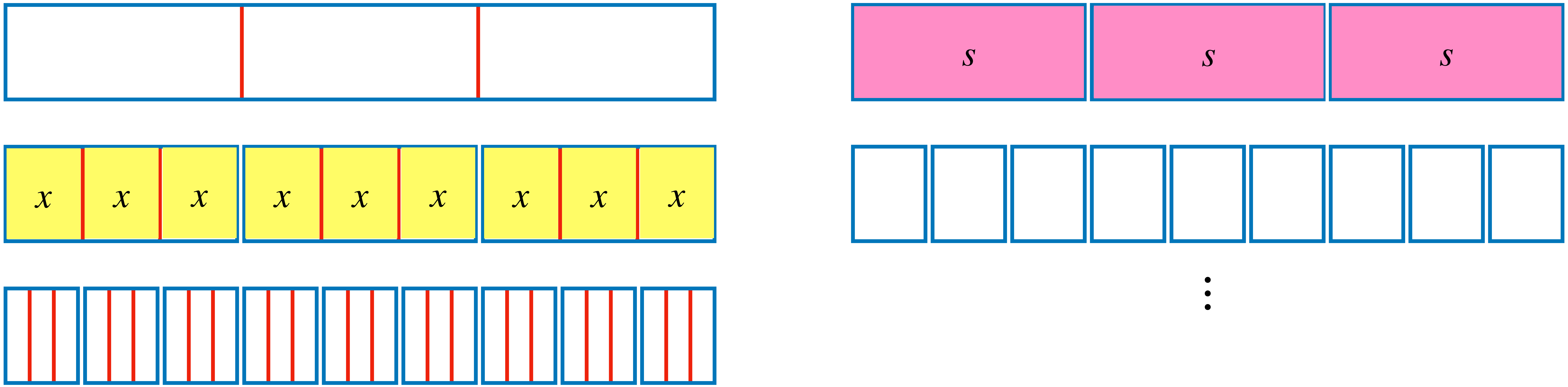
Swap Regret = Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(S_1 - R_2 + S_2 - R_3 + S_3 - R_4 + \cdots)$$

$\leq$ average of external regret of row 2 Lazy-MWU

Tree-Swap Swap Regret



For $M = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

External Regret(Lazy MWU) $\leq \epsilon$

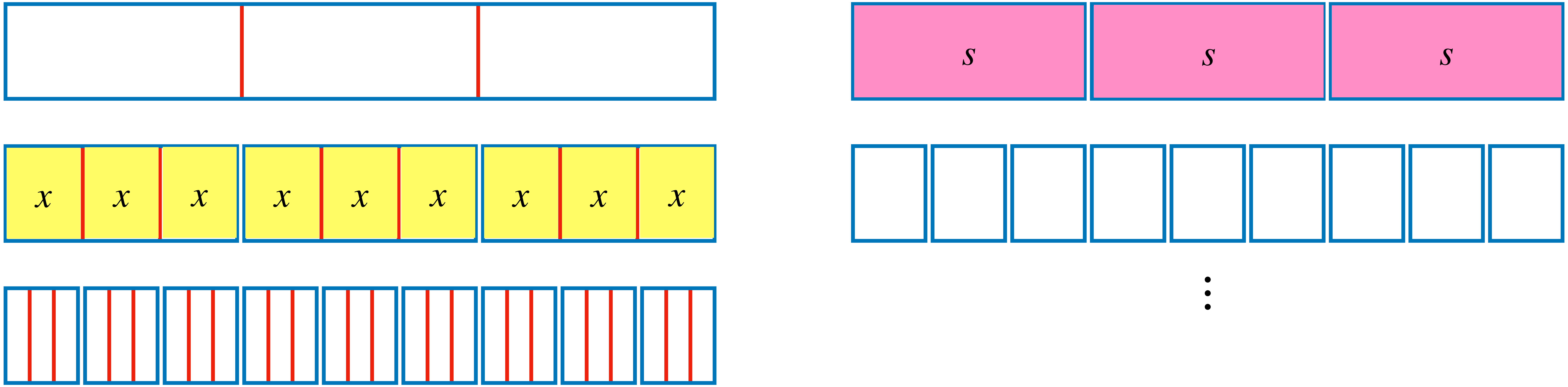
Swap Regret = Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(S_1 - R_2 + S_2 - R_3 + S_3 - R_4 + \cdots)$$

