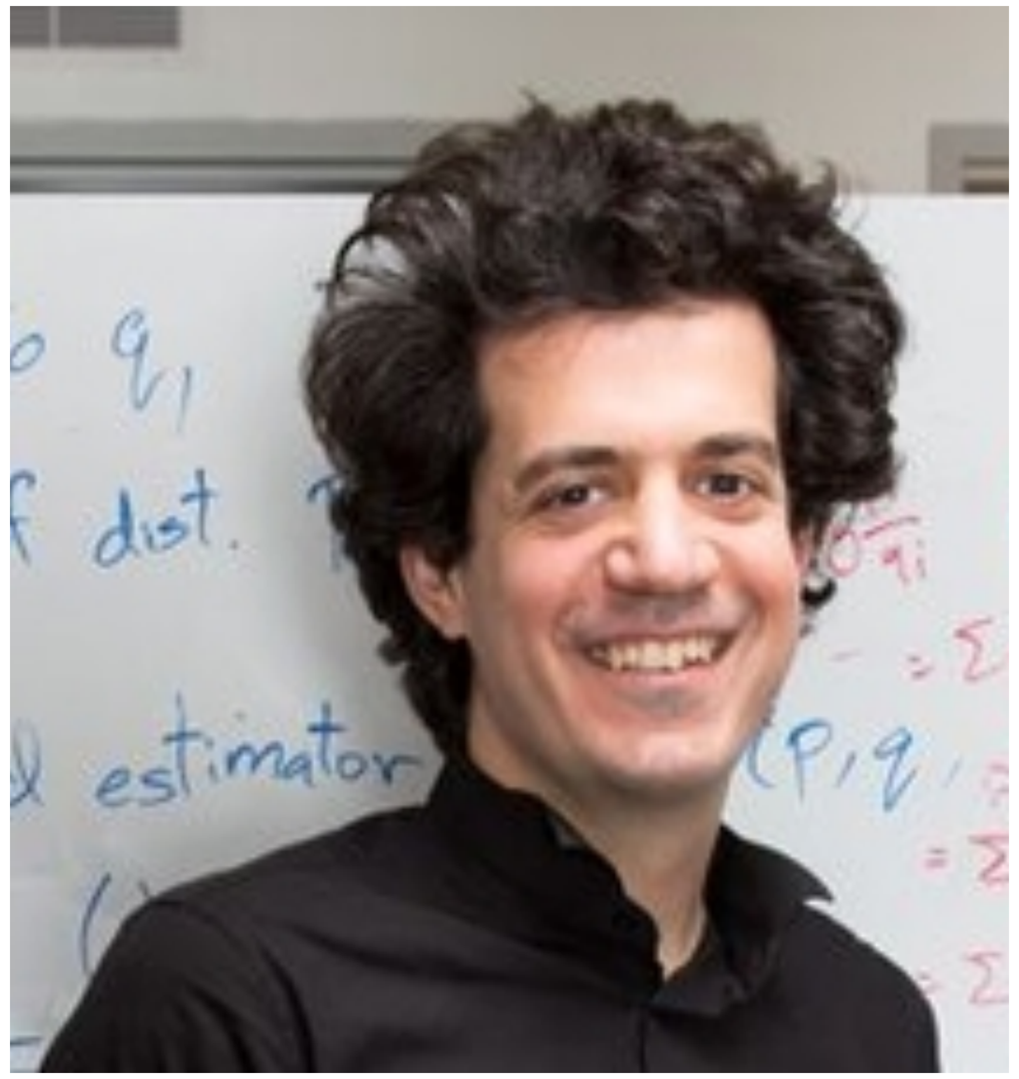
From External to Swap Regret 2.0: An Efficient Reduction for Large Action Spaces

Fast Swap Regret Minimization and Applications to Approximate Correlated Equilibria

Presentation by: Maxwell Fishelson

Yuval Dagan

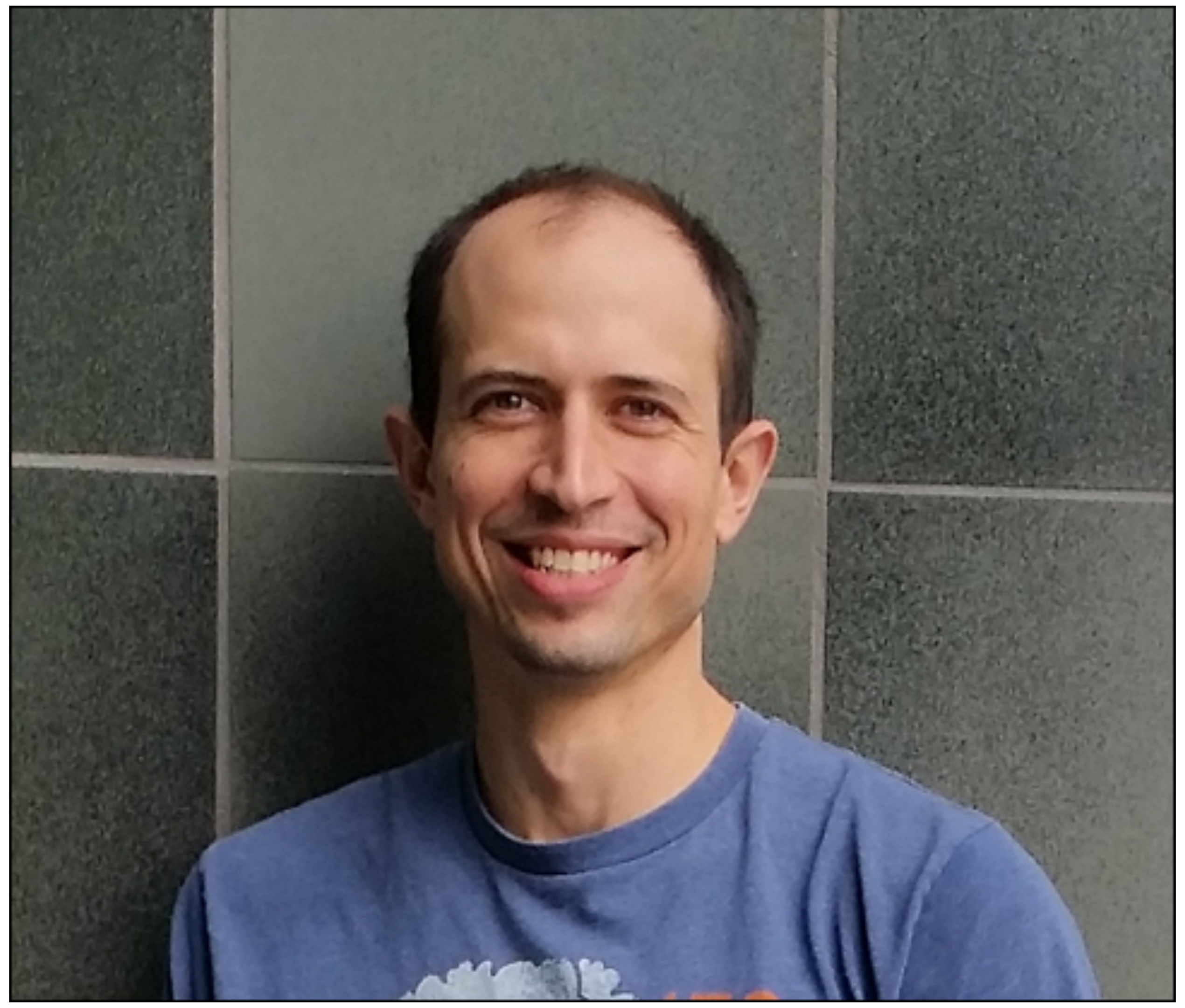


Costis Daskalakis

Maxwell Fishelson

Noah Golowich

Binghui Peng



Aviad Rubinstein

Presentation by: Maxwell Fishelson

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

$$x^{(t)} \in \Delta(N)$$

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

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

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

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

Regret

regret = reward obtained vs benchmark

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

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

Regret

regret = reward obtained vs benchmark

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

ϕ $\longrightarrow$

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

Regret

regret = reward obtained vs benchmark

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

$\xrightarrow{\phi}$

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

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

Regret

regret = reward obtained vs benchmark

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

$$\xrightarrow{\phi}$$

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

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

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

Regret

regret = reward obtained vs benchmark

$$\begin{matrix} x^{(1)} & u^{(1)} \\ x^{(2)} & u^{(2)} \\ x^{(3)} & u^{(3)} \\ \vdots & \\ x^{(T)} & u^{(T)} \end{matrix} \quad \xrightarrow{\ \phi\ } \quad \begin{matrix} \phi(x^{(1)}) & u^{(1)} \\ \phi(x^{(2)}) & u^{(2)} \\ \phi(x^{(3)}) & u^{(3)} \\ \vdots & \\ \phi(x^{(T)}) & u^{(T)} \end{matrix}$$

Benchmark:
Best $\phi \in \Phi$

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

Regret

regret = reward obtained vs benchmark

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

ϕ $\longrightarrow$

Benchmark:
Best $\phi \in \Phi$

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

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

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

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

Regret

regret = reward obtained vs benchmark

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

ϕ $\longrightarrow$

Benchmark:
Best $\phi \in \Phi$

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

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

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

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

ϵ-Regret: $\leq \epsilon T$

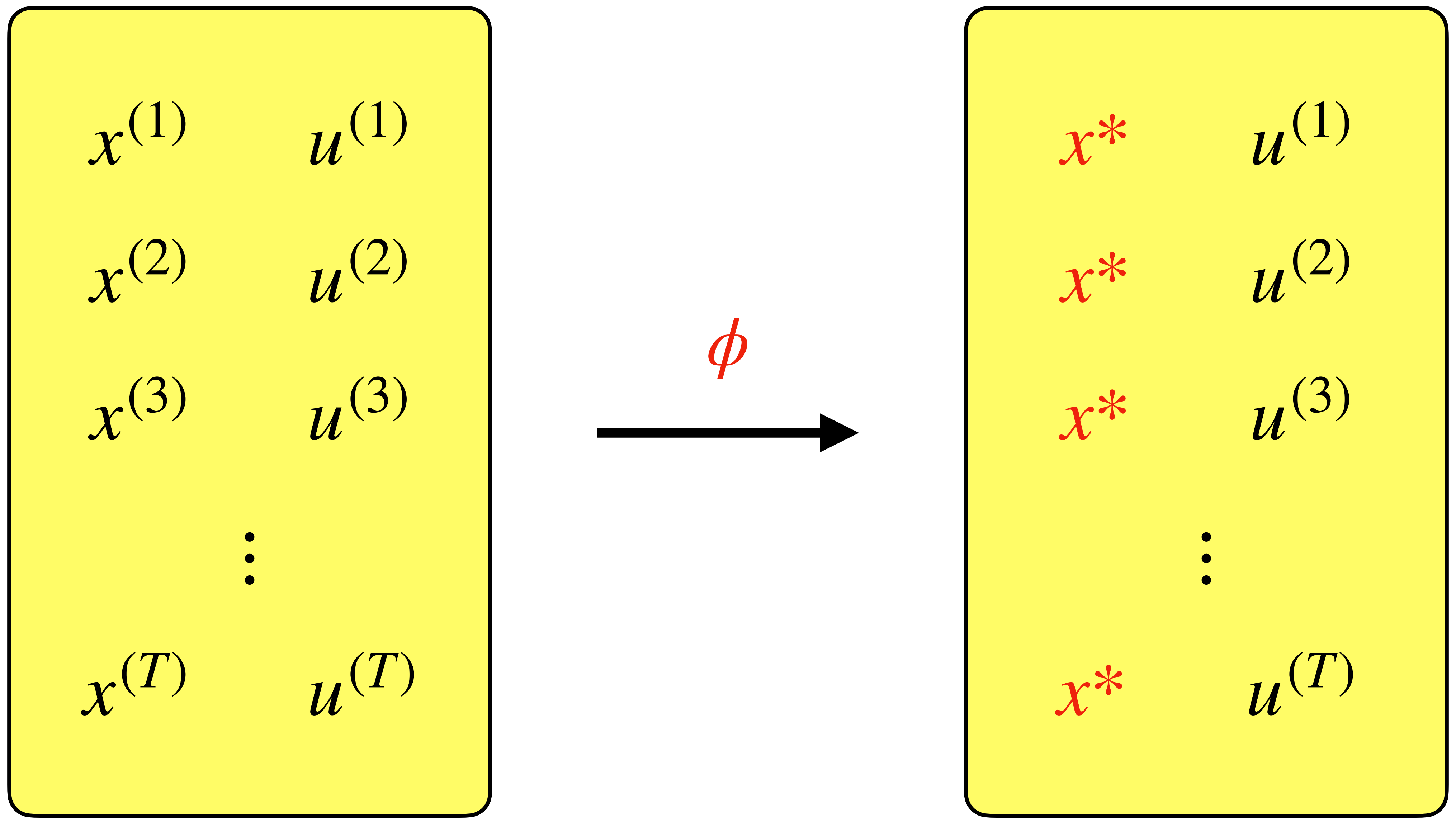
External Regret

External Regret

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

External Regret

Φ = **all constant maps**



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	$u^{(3)}$
$\vdots$	
B	$u^{(T)}$