$\leq$ average of external regret of row 2 Lazy-MWU

Tree-Swap Swap Regret



For $M = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

External Regret(Lazy MWU) $\leq \epsilon$

Swap Regret = Swap Reward - Reward

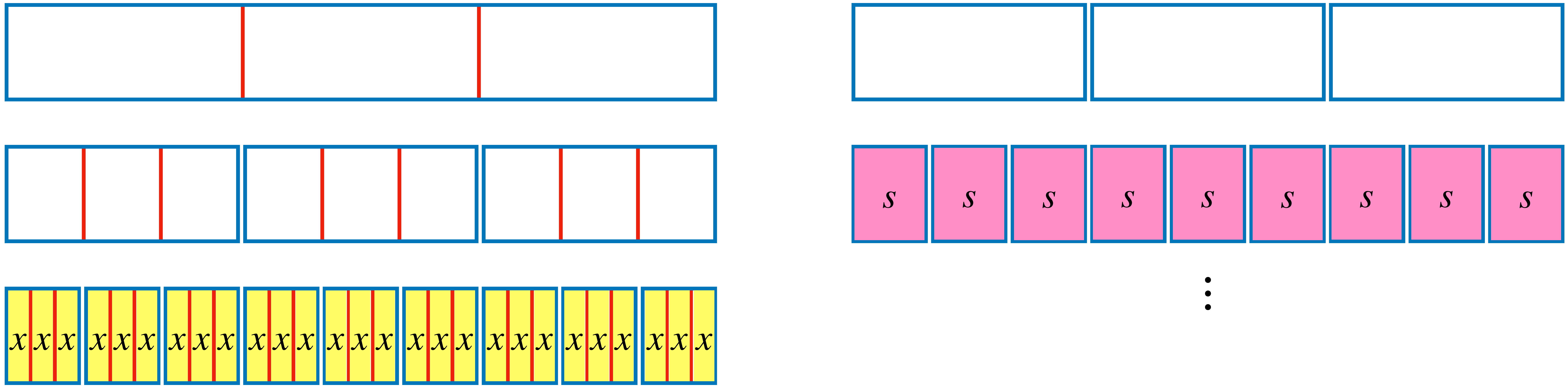
$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(S_1 - R_2 + S_2 - R_3 + S_3 - R_4 + \cdots)$$

$\leq$ average of external regret of row 2 Lazy-MWU

$\leq \epsilon$

Tree-Swap Swap Regret



For $M = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

External Regret(Lazy MWU) $\leq \epsilon$

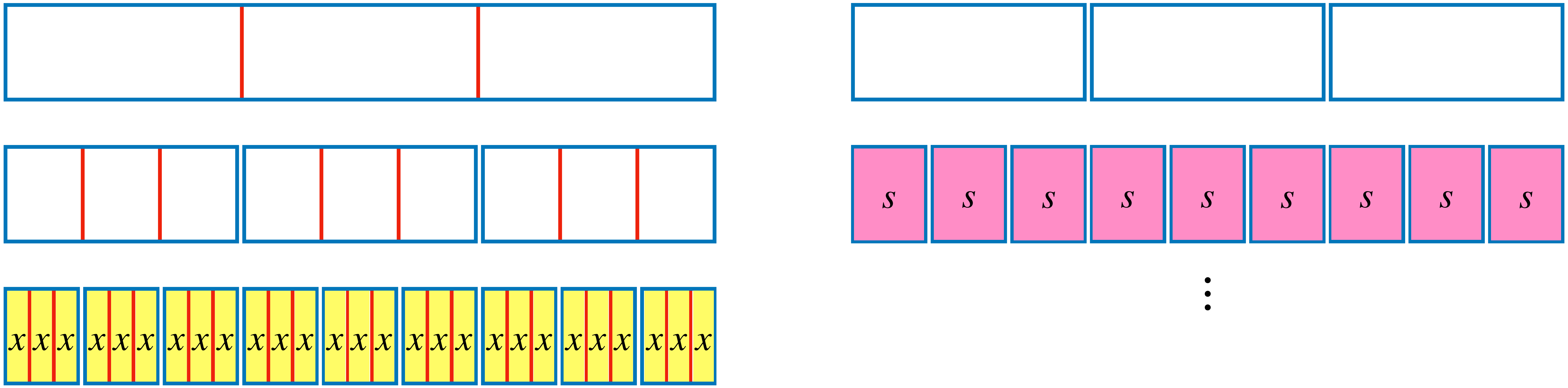
Swap Regret = Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(S_1 - R_2 + S_2 - R_3 + S_3 - R_4 + \cdots)$$

$$\leq \epsilon$$

Tree-Swap Swap Regret



For $M = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

External Regret(Lazy MWU) $\leq \epsilon$

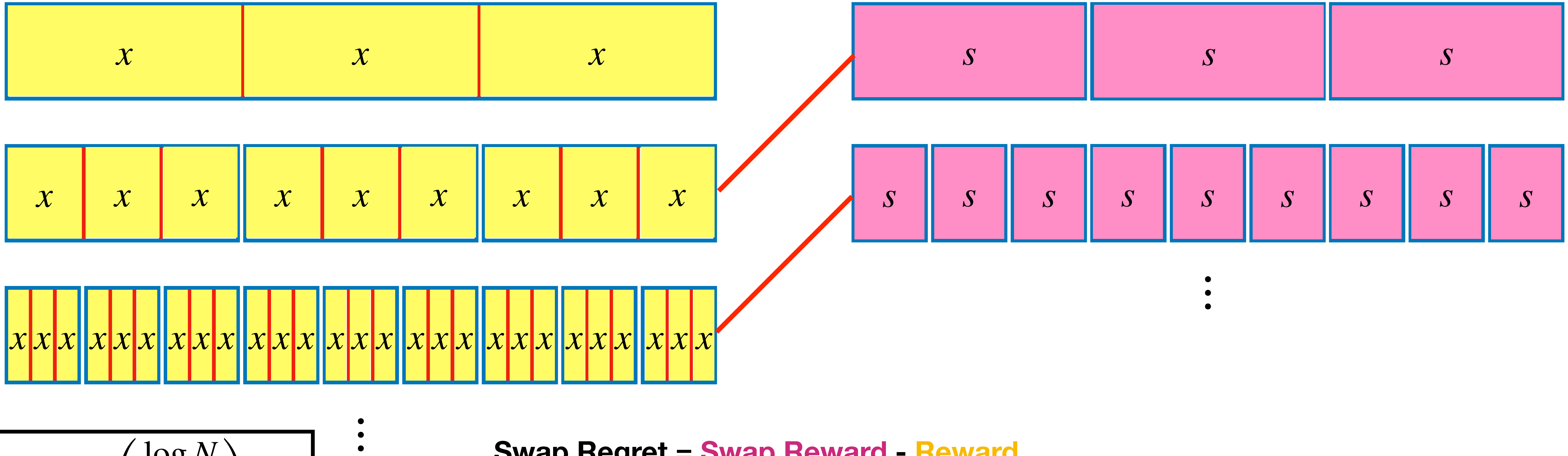
Swap Regret = **Swap Reward** - **Reward**

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(S_1 - R_2 + \boxed{S_2 - R_3} + S_3 - R_4 + \cdots)$$