Swap Regret

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

✓ **A** $u^{(1)}$

✗ **B** $u^{(2)}$

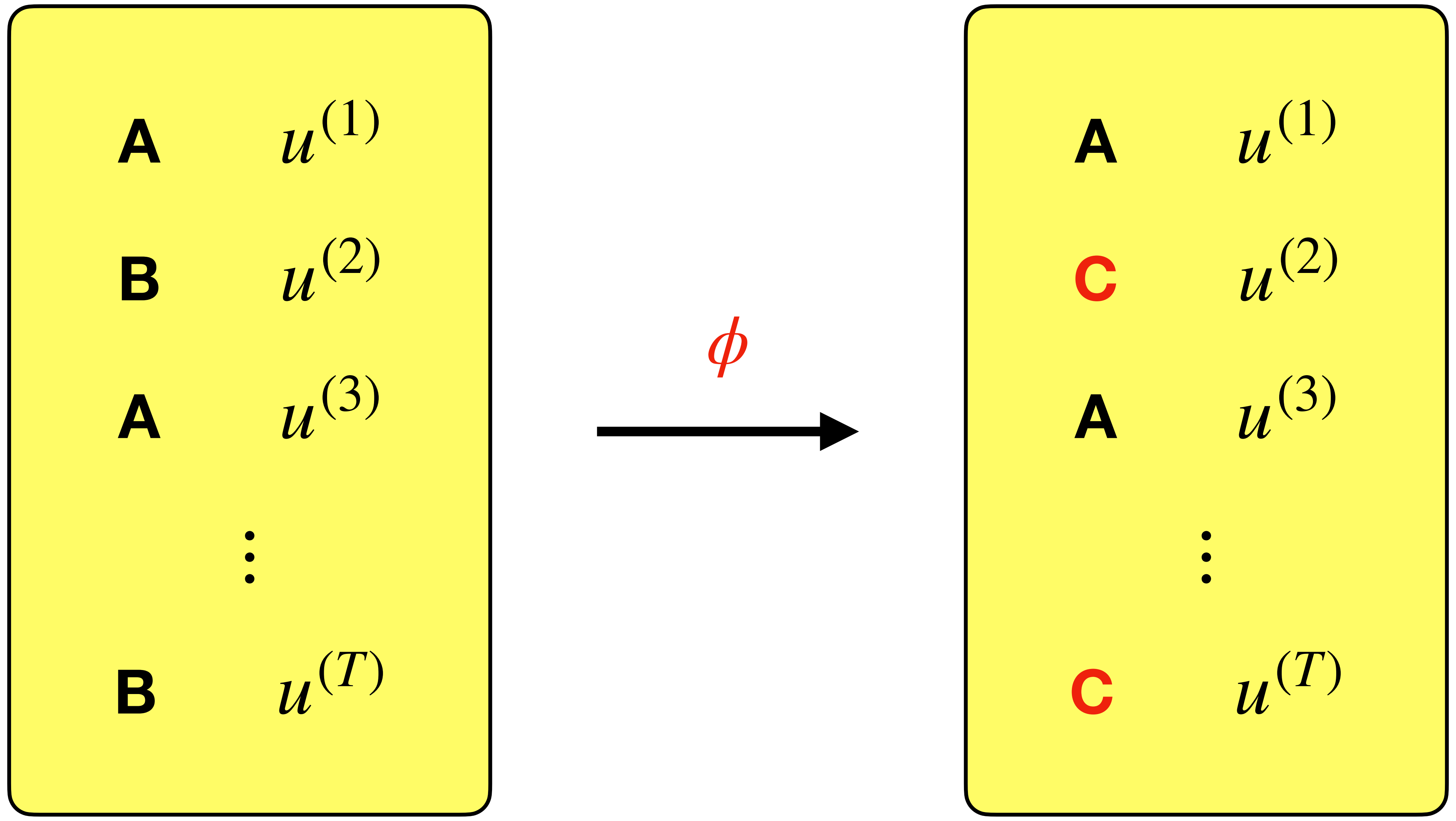
✓ **A** $u^{(3)}$

$\vdots$

✗ **B** $u^{(T)}$

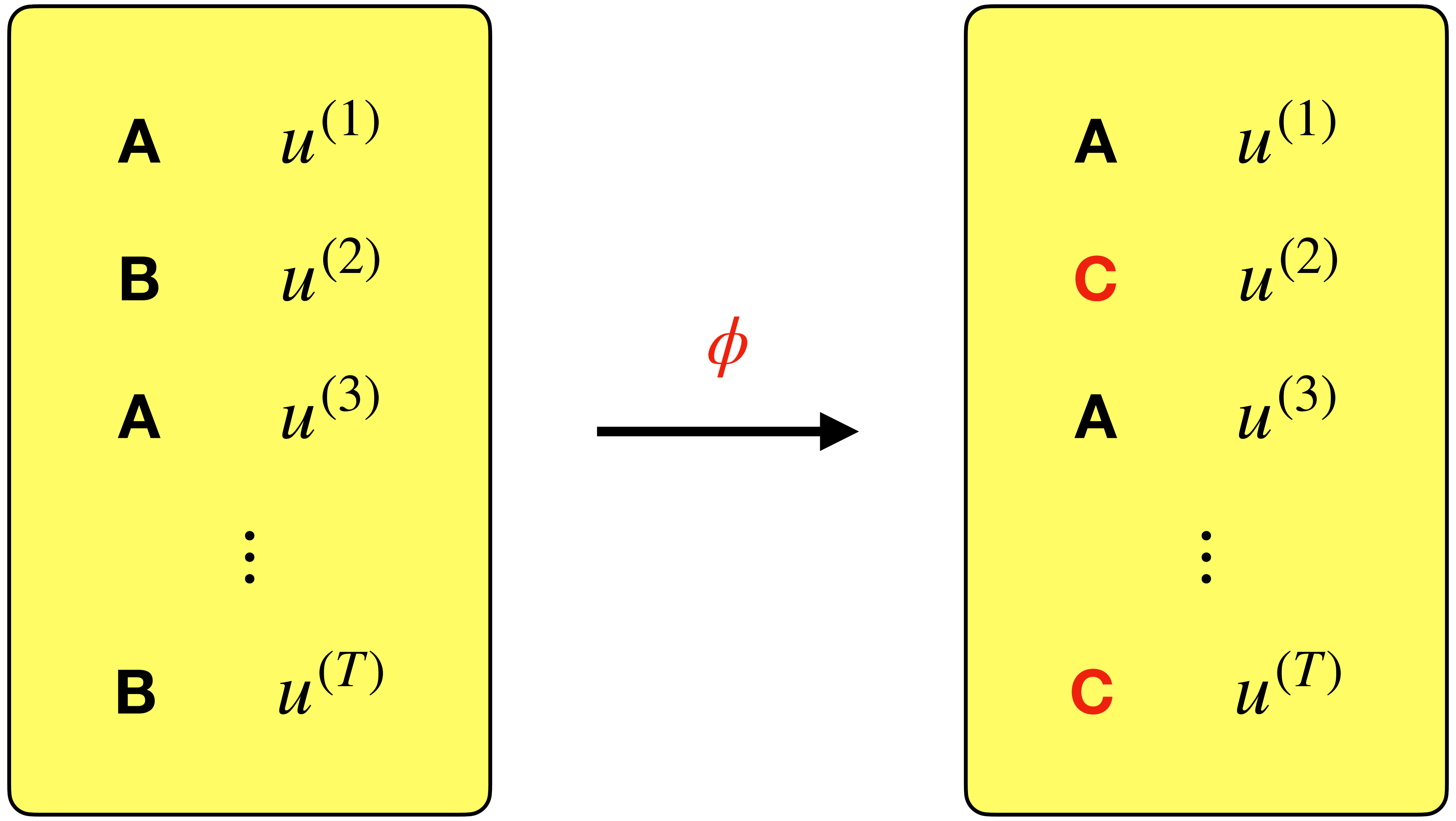
Swap Regret

Φ = **all maps** $[N] \to [N]$



Swap Regret

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

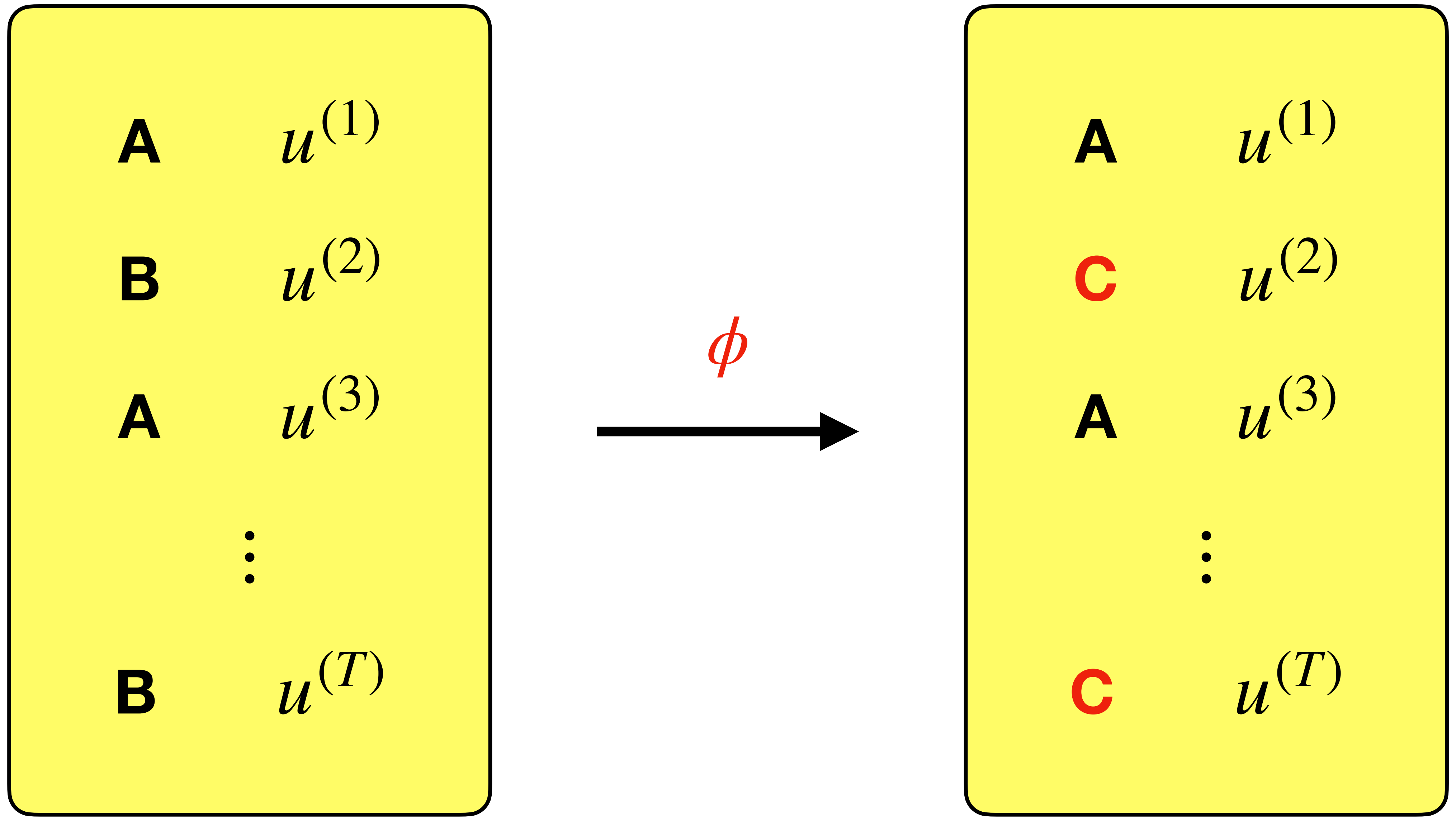


ϕ

A	A
B	C
C	C
⋮	⋮

Swap Regret

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



ϕ

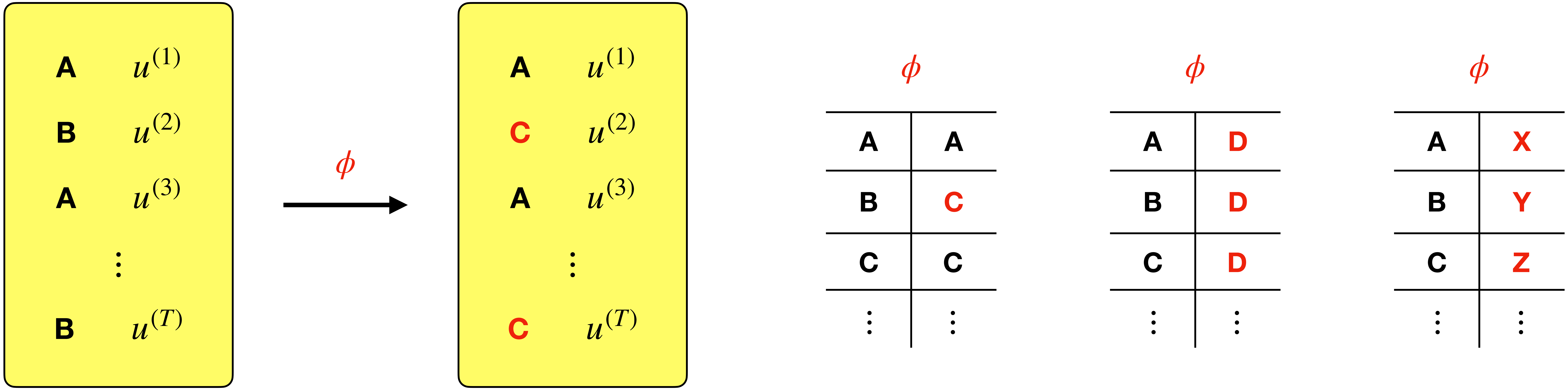
A	A
B	C
C	C
⋮	⋮

ϕ

A	D
B	D
C	D
⋮	⋮

Swap Regret

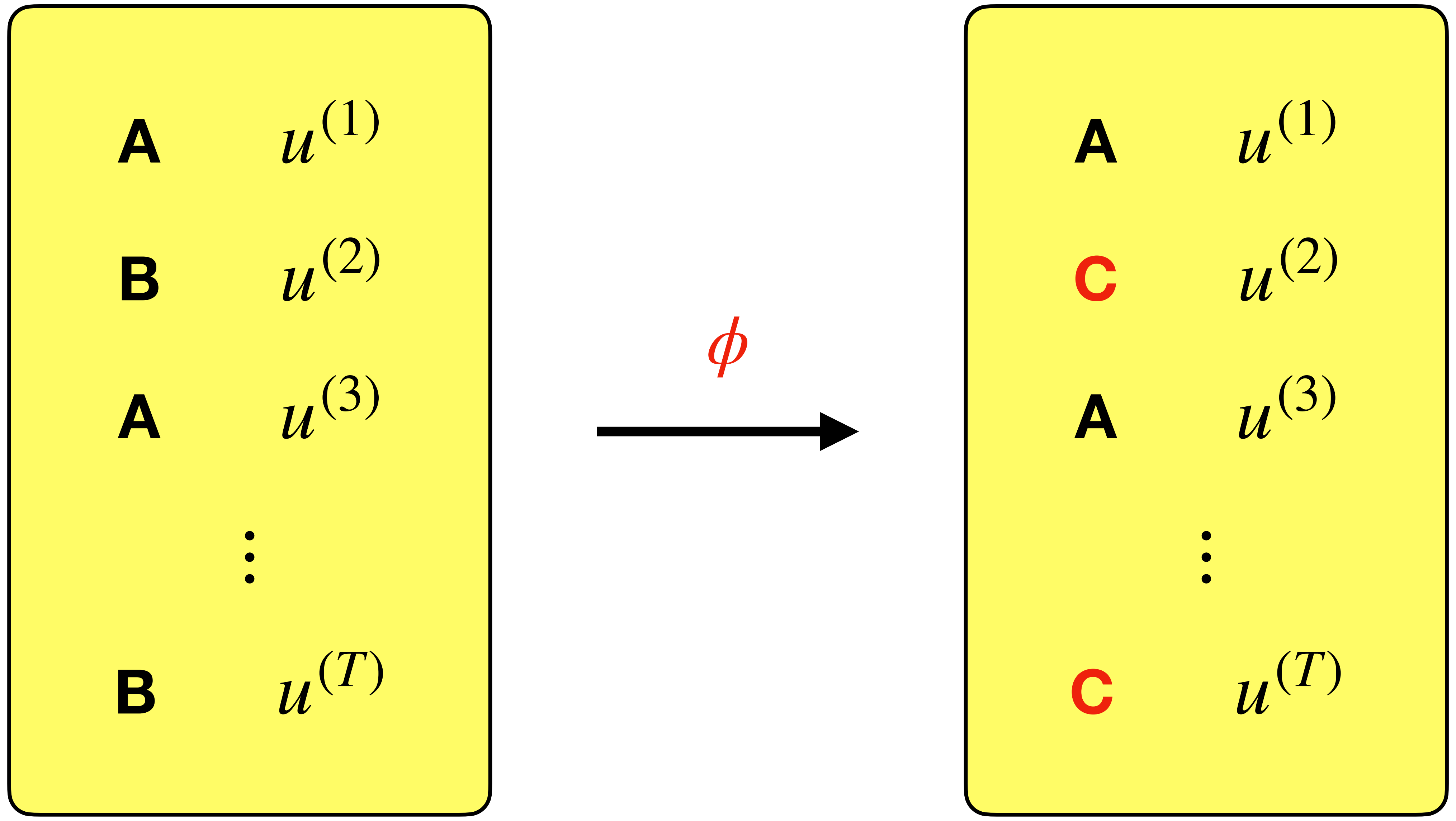
Φ = all maps $[N] \rightarrow [N]$



Swap Regret

Φ = all maps $[N] \to [N]$

Φ = all linear maps $\Delta(N) \to \Delta(N)$



ϕ

A	A
B	C
C	C
⋮	⋮

ϕ

A	D
B	D
C	D
⋮	⋮

ϕ

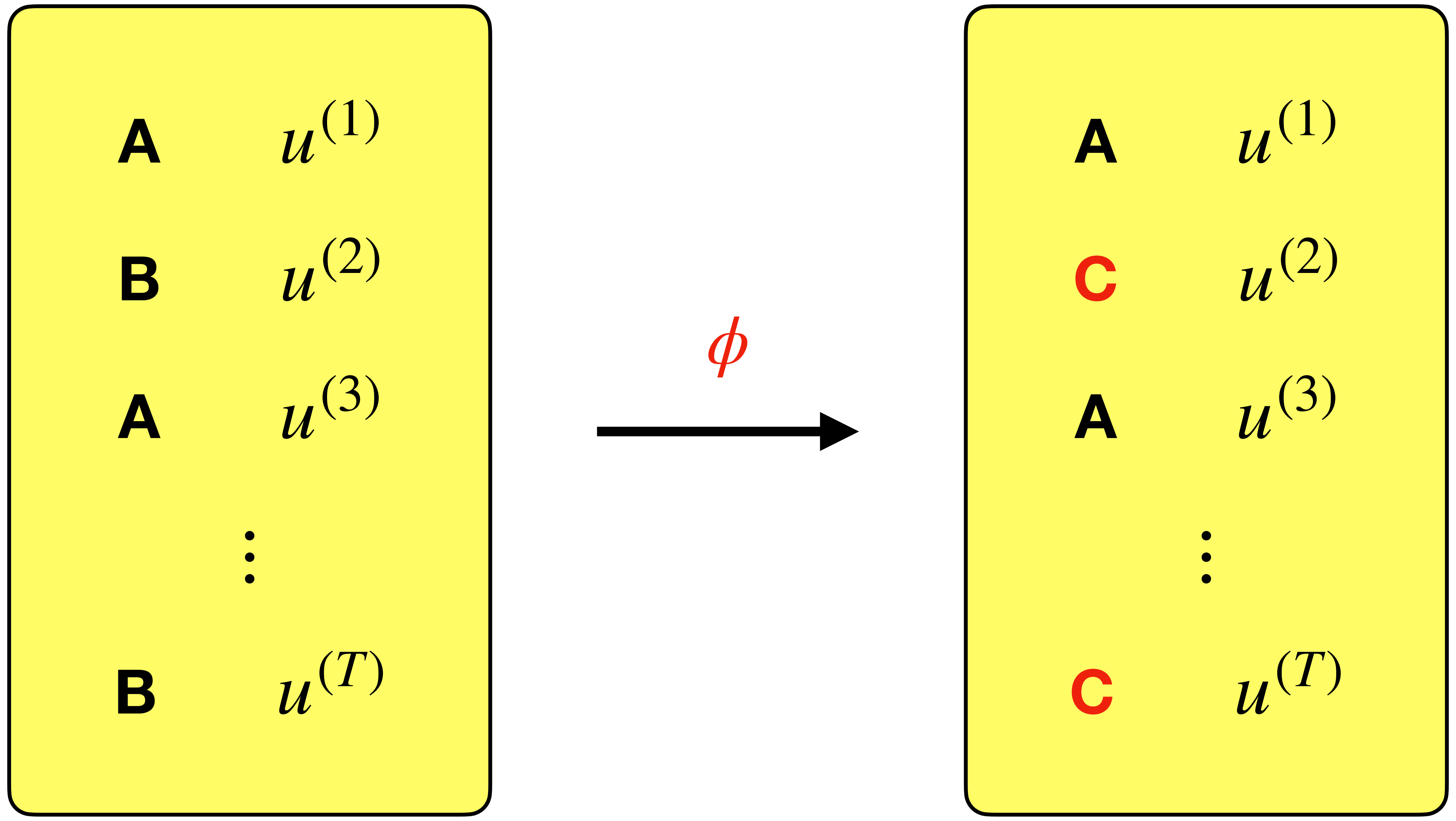
A	X
B	Y
C	Z
⋮	⋮

Swap Regret

Φ = all maps $[N] \to [N]$

Φ = all linear maps $\Delta(N) \to \Delta(N)$

This paper: Φ = ALL maps $\Delta(N) \to \Delta(N)$



ϕ

A	A
B	C
C	C
⋮	⋮

ϕ

A	D
B	D
C	D
⋮	⋮

ϕ

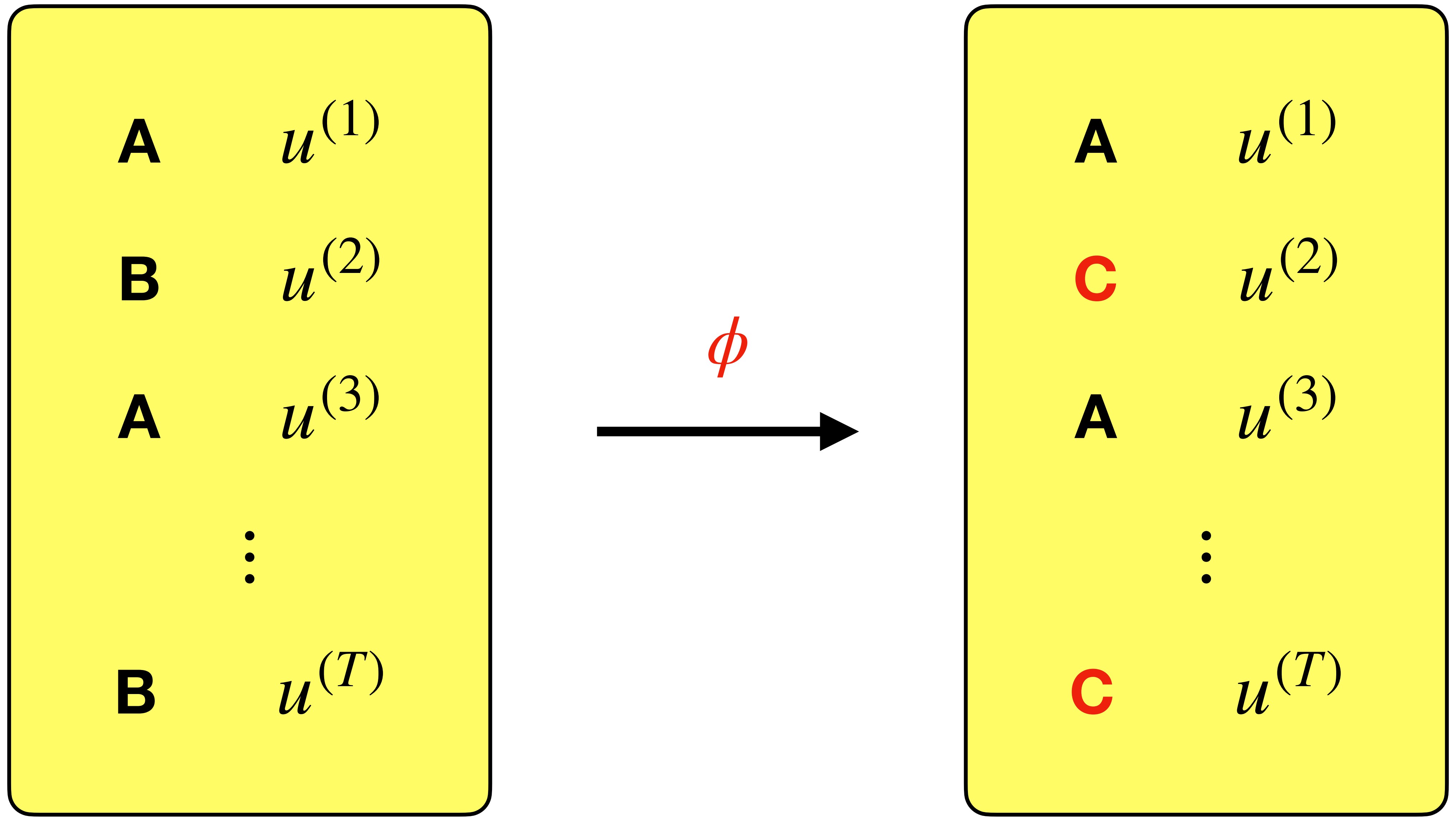
A	X
B	Y
C	Z
⋮	⋮

Swap Regret

Φ = all maps $[N] \to [N]$

Φ = all linear maps $\Delta(N) \to \Delta(N)$

This paper: Φ = ALL maps $\Delta(N) \to \Delta(N)$

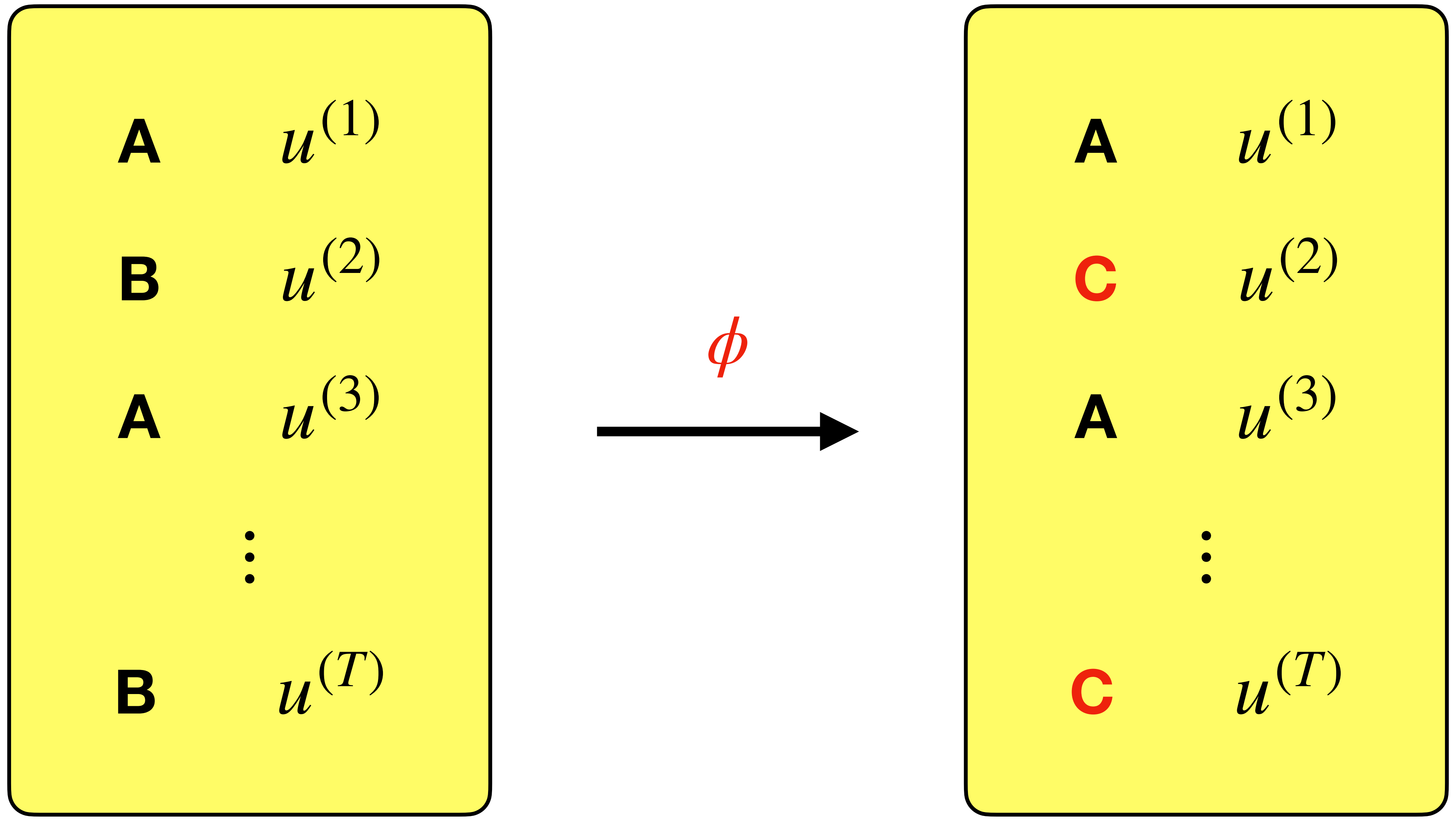


Swap Regret

Φ = all maps $[N] \to [N]$

Φ = all linear maps $\Delta(N) \to \Delta(N)$

This paper: Φ = **ALL** maps $\Delta(N) \to \Delta(N)$



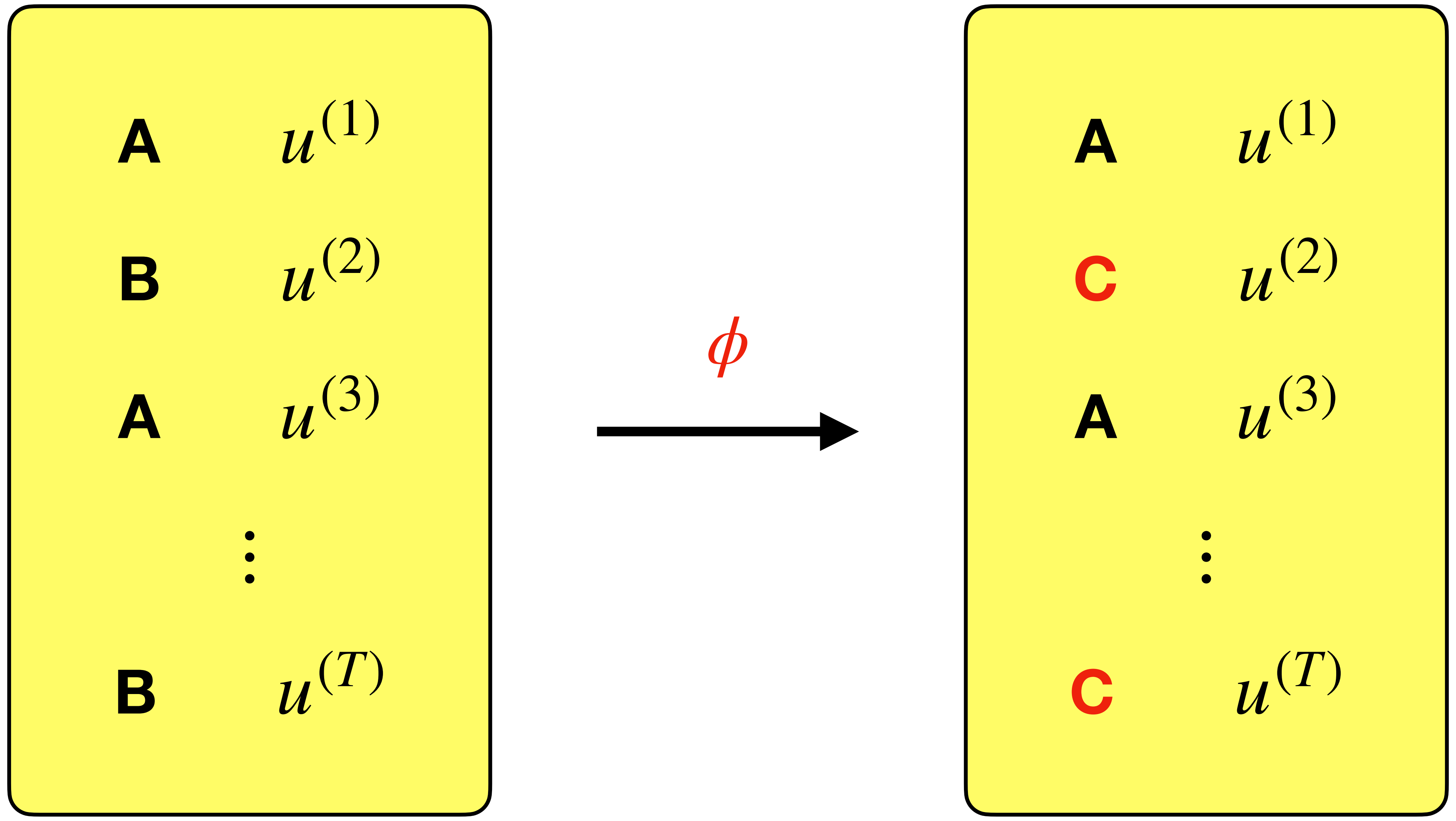
Why?

Swap Regret

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

$\Phi =$ **all linear maps** $\Delta(N) \to \Delta(N)$

This paper: $\Phi =$ **ALL maps** $\Delta(N) \to \Delta(N)$



Why?

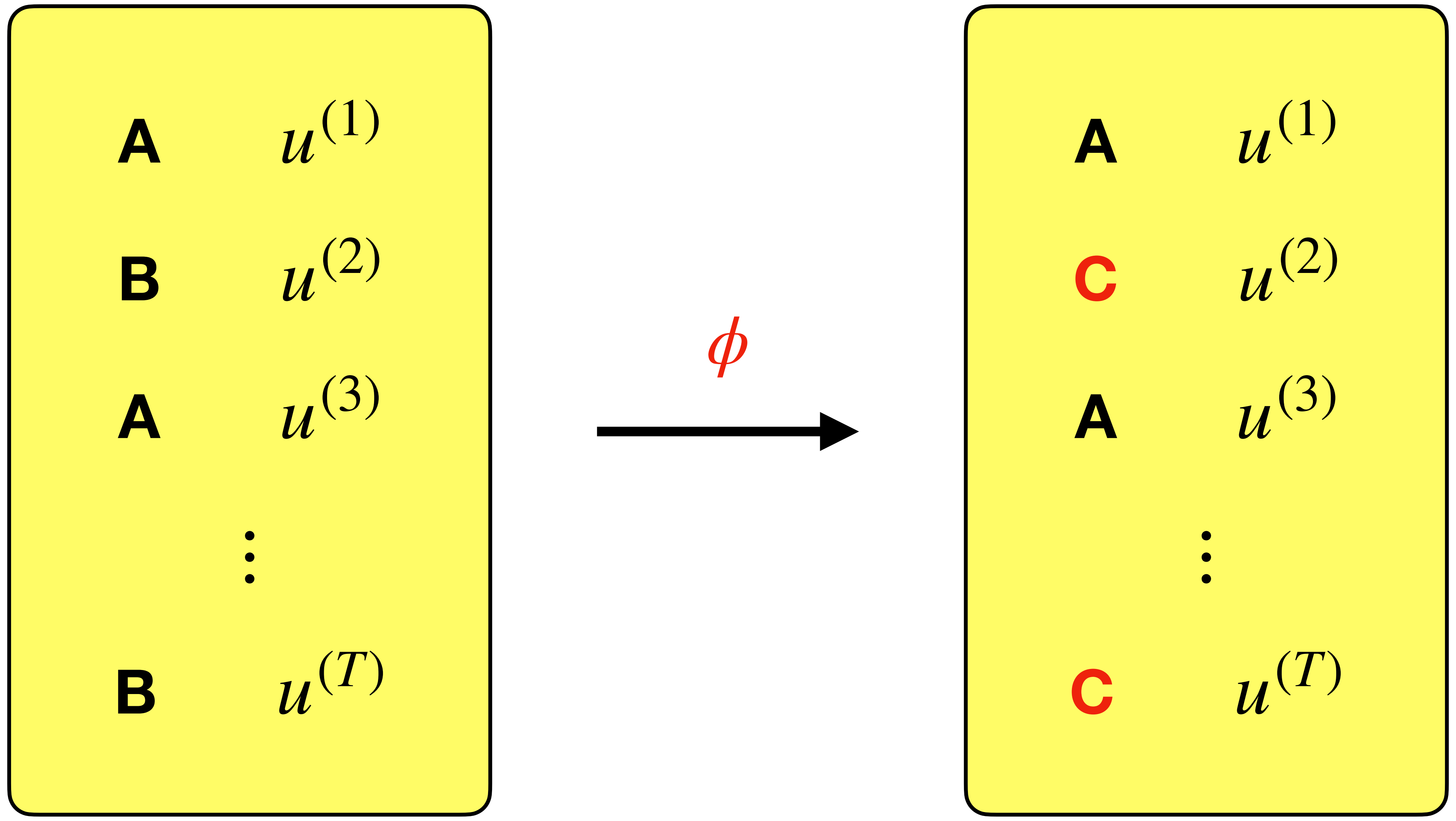
- L_2 Calibration

Swap Regret

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

$\Phi =$ **all linear maps** $\Delta(N) \to \Delta(N)$

This paper: $\Phi =$ **ALL maps** $\Delta(N) \to \Delta(N)$



Why?

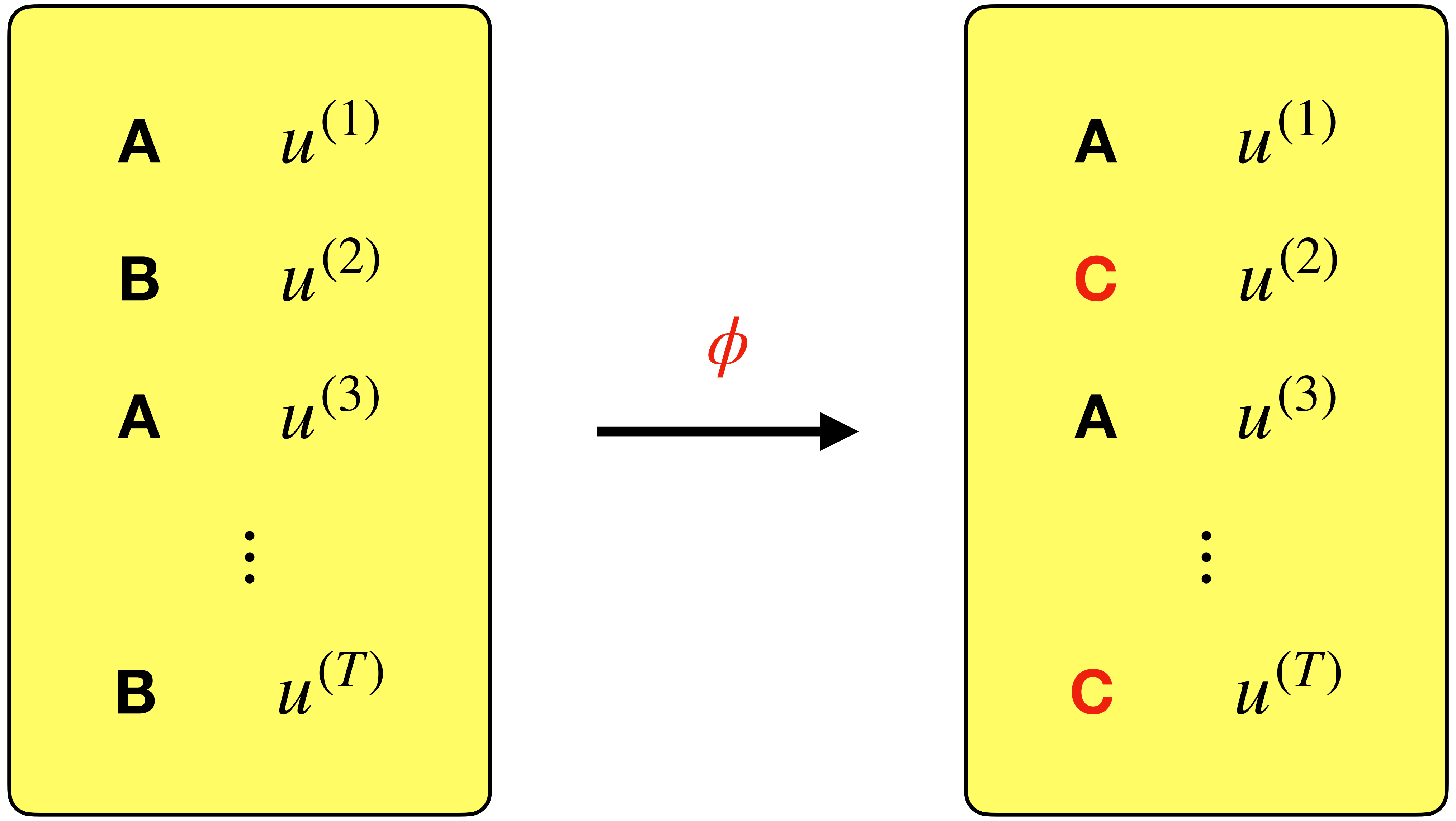
- L_2 Calibration
- Unexploitable [DSS19]

Swap Regret

Φ = all maps $[N] \to [N]$

Φ = all linear maps $\Delta(N) \to \Delta(N)$

This paper: Φ = **ALL** maps $\Delta(N) \to \Delta(N)$



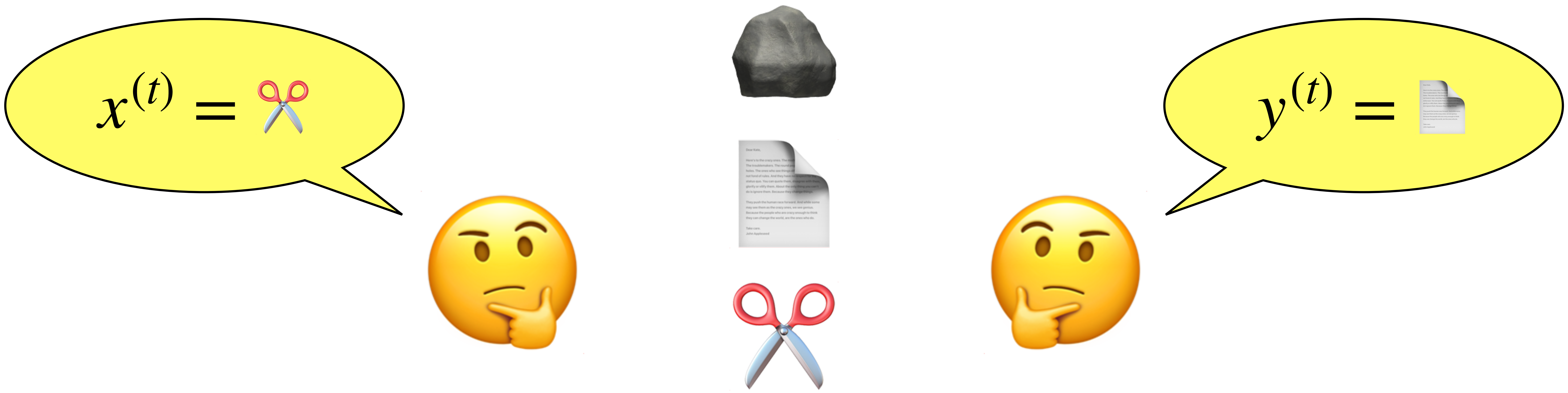
Why?

- L_2 Calibration
- Unexploitable [DSS19]
- Games!

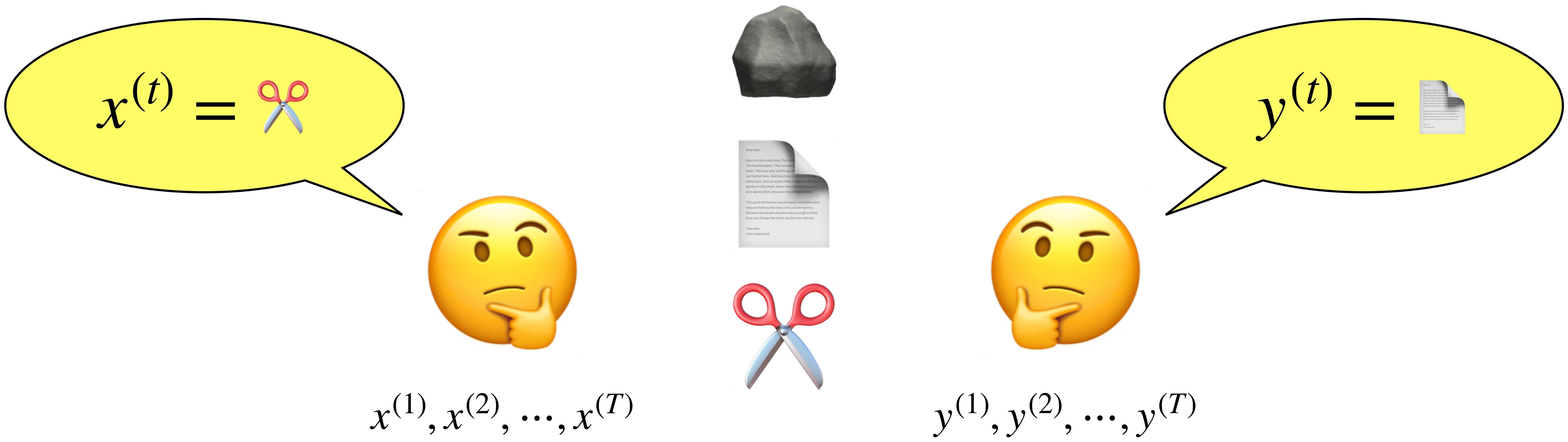
Learning in Games

Learning in Games

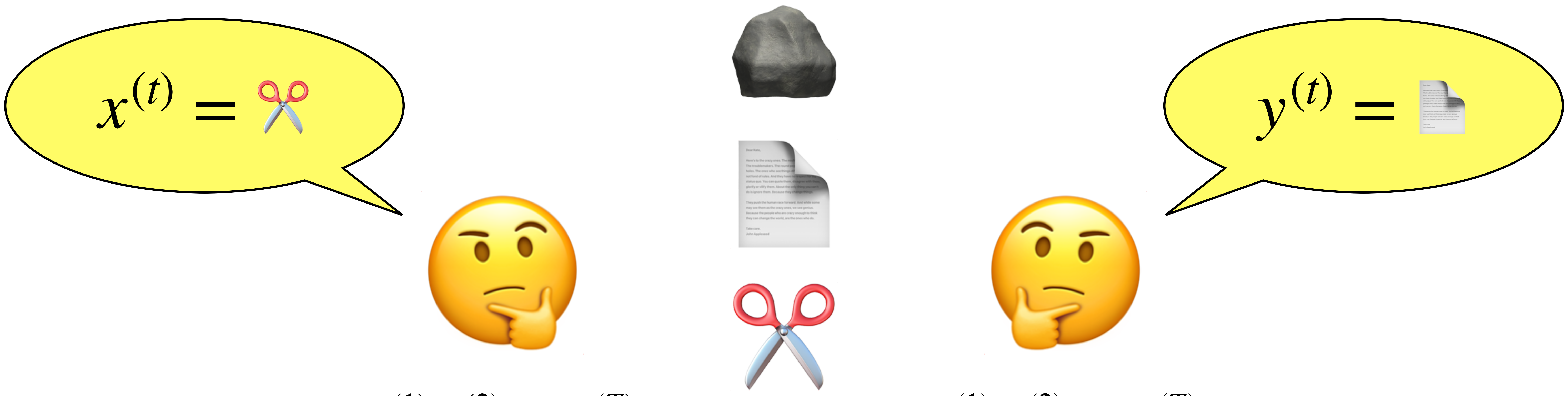
Learning in Games



Learning in Games



Learning in Games

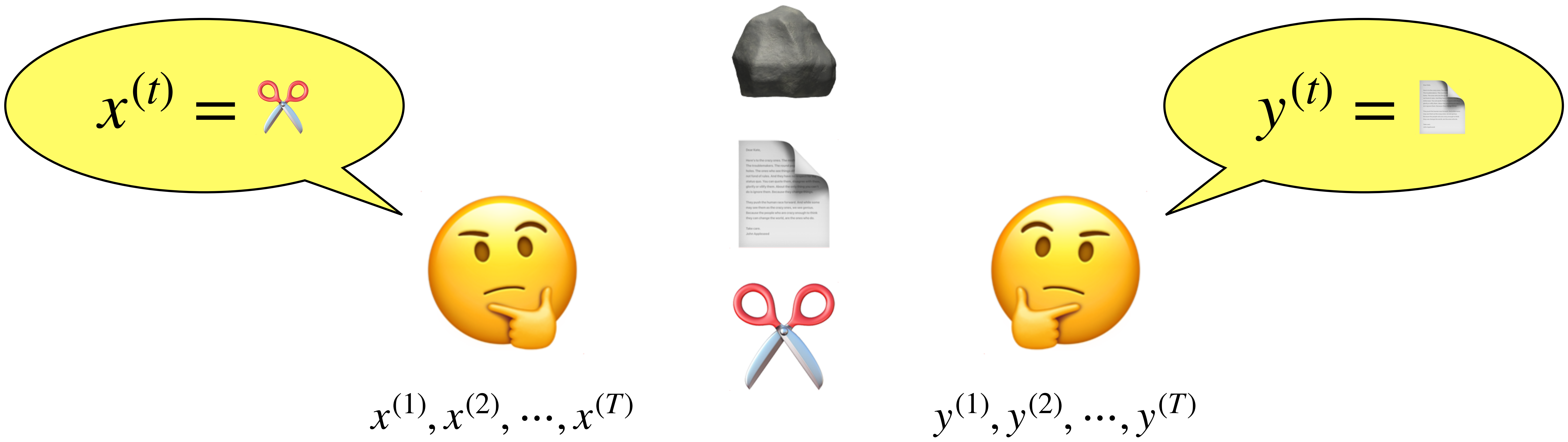


$x^{(1)}, x^{(2)}, \cdots, x^{(T)}$

$y^{(1)}, y^{(2)}, \cdots, y^{(T)}$

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

Learning in Games



ϵ external regret over T rounds → average strategies are ϵ coarse-correlated-equilibrium

ϵ swap regret over T rounds → average strategies are ϵ correlated-equilibrium

Learning in Games

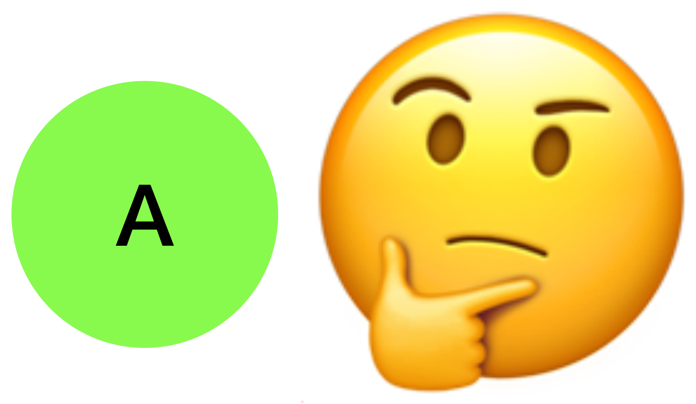
Game Equilibrium: Strategy profile where no player wants to deviate

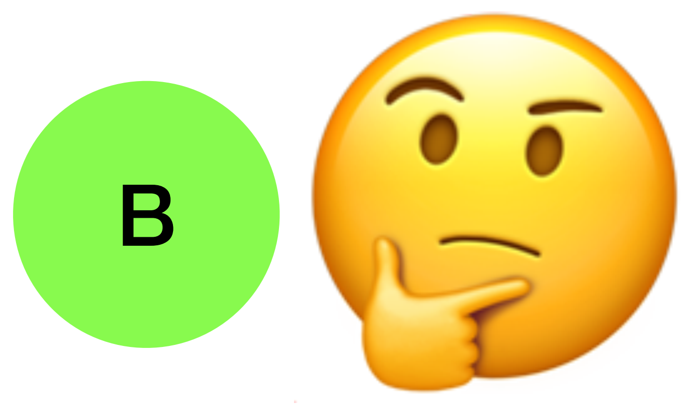
Learning in Games

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

Coarse Deviations

A	C
B	C

Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate

Coarse Deviations

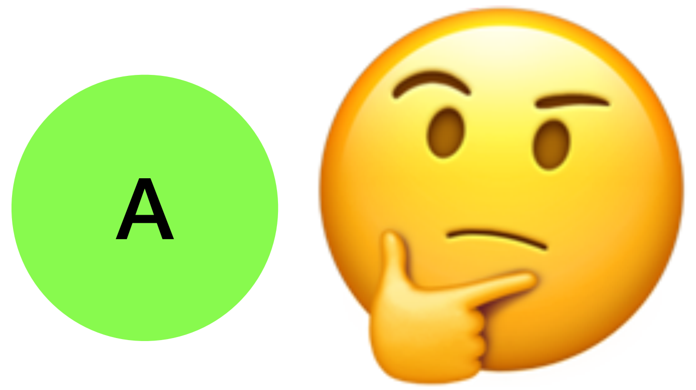
A	C
B	C

Swap Deviations

A	A
B	C

Learning in Games

Game Equilibrium: Strategy profile where no player wants to deviate



Coarse Deviations

A	C
B	C

Swap Deviations

A	A
B	C

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	C
B	C

Swap Deviations

A	A
B	C

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	C
B	C

Swap Deviations

A	A
B	C

CCE: No player wants to make a coarse deviation

CE: No player wants to make a swap 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]
[PR23]