$\leq \epsilon$ $\leq \epsilon$

Tree-Swap Swap Regret



For $M = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

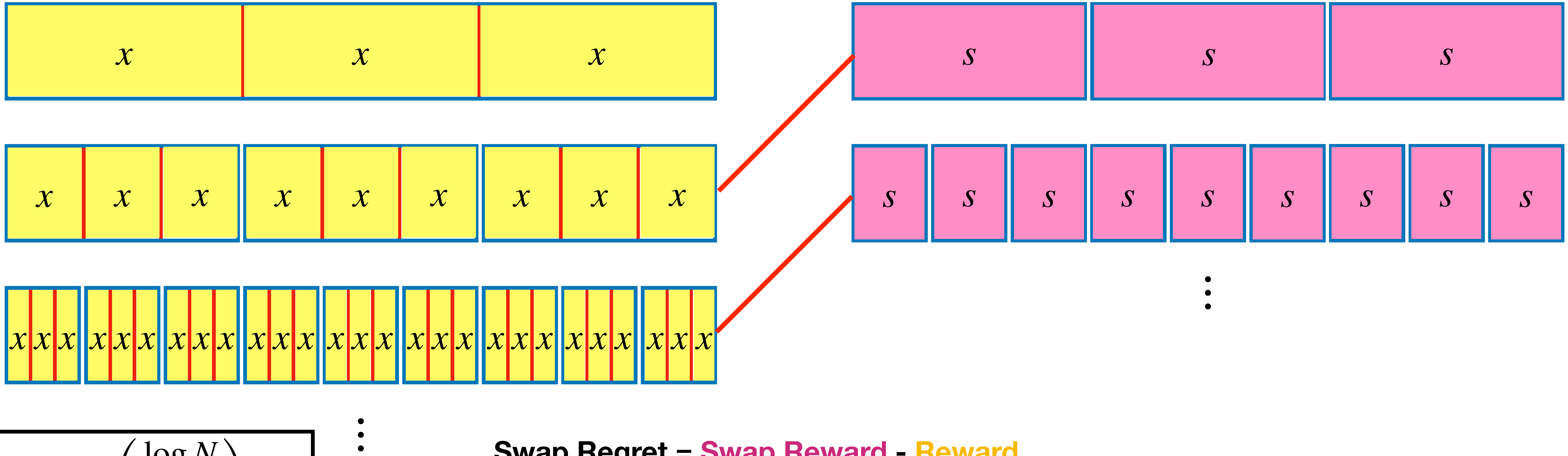
External Regret(Lazy MWU) $\leq \epsilon$

Swap Regret = Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(S_1 - R_2 + S_2 - R_3 + S_3 - R_4 + \cdots)$$

Tree-Swap Swap Regret



For $M = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

External Regret(Lazy MWU) $\leq \epsilon$

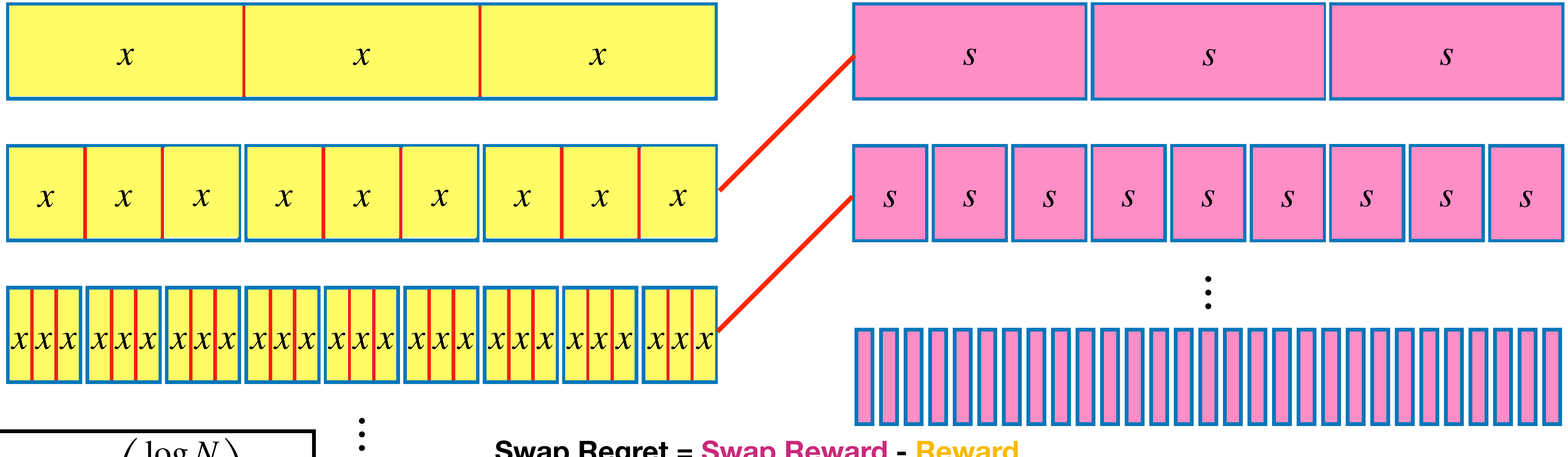
Swap Regret = Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(S_1 - R_2 + S_2 - R_3 + S_3 - R_4 + \cdots)$$