Main Results

Main Results

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

Main Results

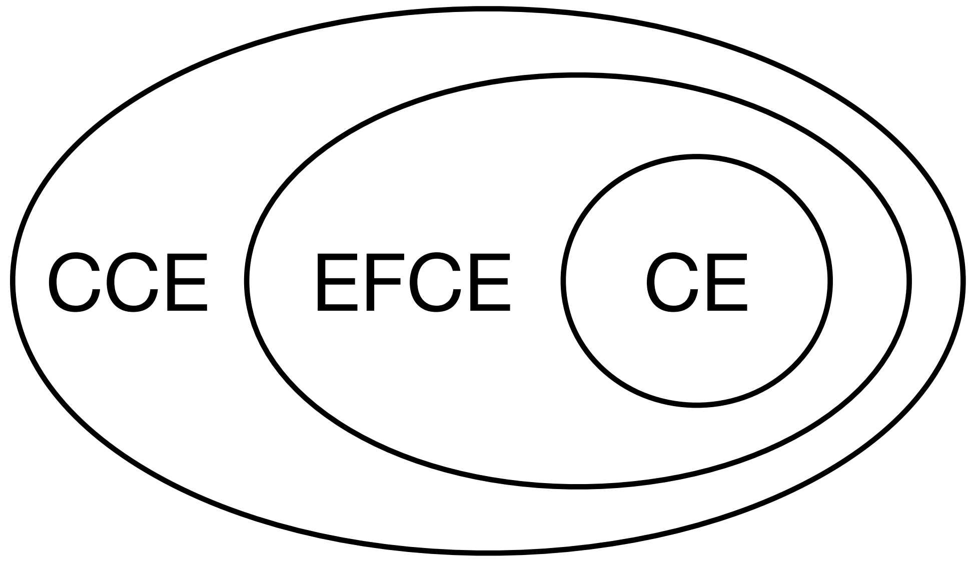
Tree-Swap: ϵ-Swap-Regret for $T = \log(N)^{\tilde{\Omega}(1/\epsilon)}$