$$\leq \frac{1}{d}(\epsilon + \epsilon + \cdots S_d)$$

Tree-Swap Swap Regret



For $M = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

External Regret(Lazy MWU) $\leq \epsilon$

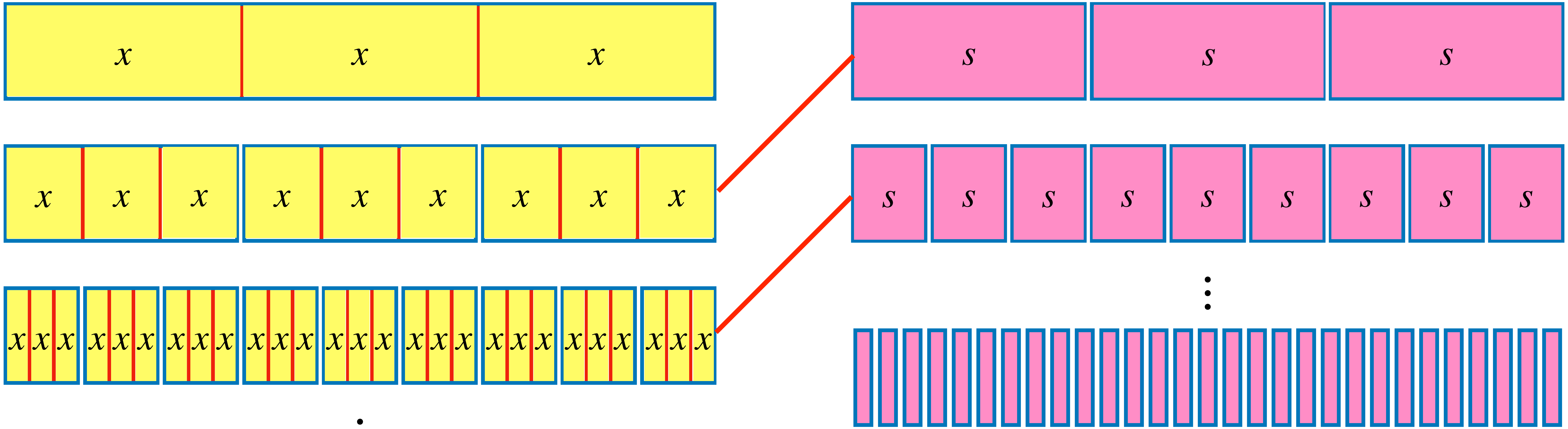
Swap Regret = Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(S_1 - R_2 + S_2 - R_3 + S_3 - R_4 + \cdots)$$