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

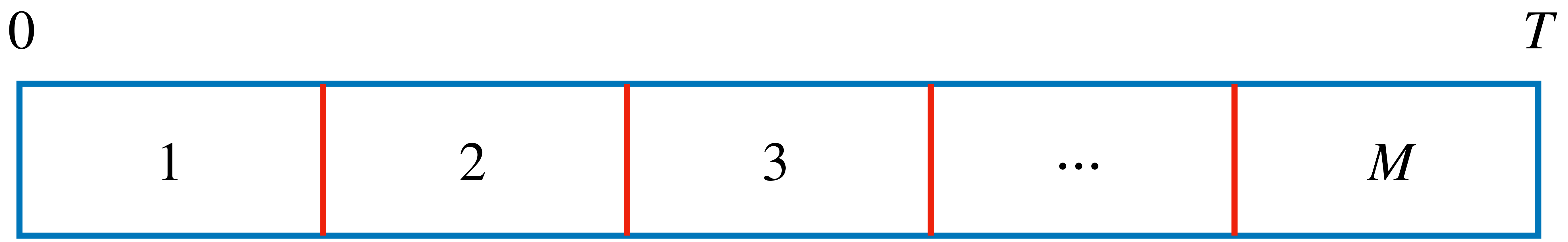


Algorithms

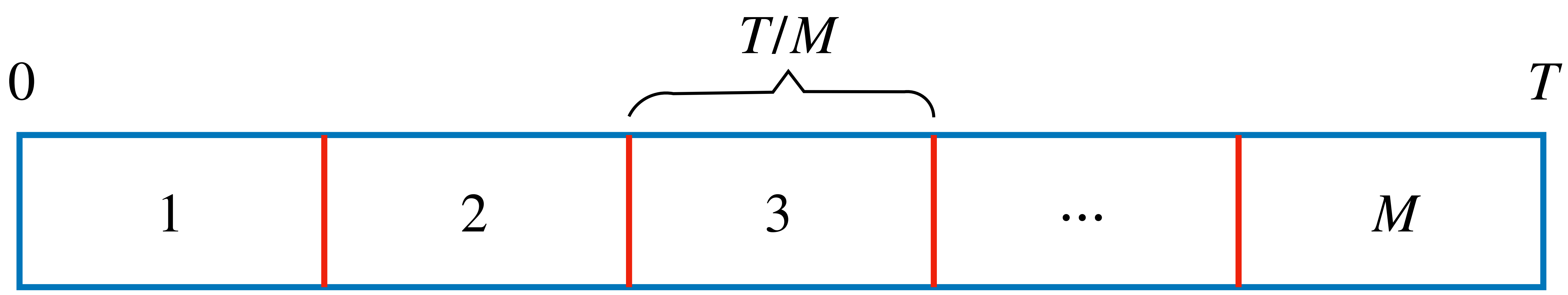
Lazy External Regret

Lazy External Regret

Lazy External Regret



Lazy External Regret



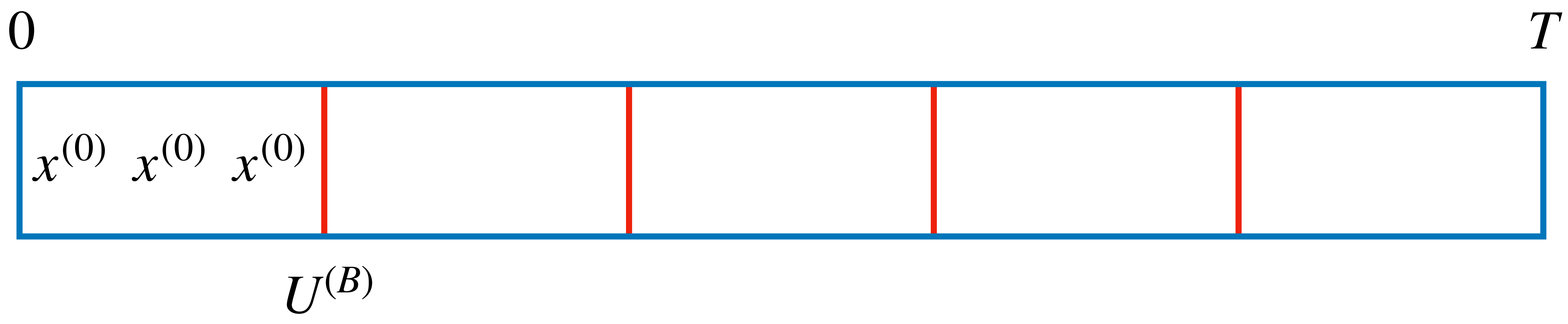
Lazy External Regret

Lazy External Regret

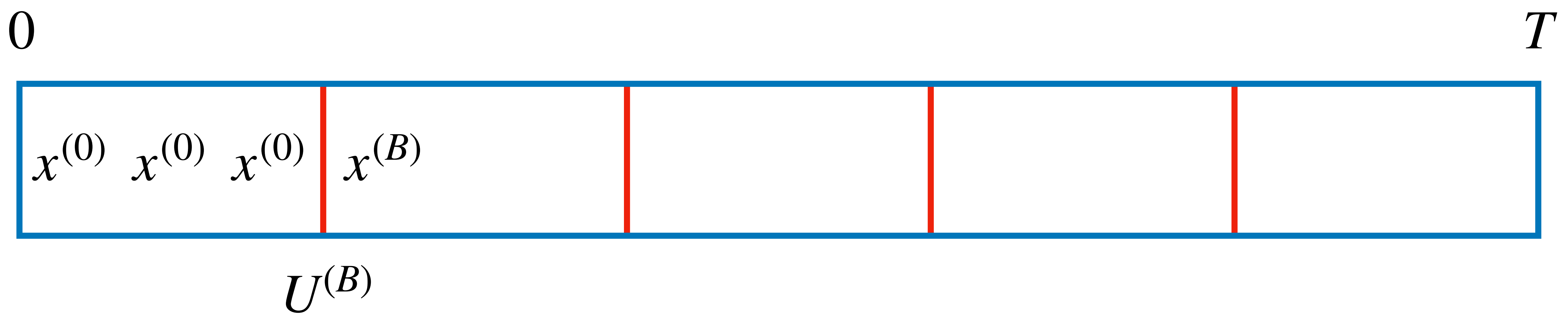
Lazy External Regret

Lazy External Regret

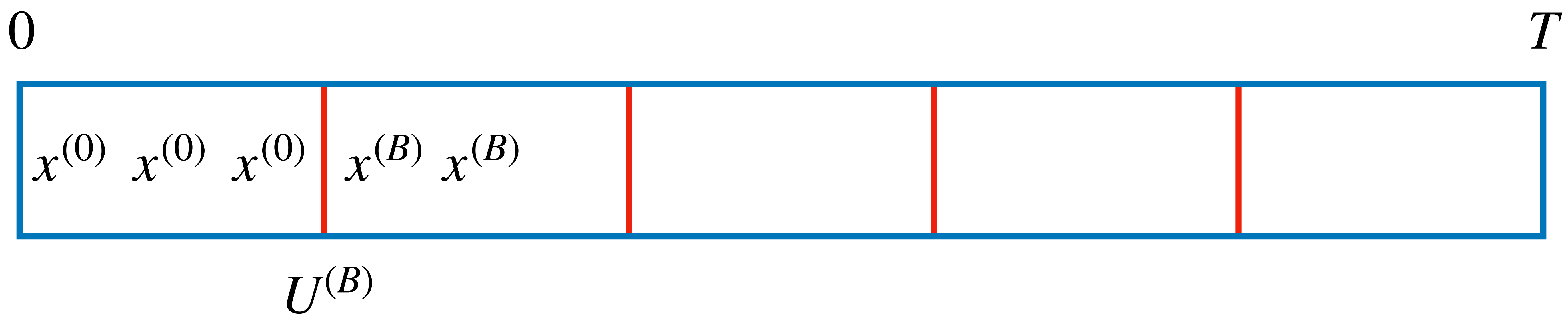
Lazy External Regret



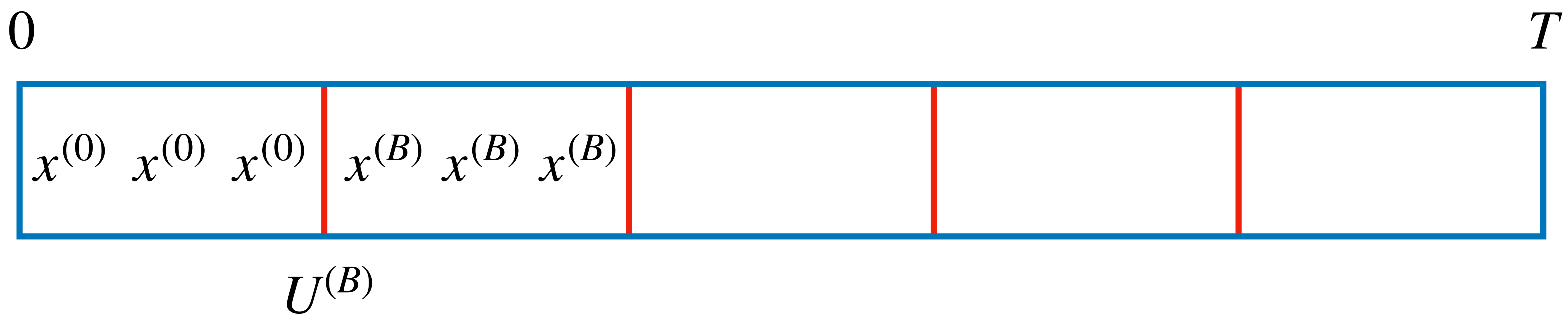
Lazy External Regret



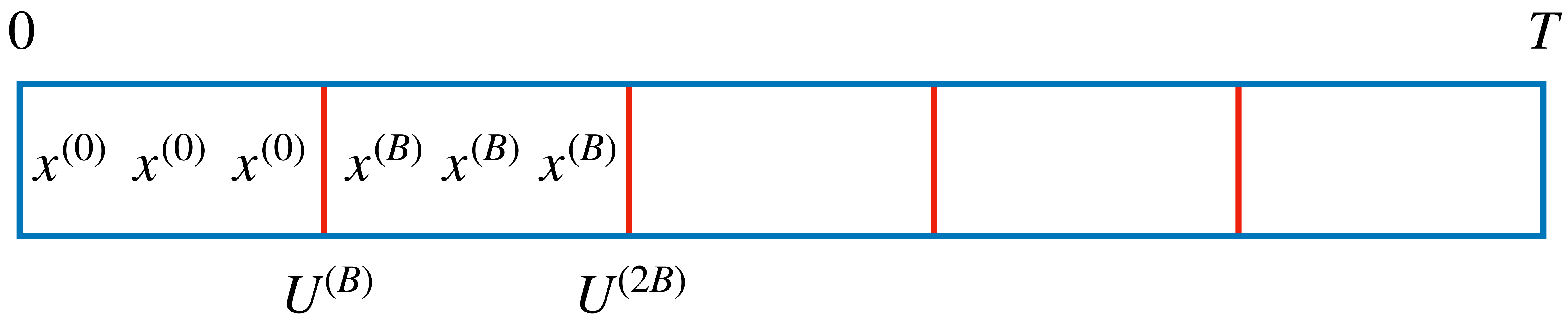
Lazy External Regret



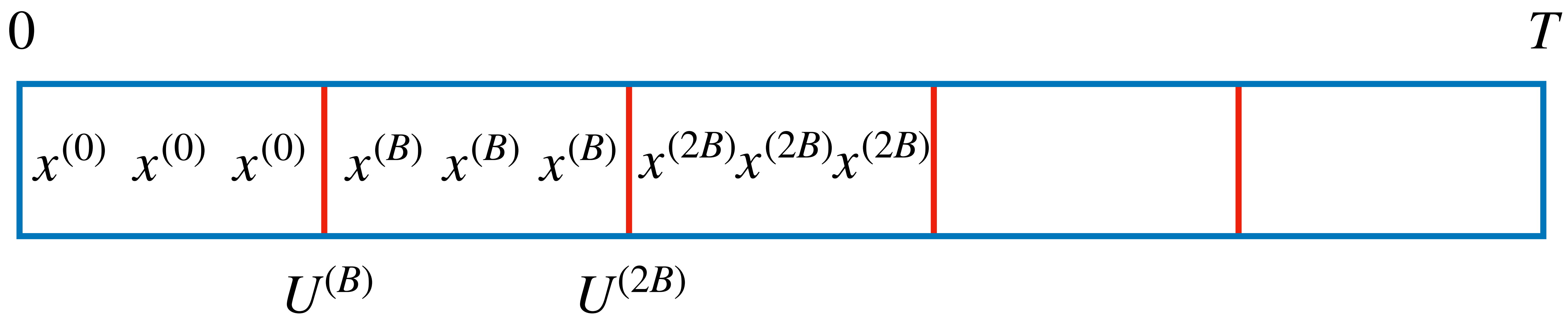
Lazy External Regret



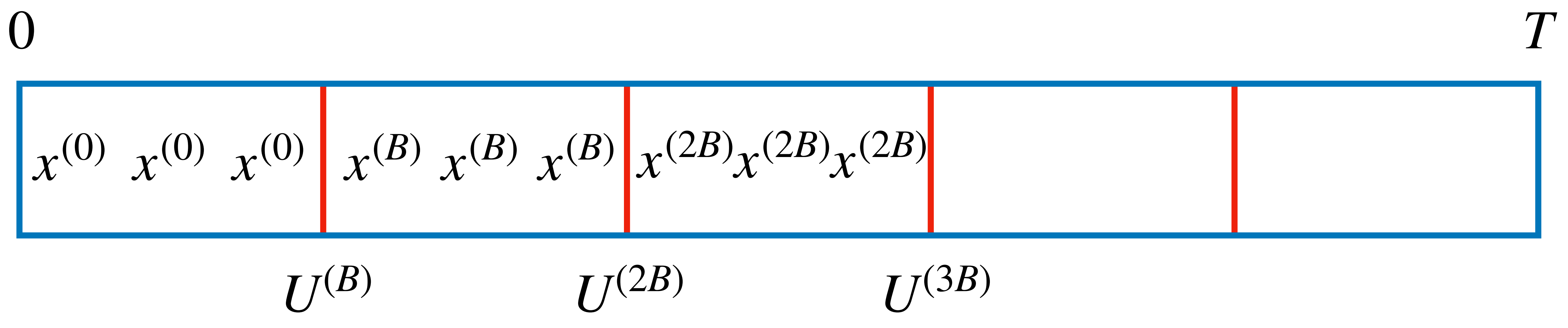
Lazy External Regret



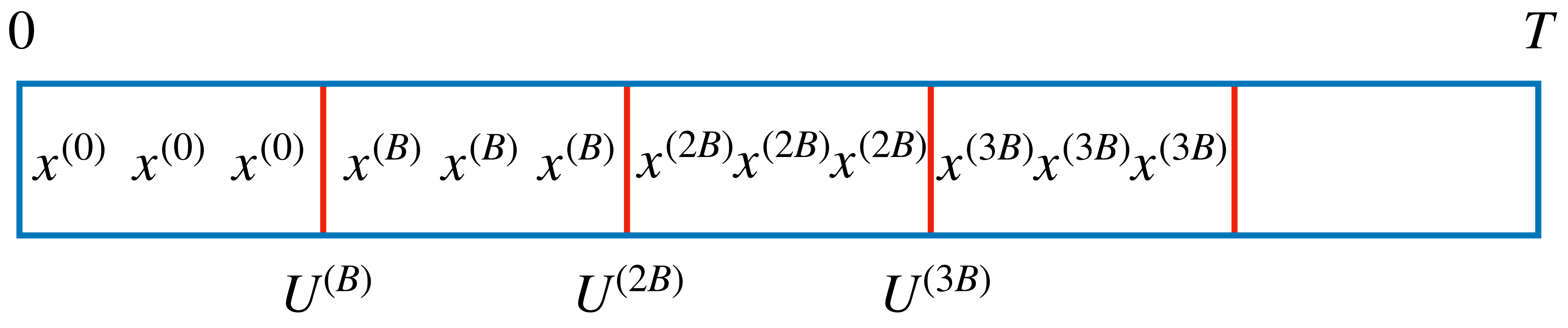
Lazy External Regret



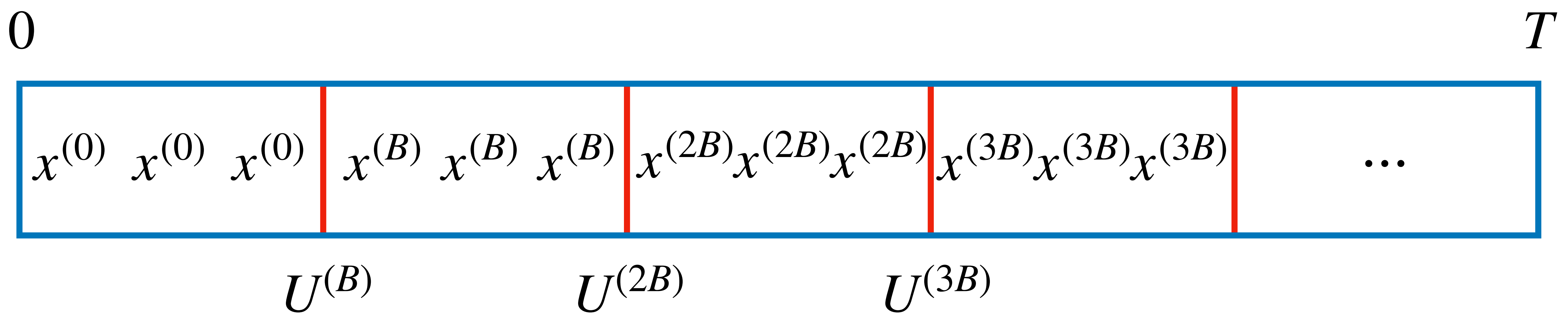
Lazy External Regret



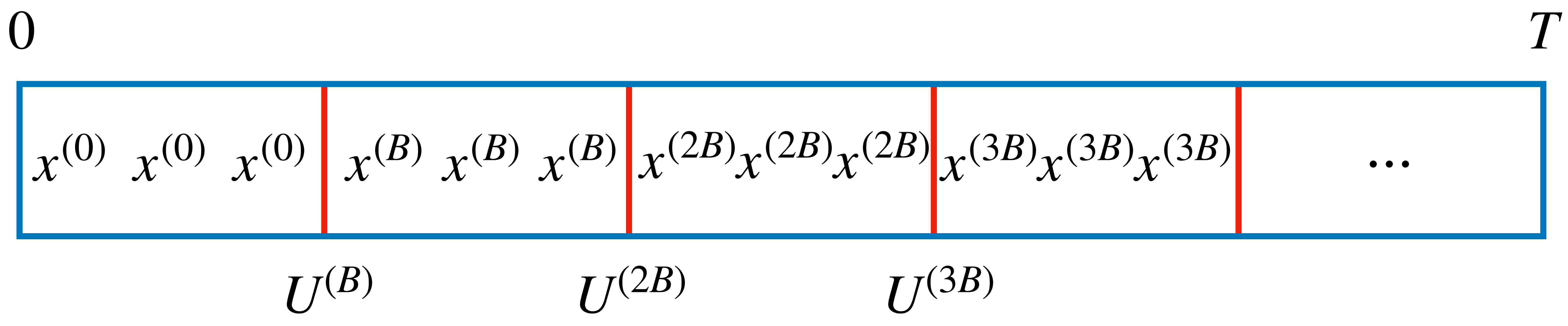
Lazy External Regret



Lazy External Regret

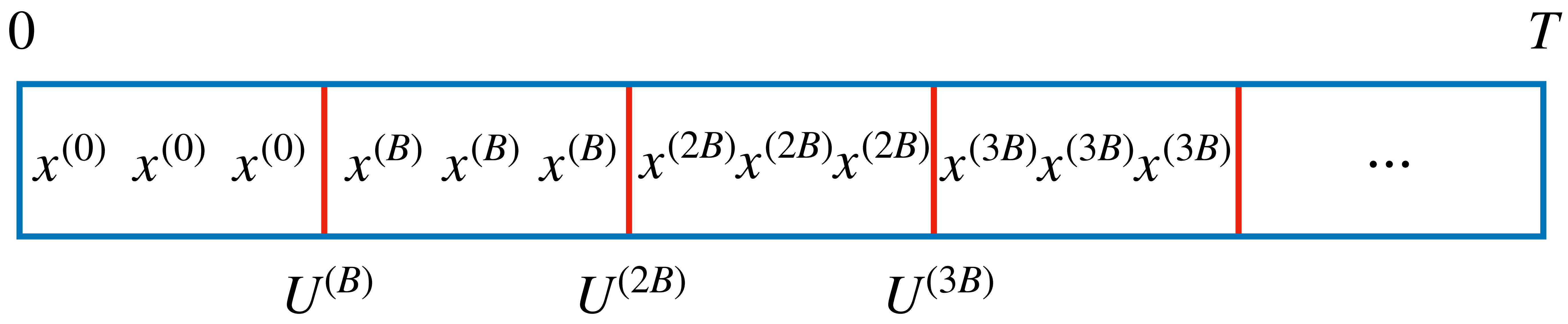


Lazy External Regret



MWU: external-regret $\epsilon = O\left(\sqrt{\frac{\log N}{T}}\right)$

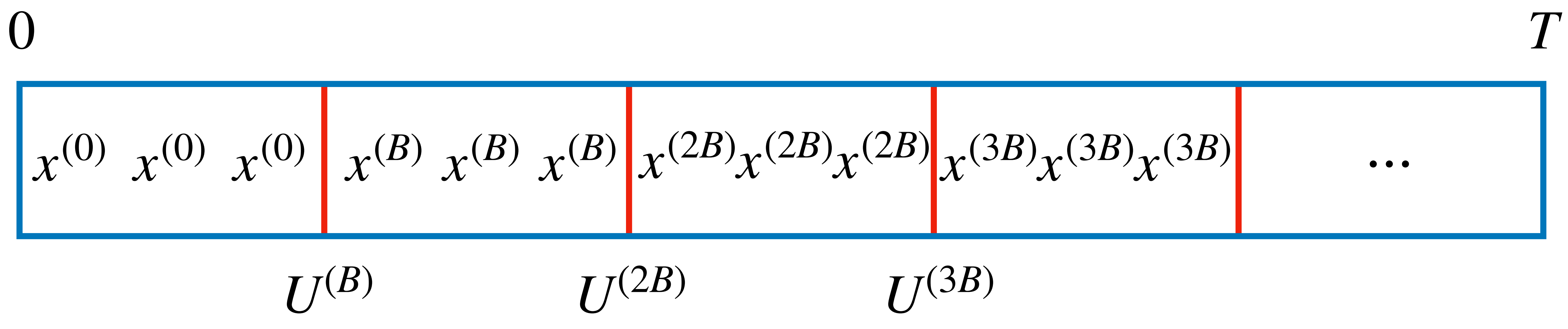
Lazy External Regret



MWU: external-regret $\epsilon = O\left(\sqrt{\frac{\log N}{T}}\right)$

Lazy MWU: external-regret $\epsilon = O\left(\sqrt{\frac{\log N}{M}}\right)$

Lazy External Regret

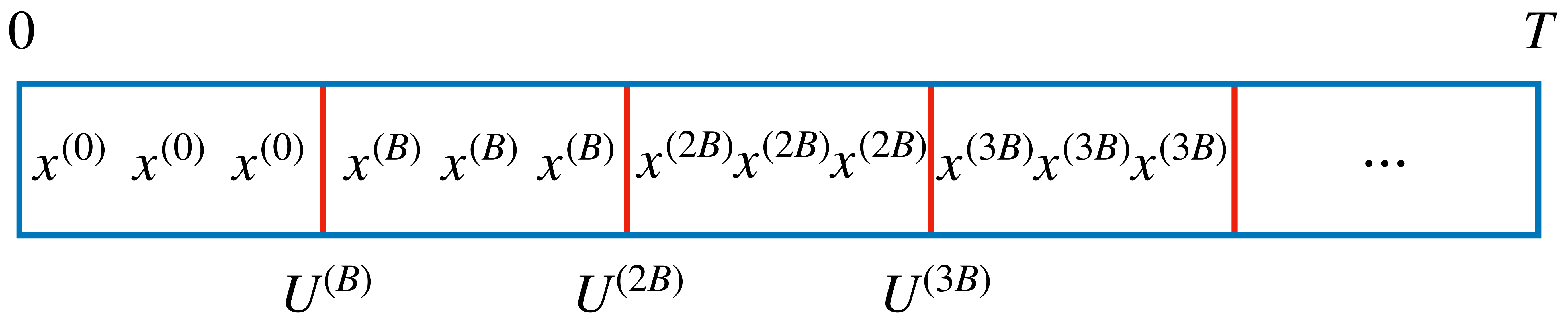


MWU: external-regret $\epsilon = O\left(\sqrt{\frac{\log N}{T}}\right)$

Lazy MWU: external-regret $\epsilon = O\left(\sqrt{\frac{\log N}{M}}\right)$

Why?

Lazy External Regret



MWU: external-regret $\epsilon = O\left(\sqrt{\frac{\log N}{T}}\right)$

Lazy MWU: external-regret $\epsilon = O\left(\sqrt{\frac{\log N}{M}}\right)$

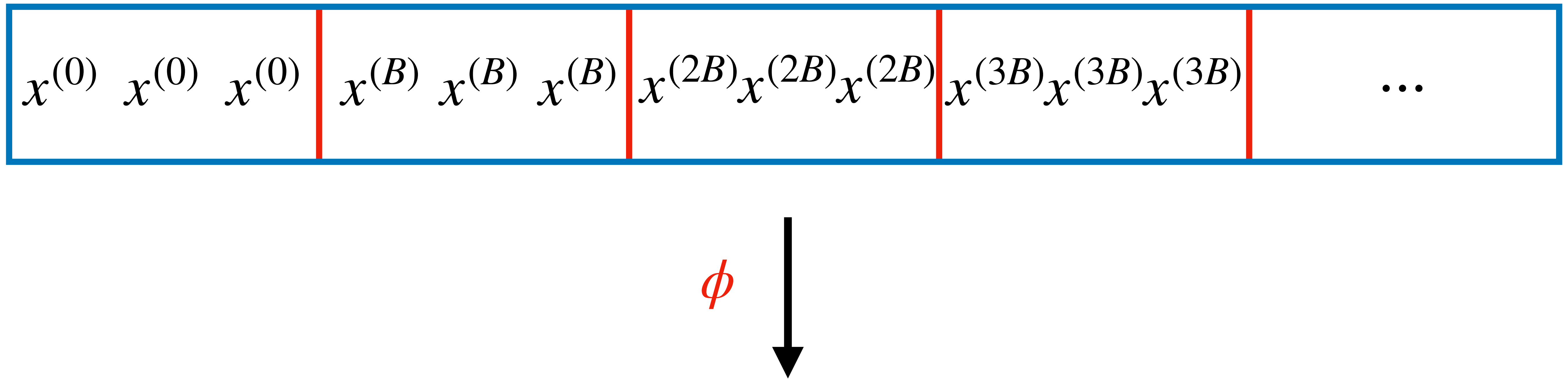
Why?

Fewer Actions

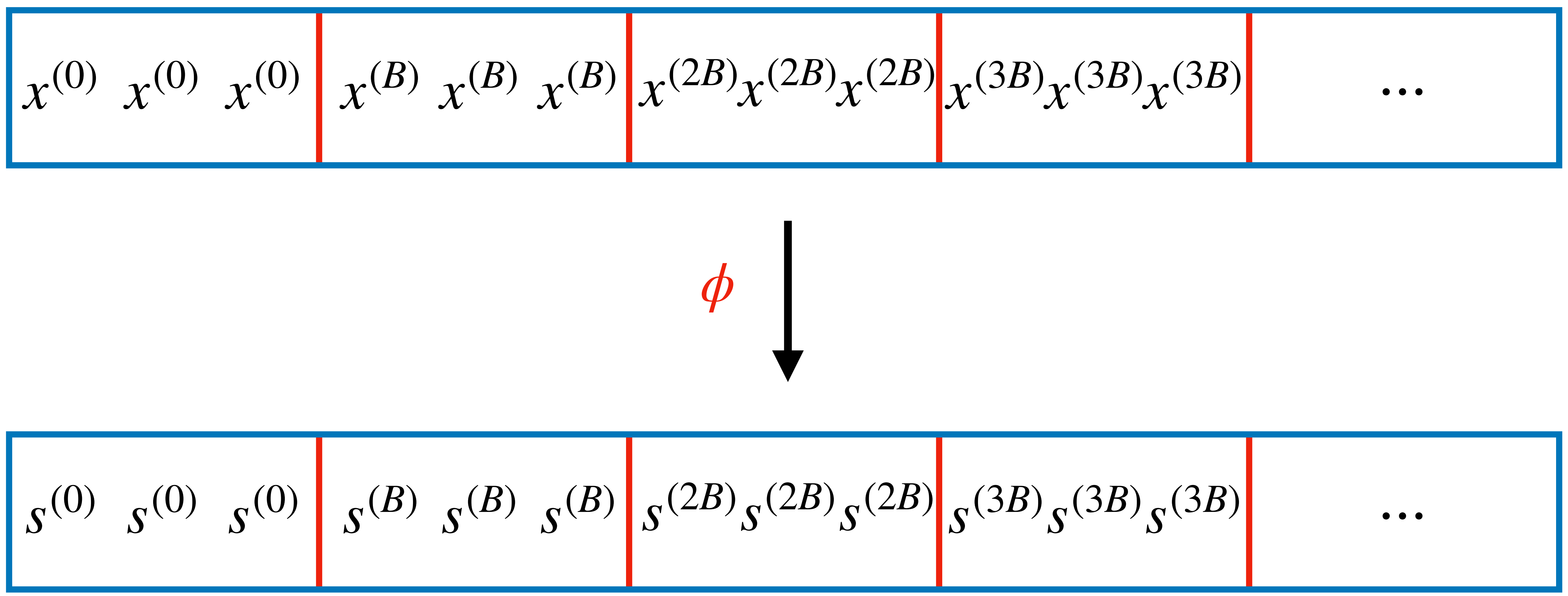
Swap Regret

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

Swap Regret



Swap Regret



$x^{(0)}$ $x^{(0)}$ $x^{(0)}$
$x^{(B)}$ $x^{(B)}$ $x^{(B)}$
$x^{(2B)}$ $x^{(2B)}$ $x^{(2B)}$
$x^{(3B)}$ $x^{(3B)}$ $x^{(3B)}$
…
ϕ
$s^{(0)}$ $s^{(0)}$ $s^{(0)}$
$s^{(B)}$ $s^{(B)}$ $s^{(B)}$
$s^{(2B)}$ $s^{(2B)}$ $s^{(2B)}$
$s^{(3B)}$ $s^{(3B)}$ $s^{(3B)}$
…

Tree-Swap

Tree-Swap

$$T = M^d$$

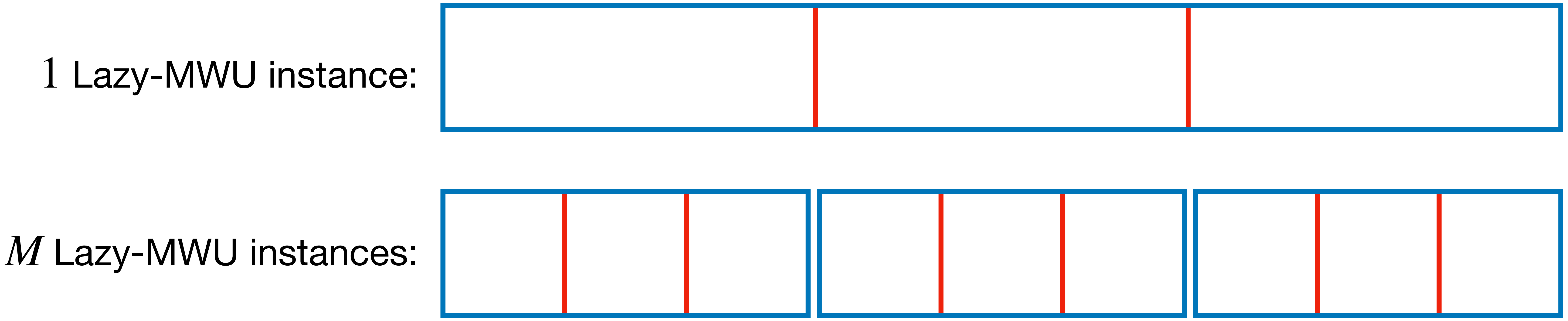
Tree-Swap

$$T = M^d$$

1 Lazy-MWU instance:

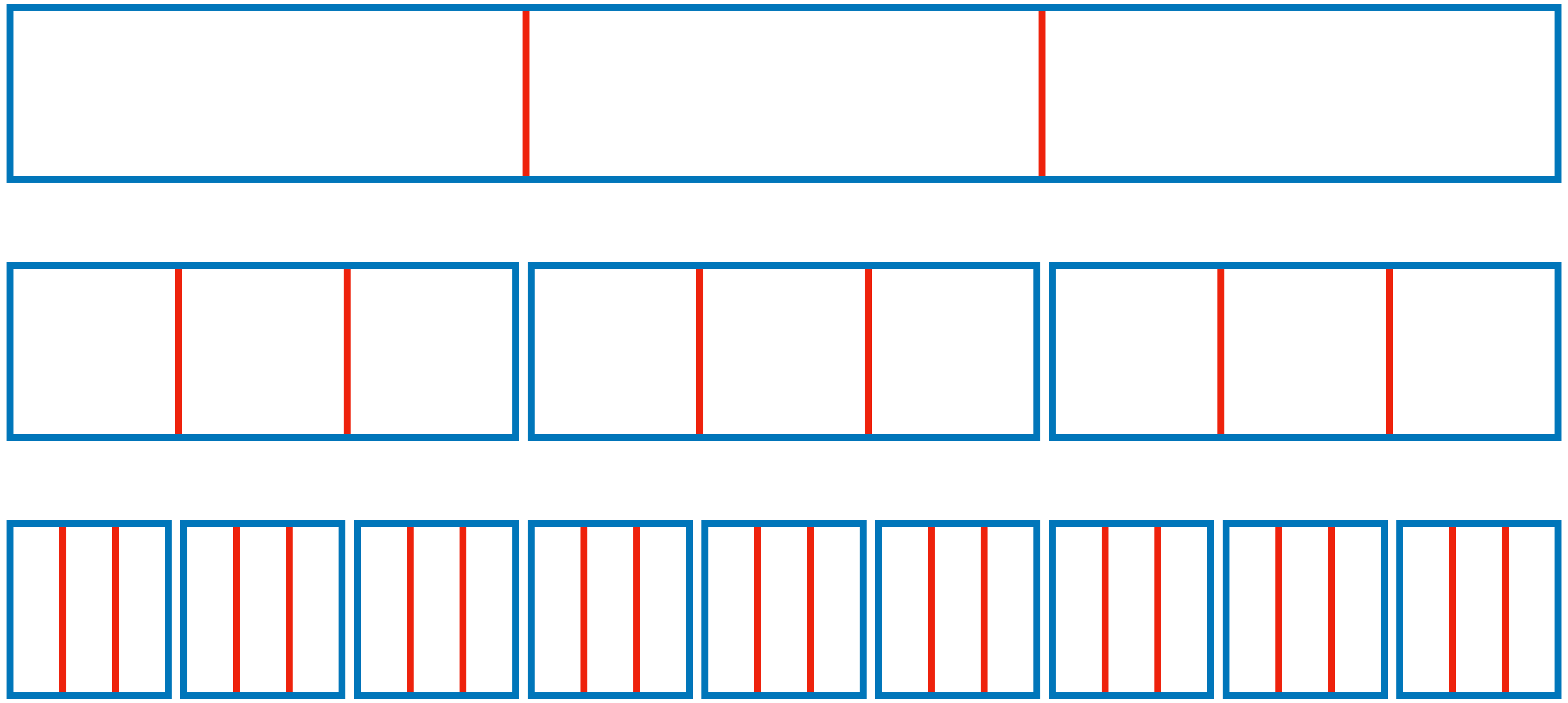
Tree-Swap

$$T = M^d$$



Tree-Swap

$$T = M^d$$



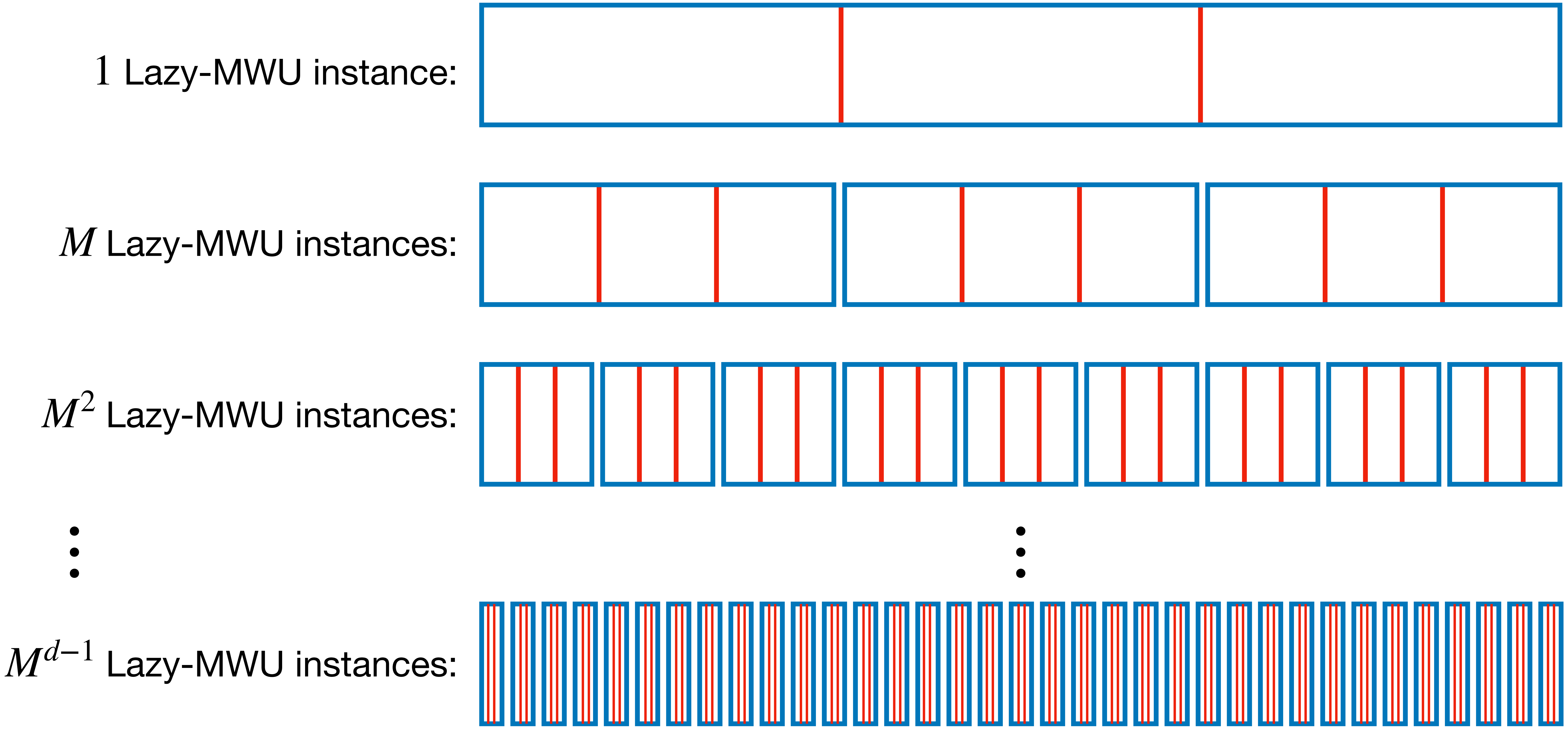
1 Lazy-MWU instance:

M Lazy-MWU instances:

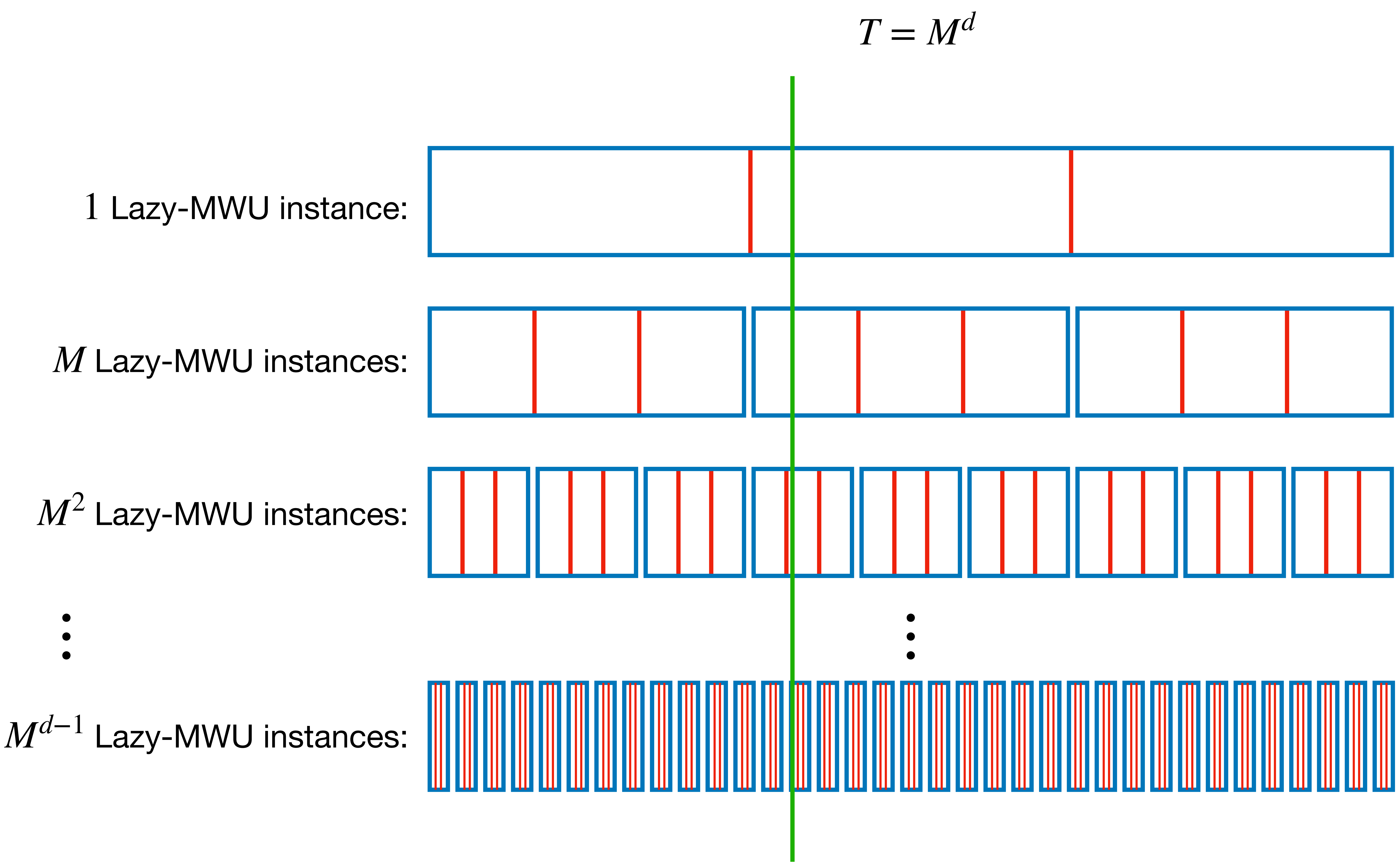
M^2 Lazy-MWU instances:

Tree-Swap

$T = M^d$



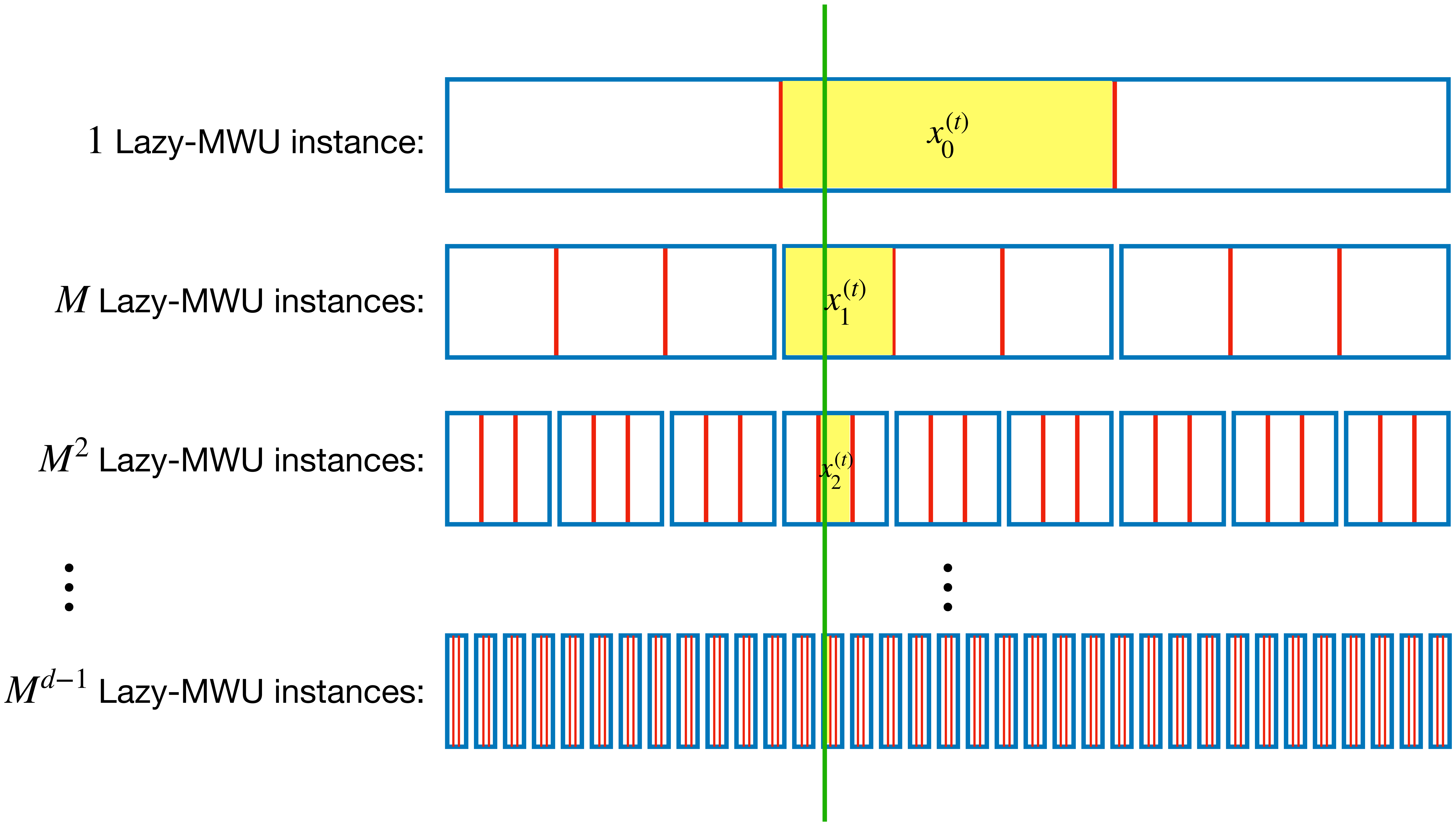
Tree-Swap



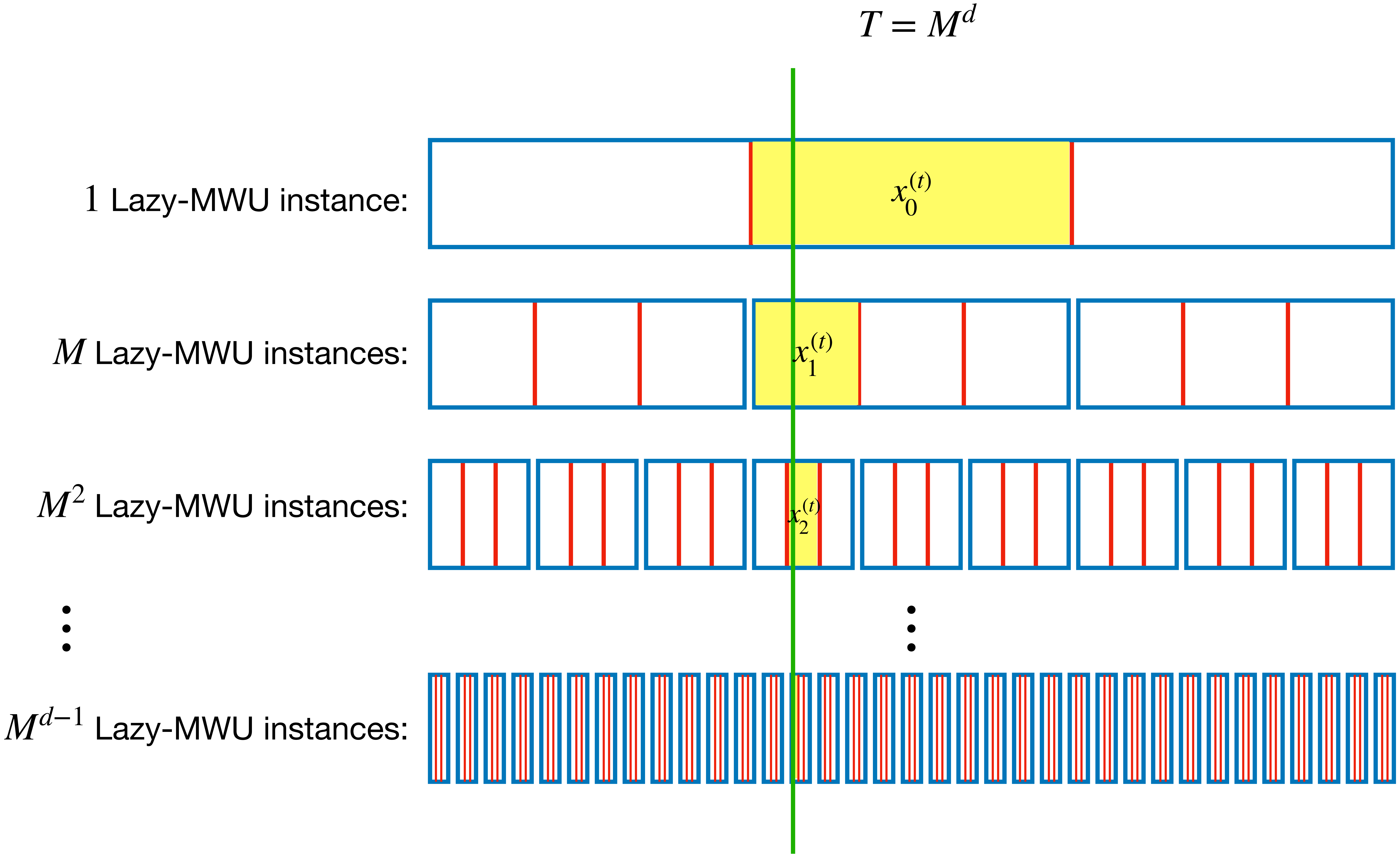
$T = M^d$
1 Lazy-MWU instance:
M Lazy-MWU instances:
M^2 Lazy-MWU instances:
M^{d-1} Lazy-MWU instances:

Tree-Swap

$T = M^d$



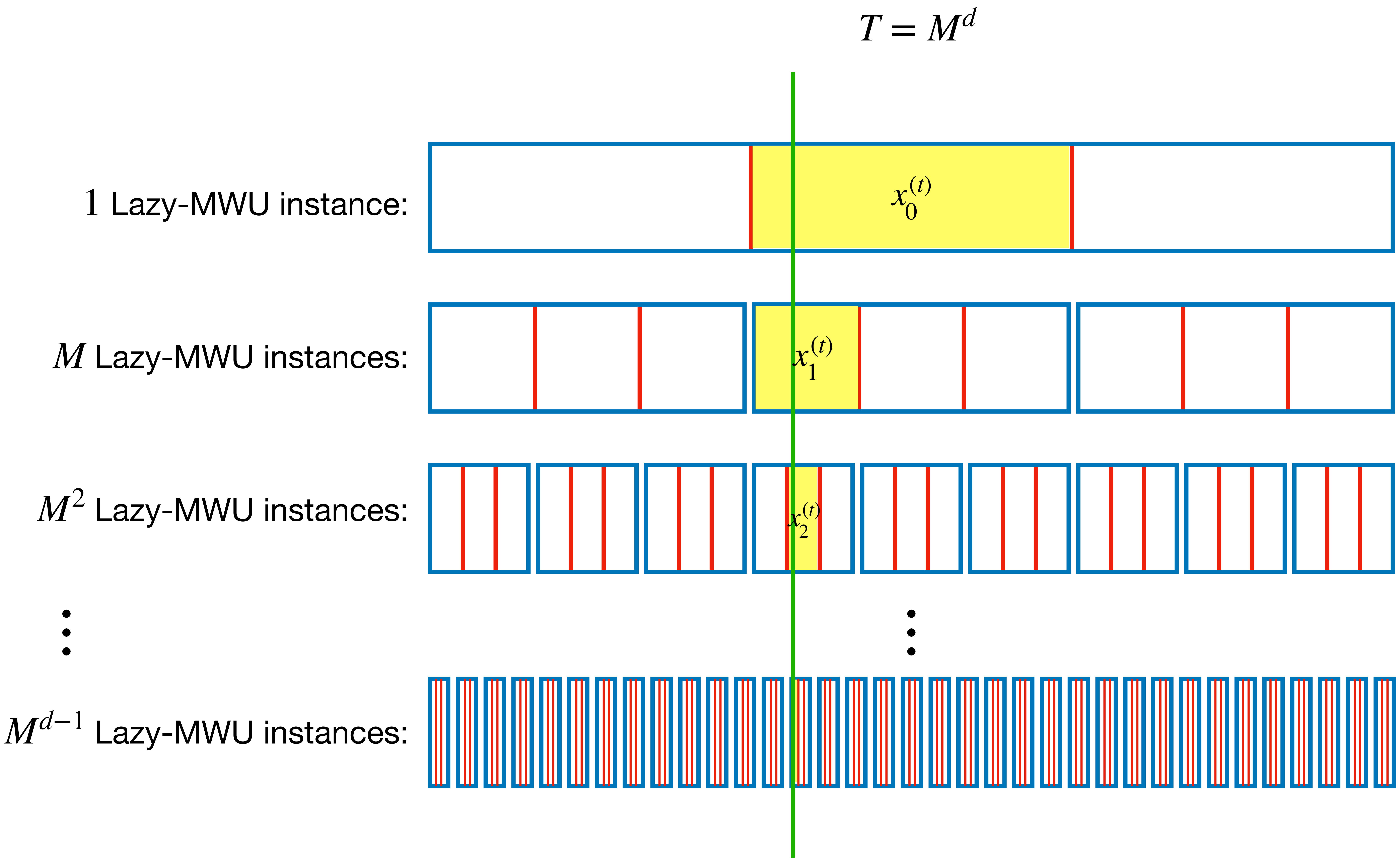
Tree-Swap



Action:

$$\frac{1}{d}\left(x_0^{(t)} + \cdots + x_{d-1}^{(t)}\right)$$

Tree-Swap

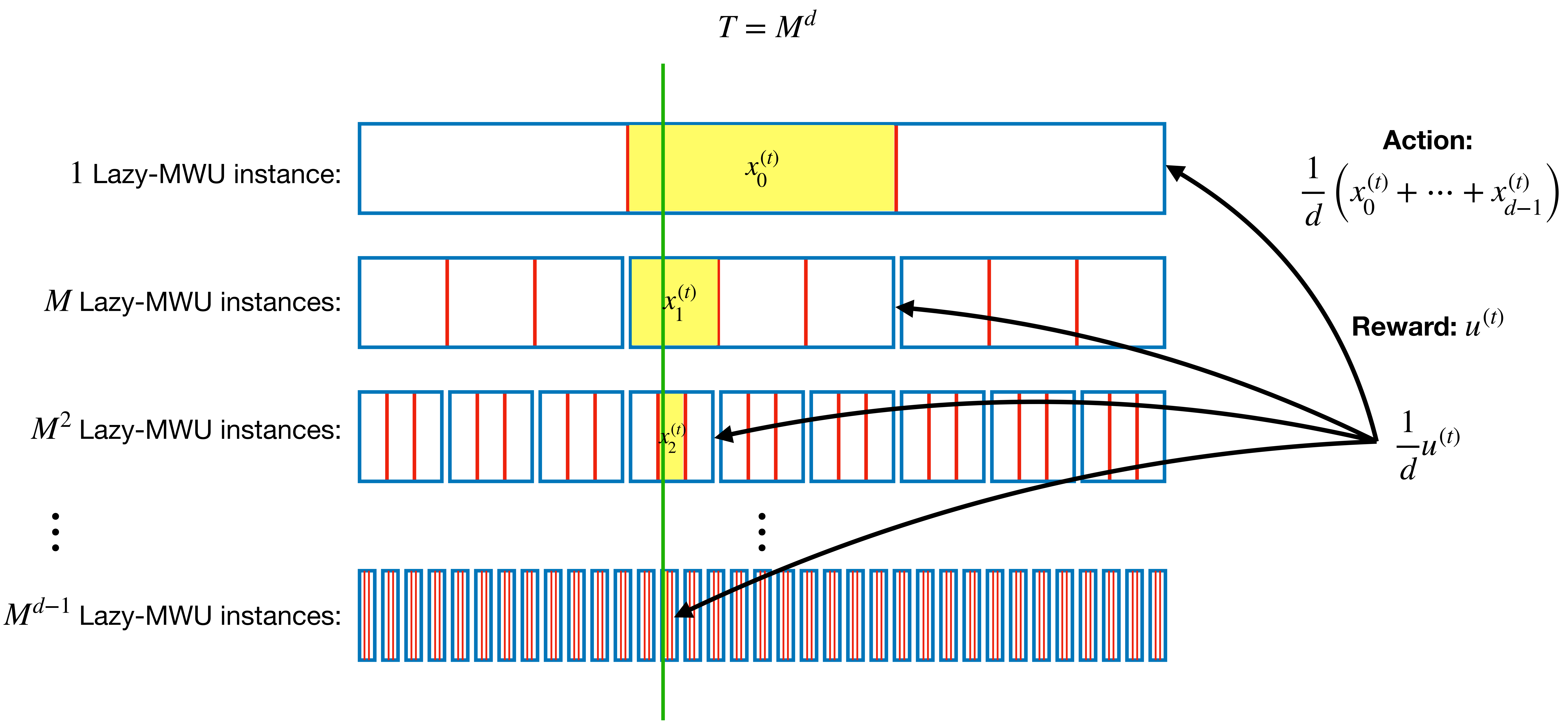


Action:

$$\frac{1}{d}\left(x_0^{(t)} + \cdots + x_{d-1}^{(t)}\right)$$

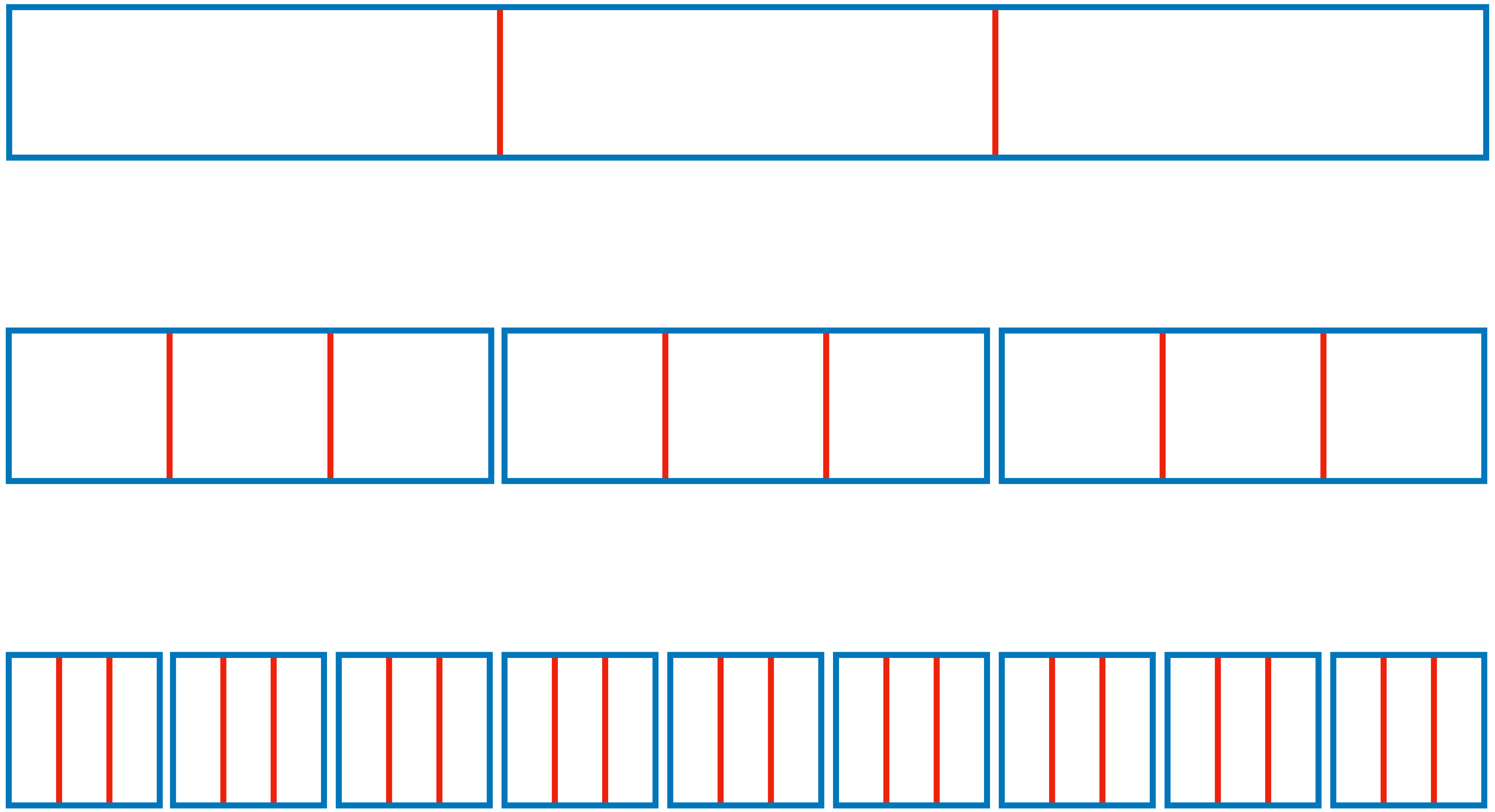
Reward: $u^{(t)}$

Tree-Swap

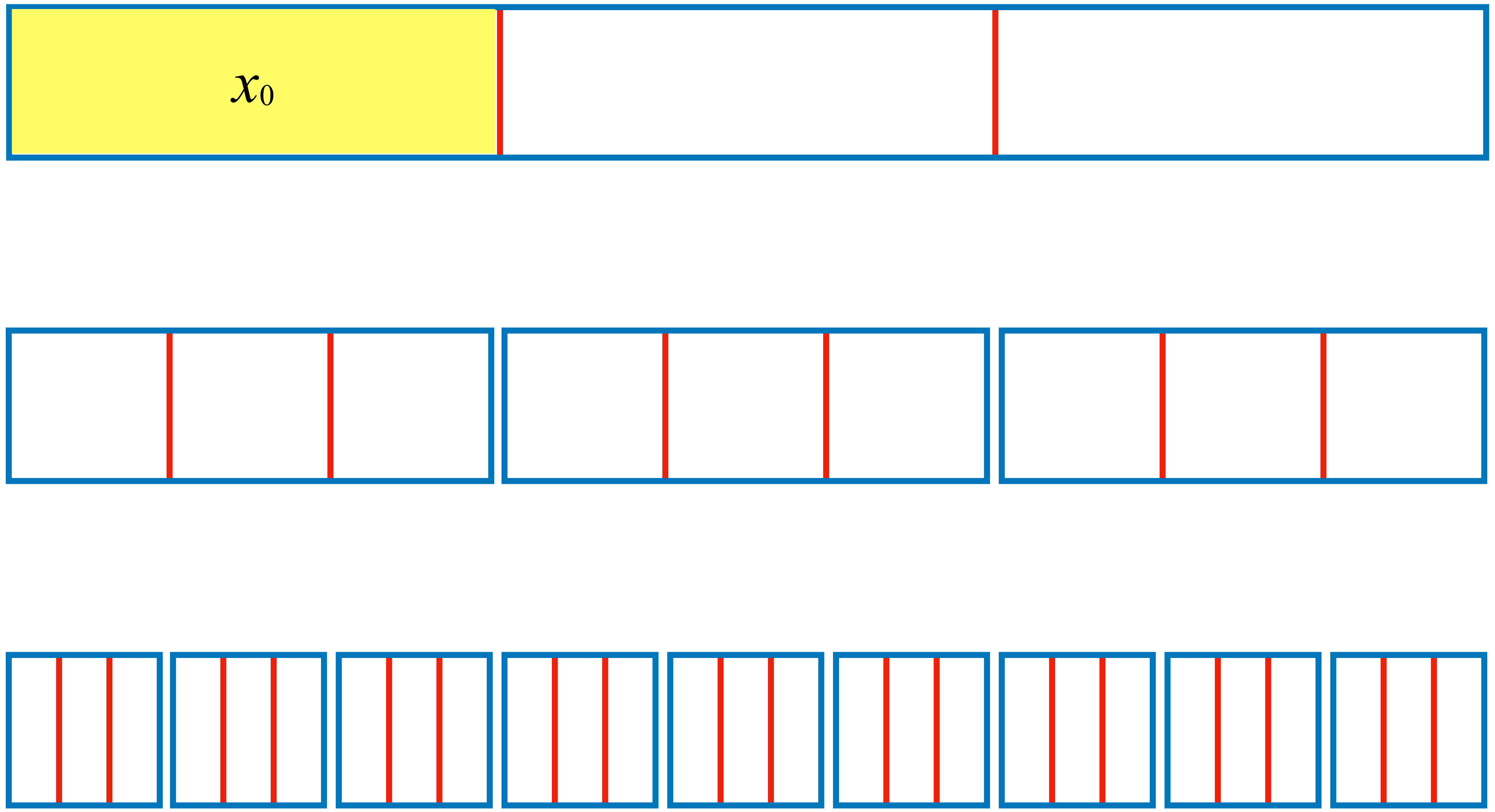


$T = M^d$
1 Lazy-MWU instance:
M Lazy-MWU instances:
M^2 Lazy-MWU instances:
M^{d-1} Lazy-MWU instances:
$x_0^{(t)}$
$x_1^{(t)}$
$x_2^{(t)}$
Action:
$\frac{1}{d}\left(x_0^{(t)} + \cdots + x_{d-1}^{(t)}\right)$
Reward: $u^{(t)}$
$\frac{1}{d}u^{(t)}$