$$\leq \frac{1}{d}(\epsilon + \epsilon + \cdots S_d)$$

Tree-Swap Swap Regret



For $M = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

External Regret(Lazy MWU) $\le \epsilon$

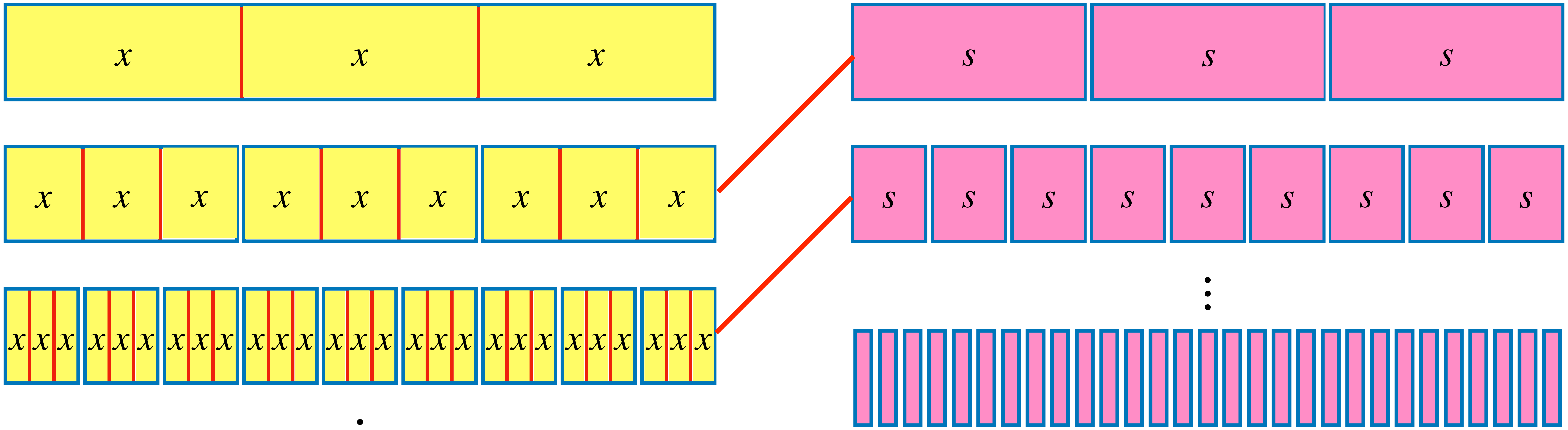
Swap Regret = Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(S_1 - R_2 + S_2 - R_3 + S_3 - R_4 + \cdots)$$

$$\le \frac{1}{d}(\epsilon + \epsilon + \cdots S_d) \le \epsilon + \frac{1}{d}$$

Tree-Swap Swap Regret



For $M = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

External Regret(Lazy MWU) $\leq \epsilon$

For $d \geq \frac{1}{\epsilon}$

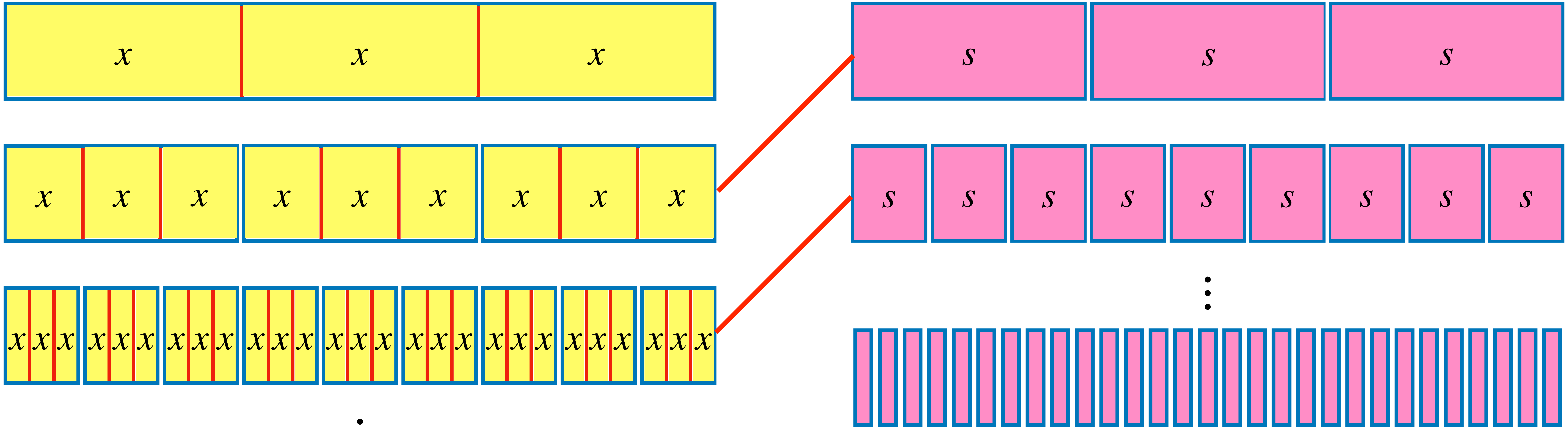
Swap Regret = Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(S_1 - R_2 + S_2 - R_3 + S_3 - R_4 + \cdots)$$

$$\leq \frac{1}{d}(\epsilon + \epsilon + \cdots S_d) \leq \epsilon + \frac{1}{d}$$

Tree-Swap Swap Regret



For $M = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

External Regret(Lazy MWU) $\leq \epsilon$

For $d \geq \frac{1}{\epsilon}$

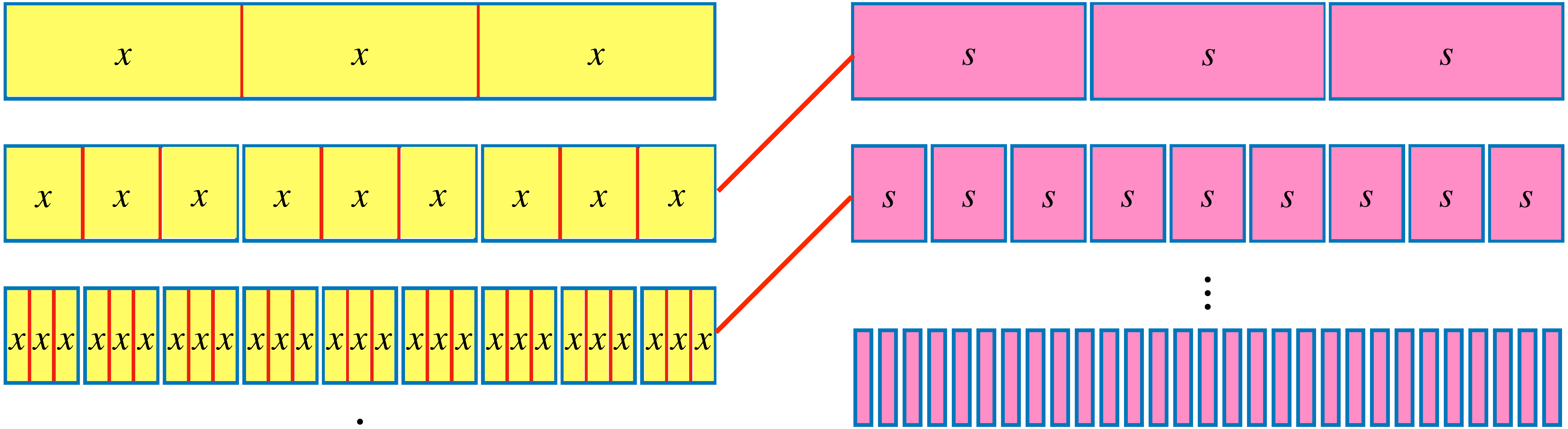
Swap Regret = Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(S_1 - R_2 + S_2 - R_3 + S_3 - R_4 + \cdots)$$

$$\leq \frac{1}{d}(\epsilon + \epsilon + \cdots S_d) \leq \epsilon + \frac{1}{d} \leq 2\epsilon$$

Tree-Swap Swap Regret



For $M = \Omega\left(\frac{\log N}{\epsilon^2}\right)$

External Regret(Lazy MWU) $\leq \epsilon$

For $d \geq \frac{1}{\epsilon}$

Swap Regret = Swap Reward - Reward

$$= \frac{1}{d}(S_1 + S_2 + \cdots + S_d - R_1 + R_2 + \cdots + R_d)$$

$$= \frac{1}{d}(S_1 - R_2 + S_2 - R_3 + S_3 - R_4 + \cdots)$$

$$\leq \frac{1}{d}(\epsilon + \epsilon + \cdots S_d) \leq \epsilon + \frac{1}{d} \leq 2\epsilon$$

For $T = M^d = \left(\frac{\log N}{\epsilon^2}\right)^{\Omega(1/\epsilon)}$

Swap Regret(Tree-Swap) $\leq \epsilon$

Thanks for listening!

Questions? maxfish@mit.edu

Intuition? maxkfish.com

References:

[DD🐟G23] **https://arxiv.org/pdf/2310.19786.pdf**
[PR23] **https://arxiv.org/pdf/2310.19647.pdf**
[BM07] **https://www.jmlr.org/papers/volume8/blum07a/blum07a.pdf**