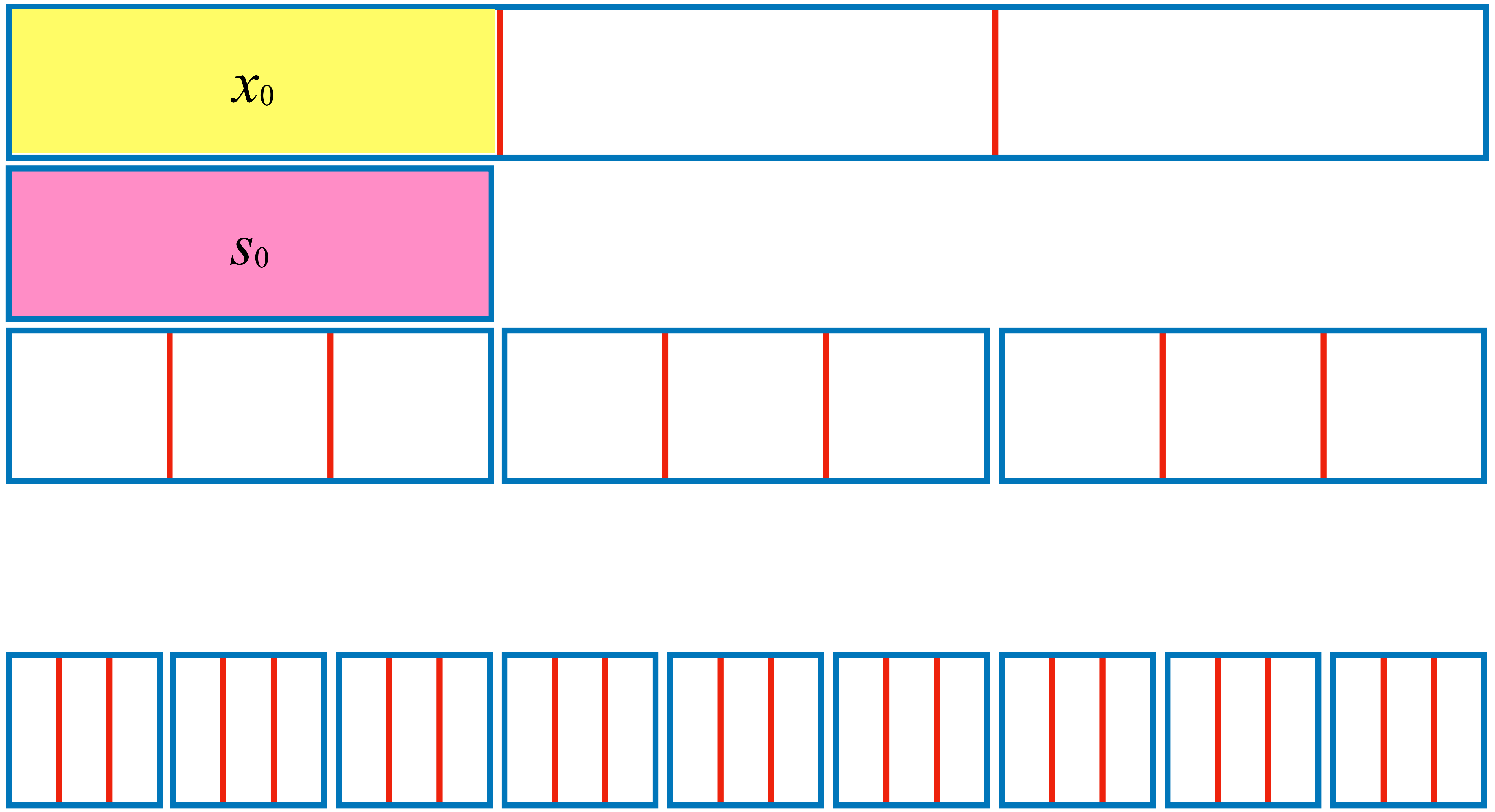
Tree-Swap Swap Regret



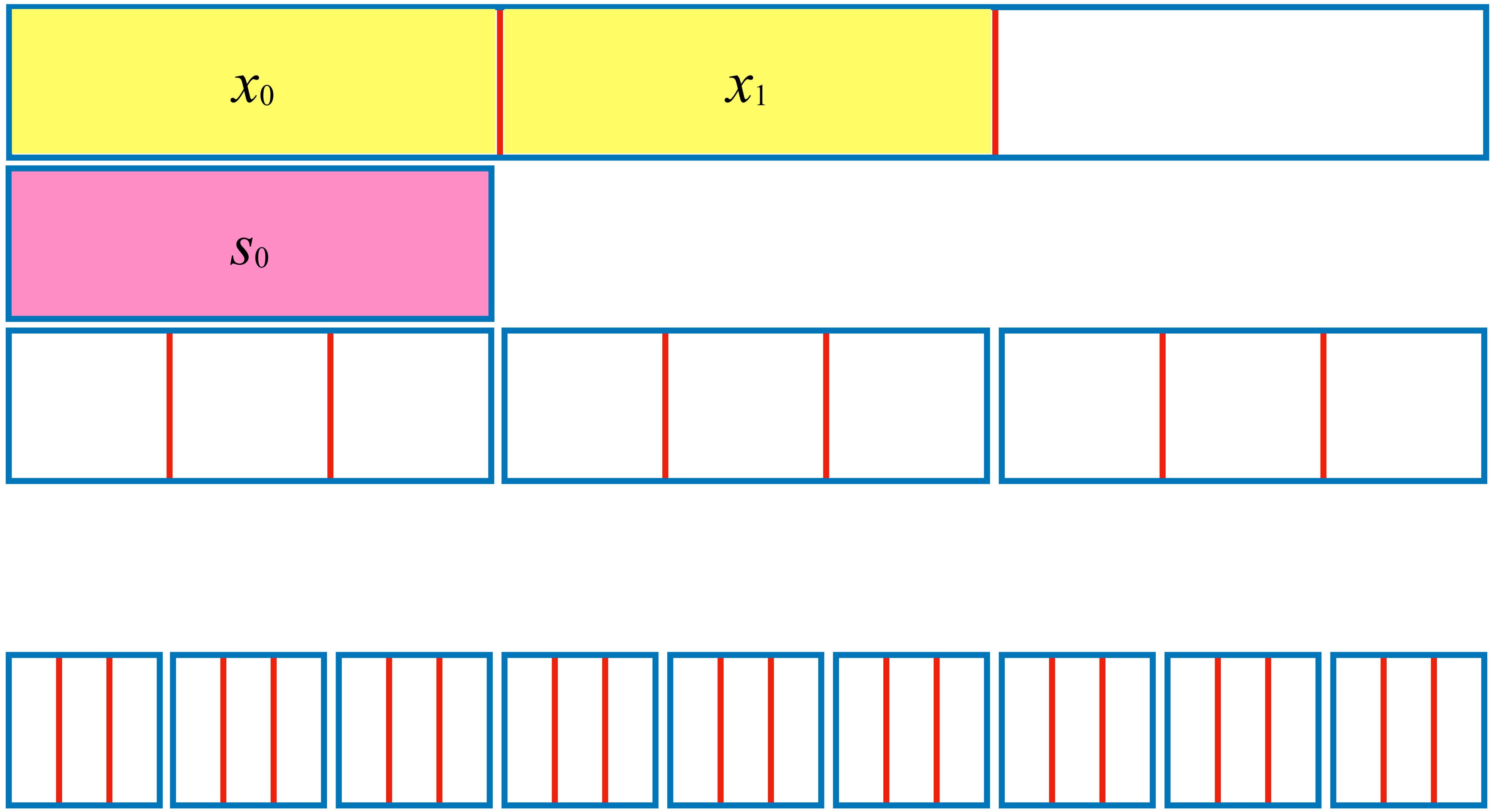
Tree-Swap Swap Regret



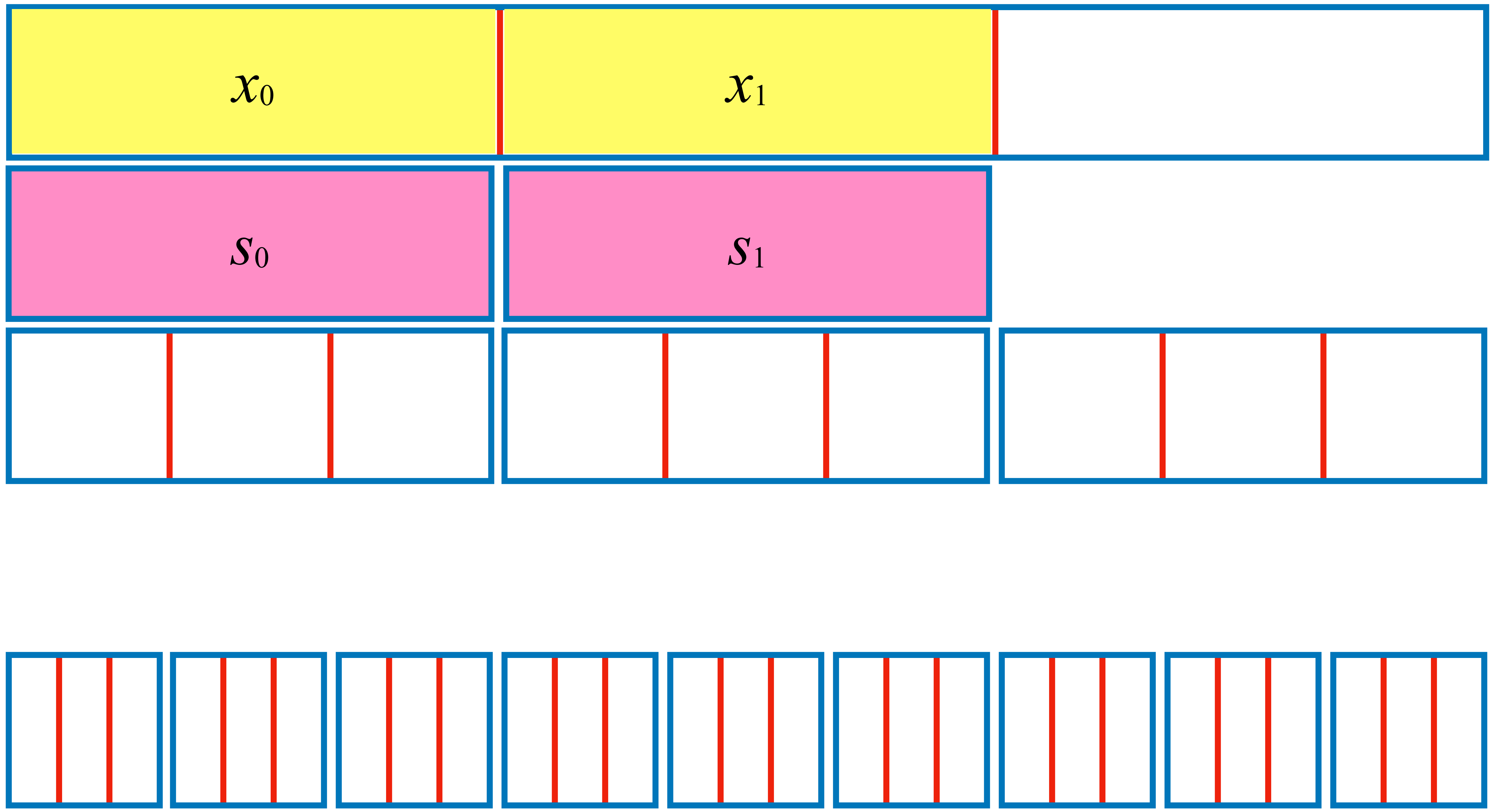
Tree-Swap Swap Regret



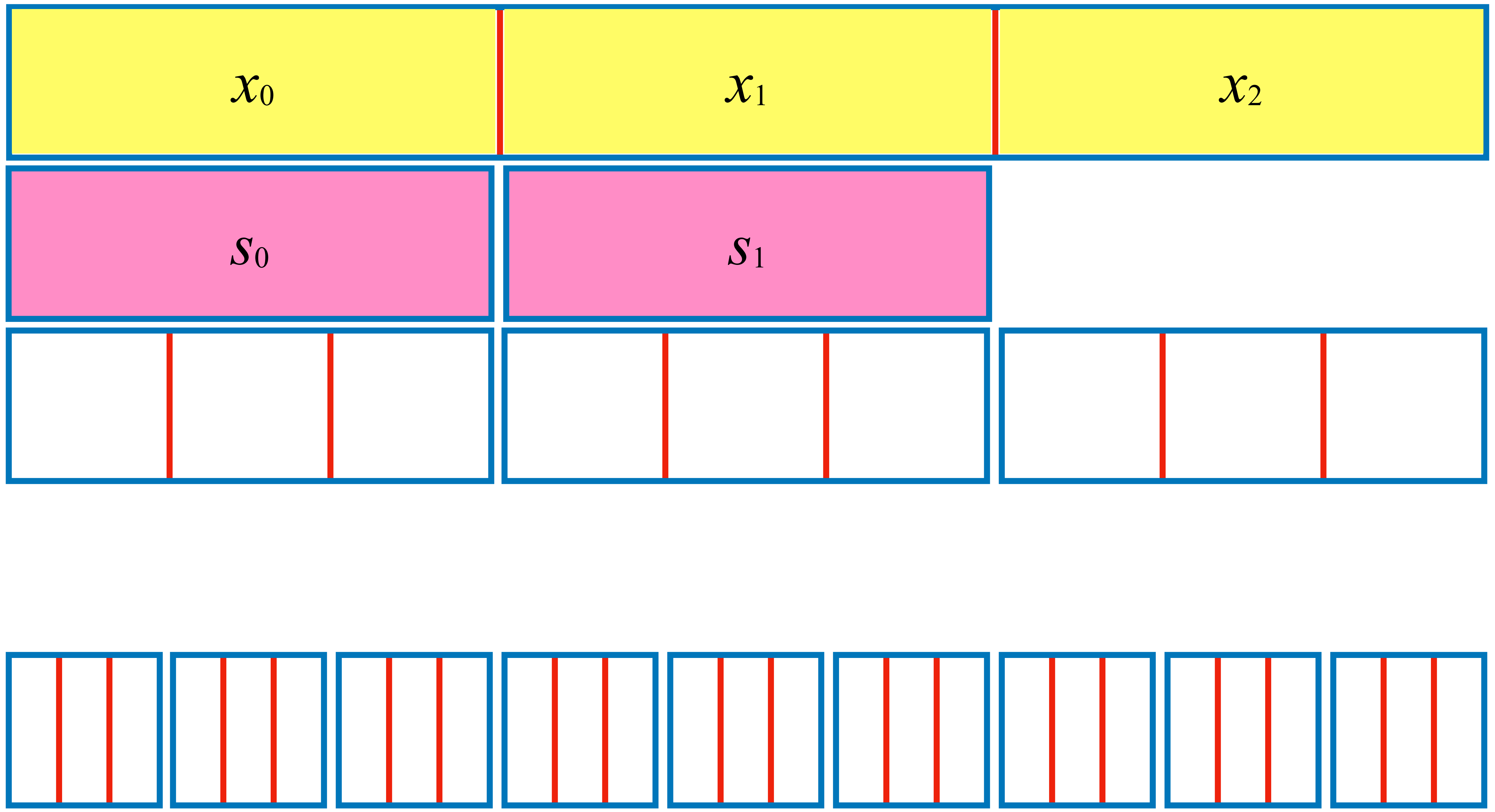
Tree-Swap Swap Regret



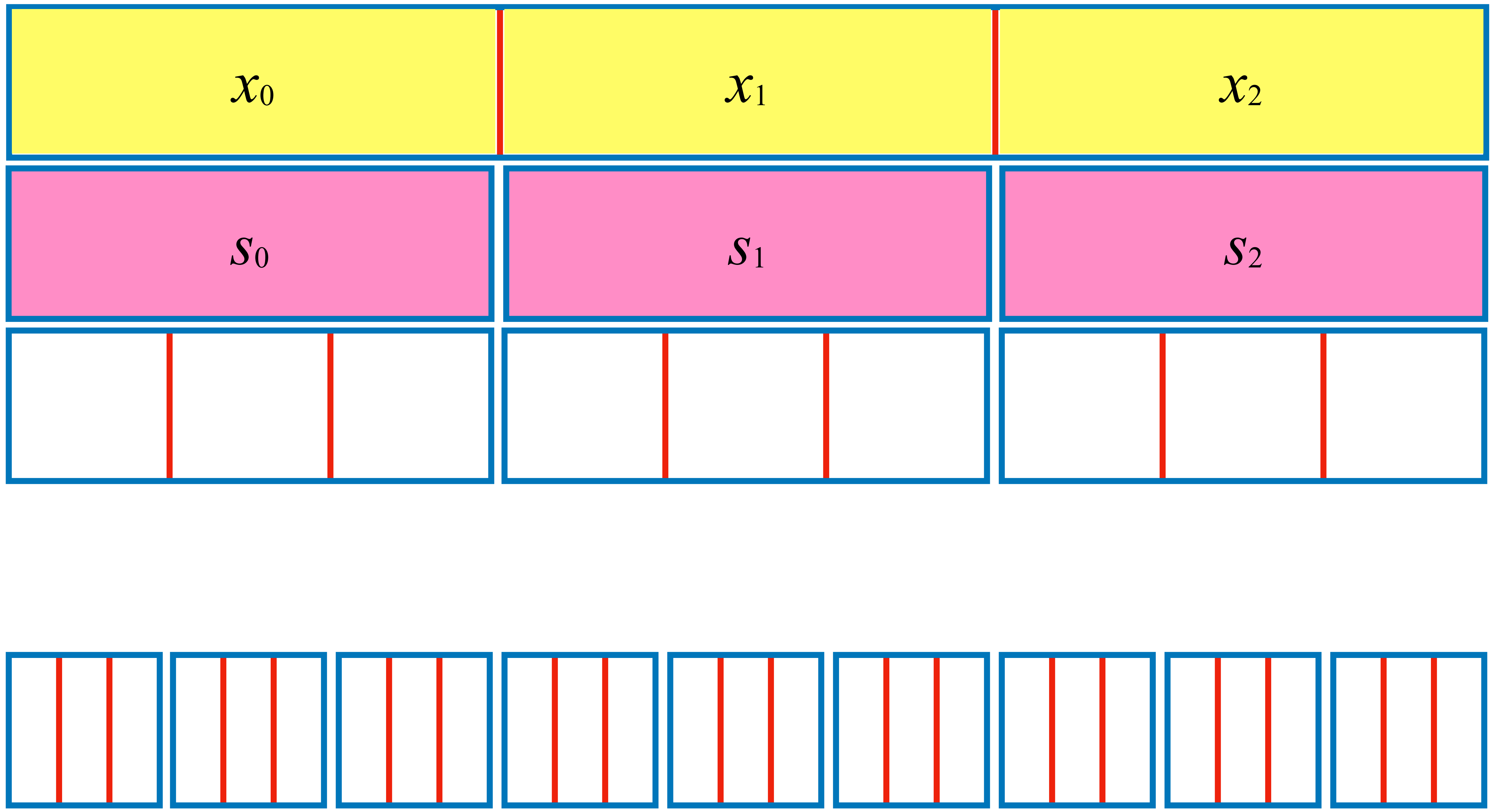
Tree-Swap Swap Regret



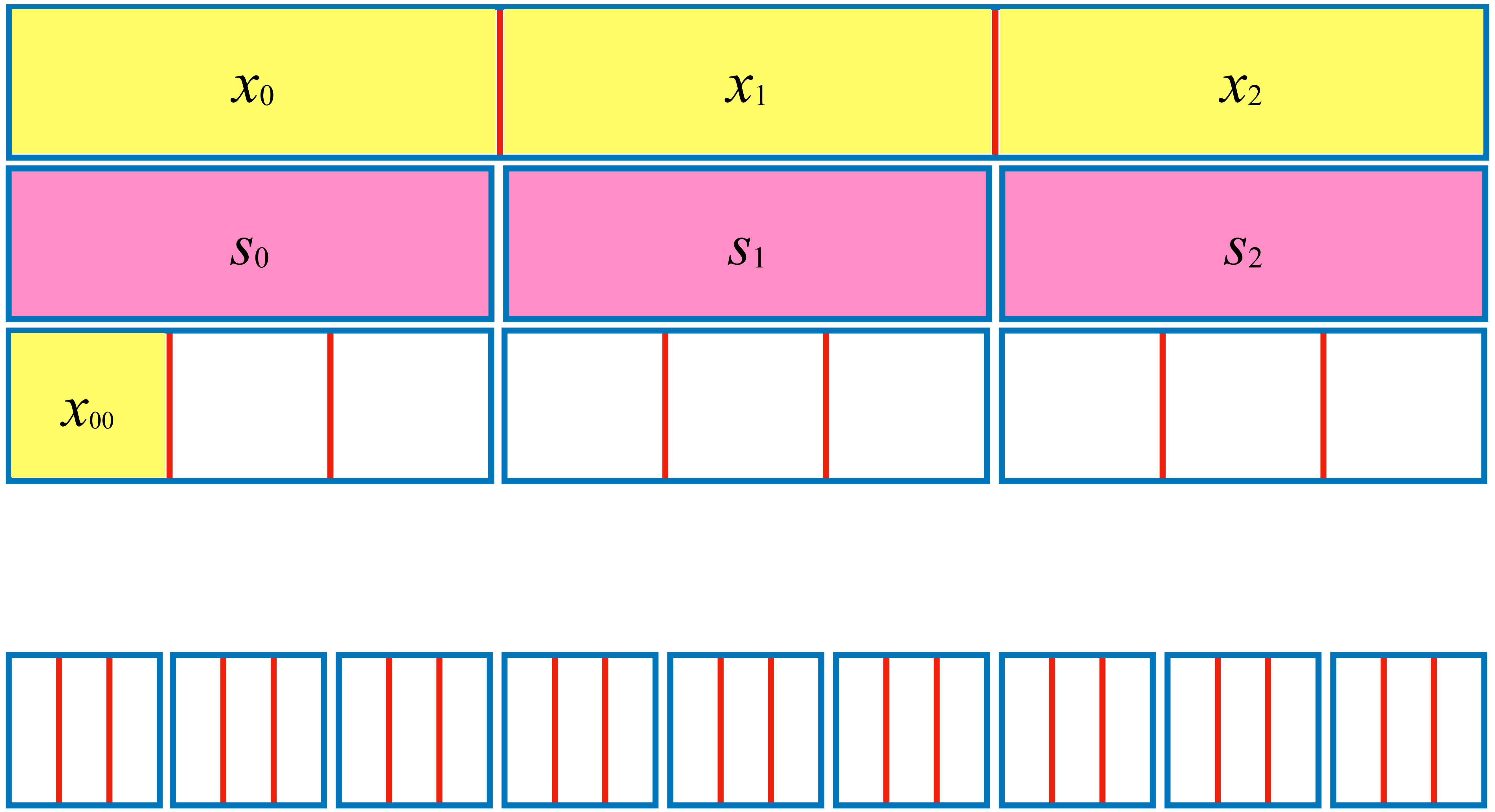
Tree-Swap Swap Regret



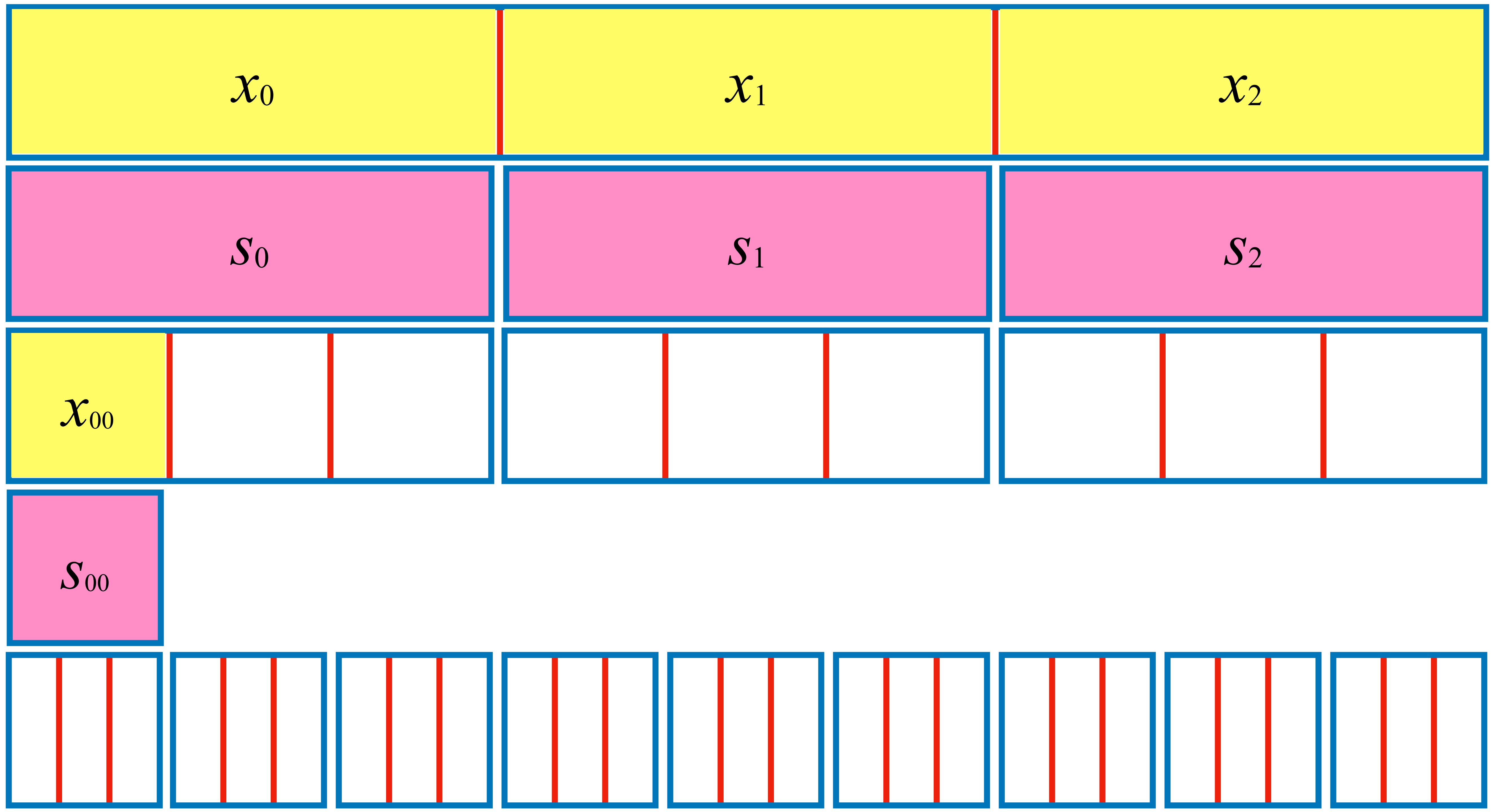
Tree-Swap Swap Regret



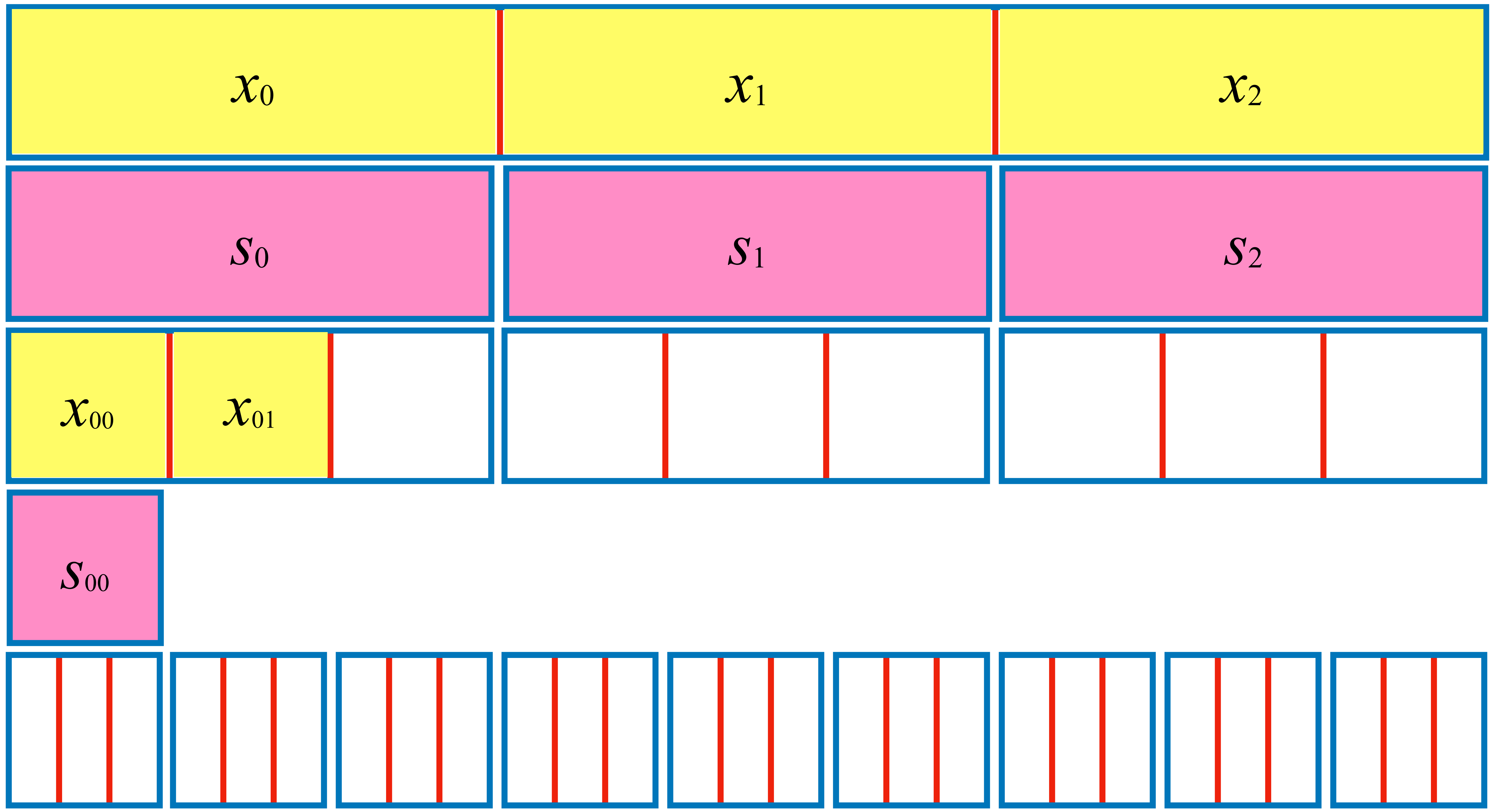
Tree-Swap Swap Regret



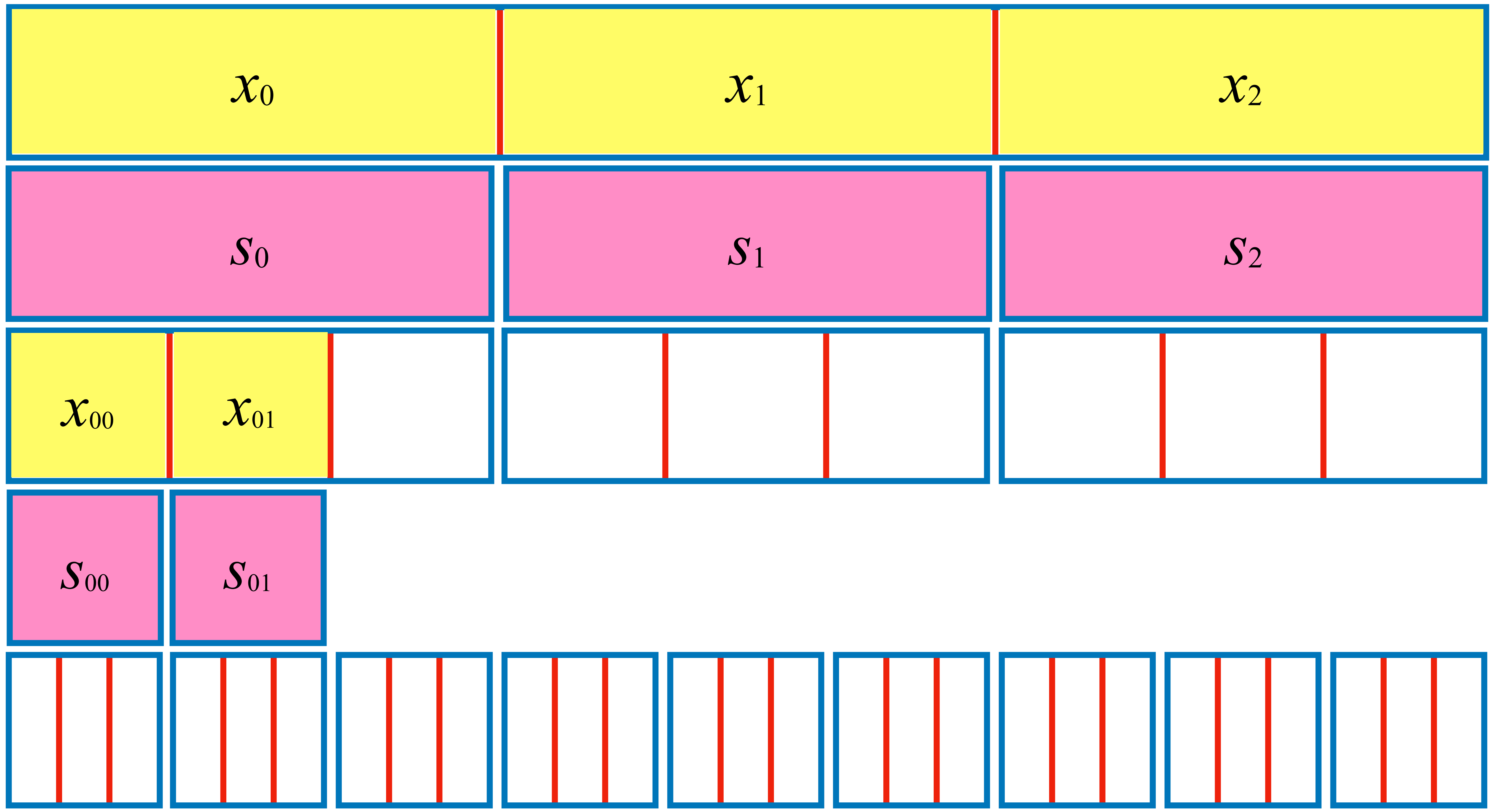
Tree-Swap Swap Regret



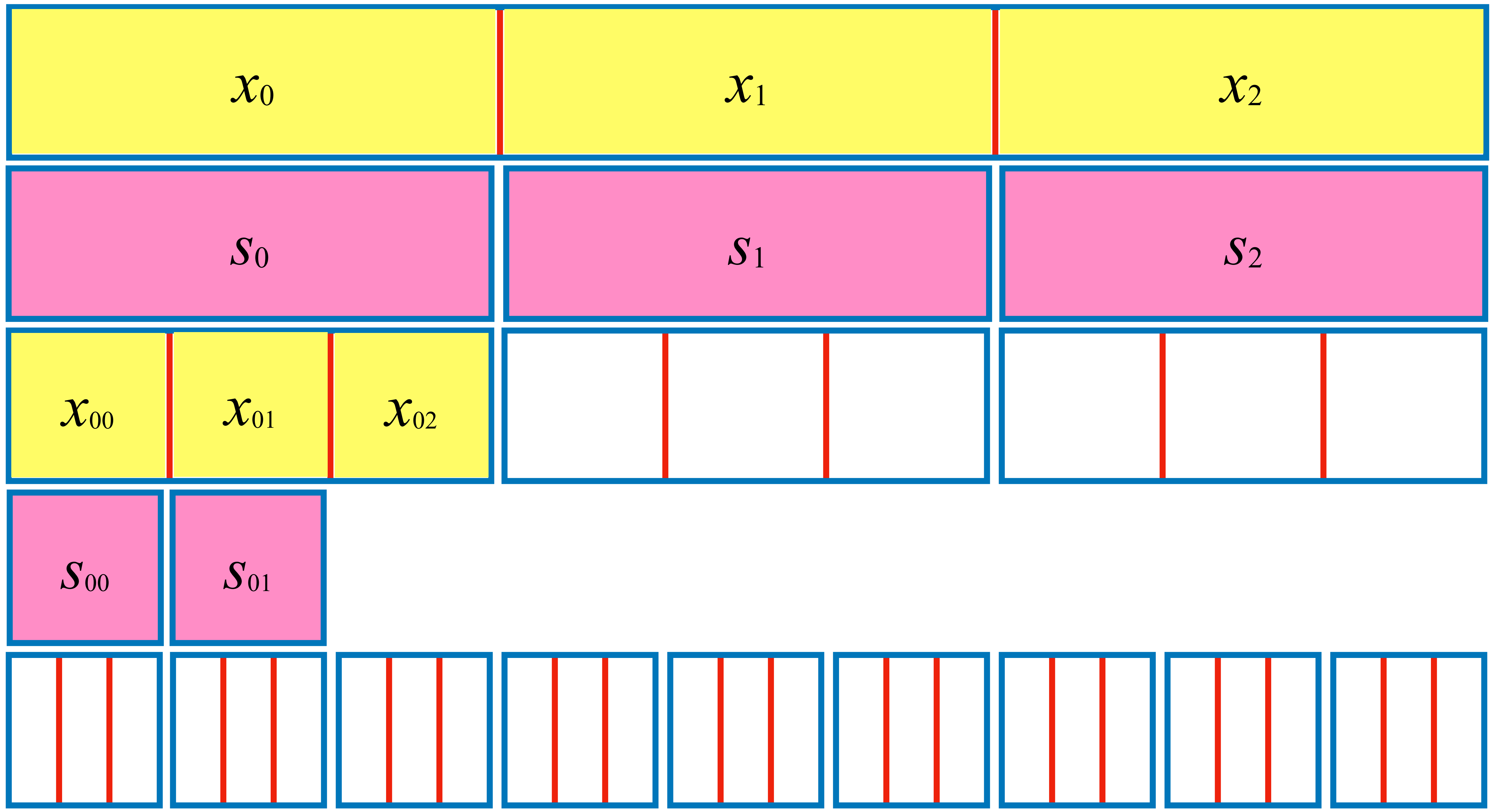
Tree-Swap Swap Regret



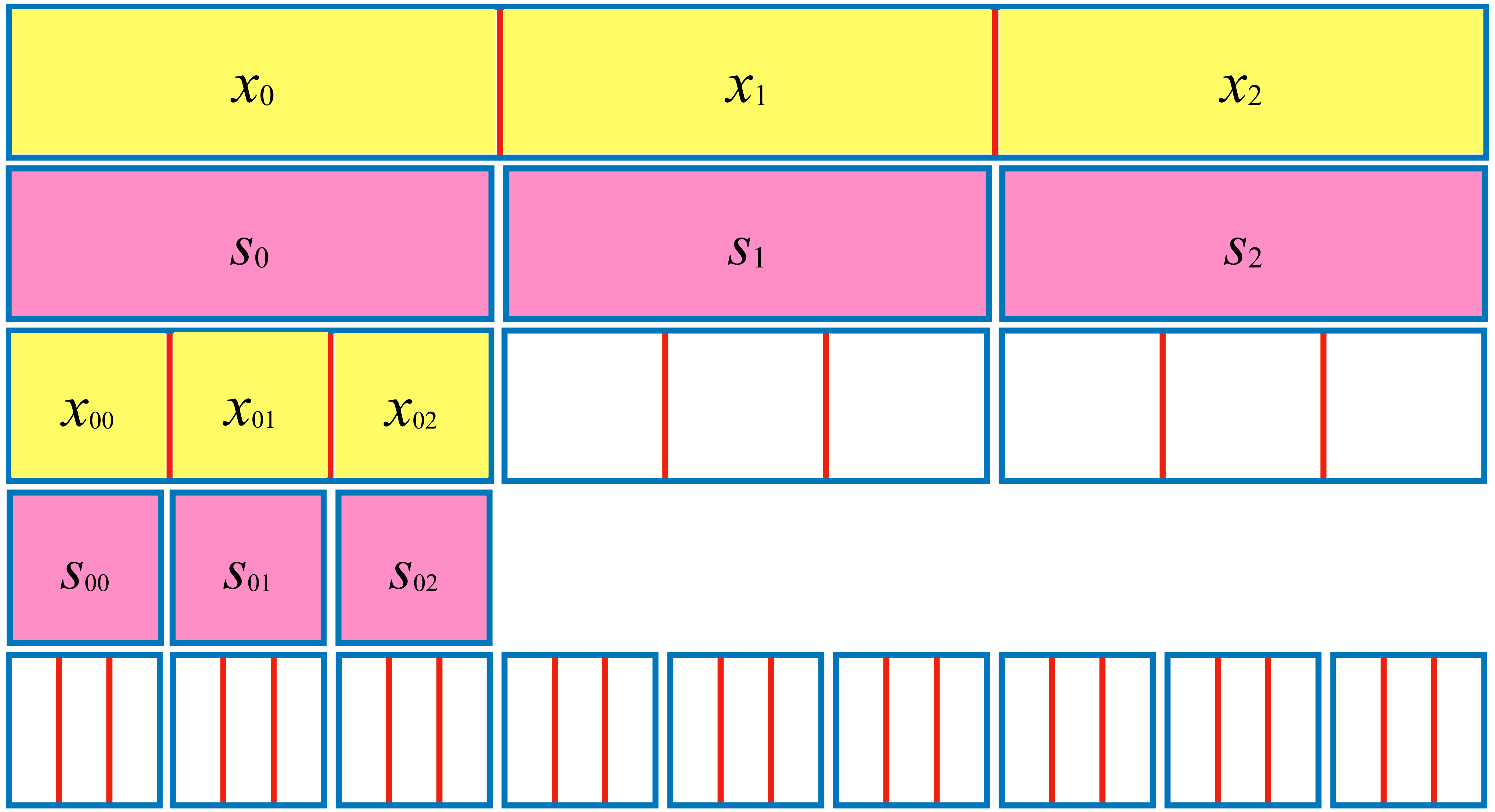
Tree-Swap Swap Regret



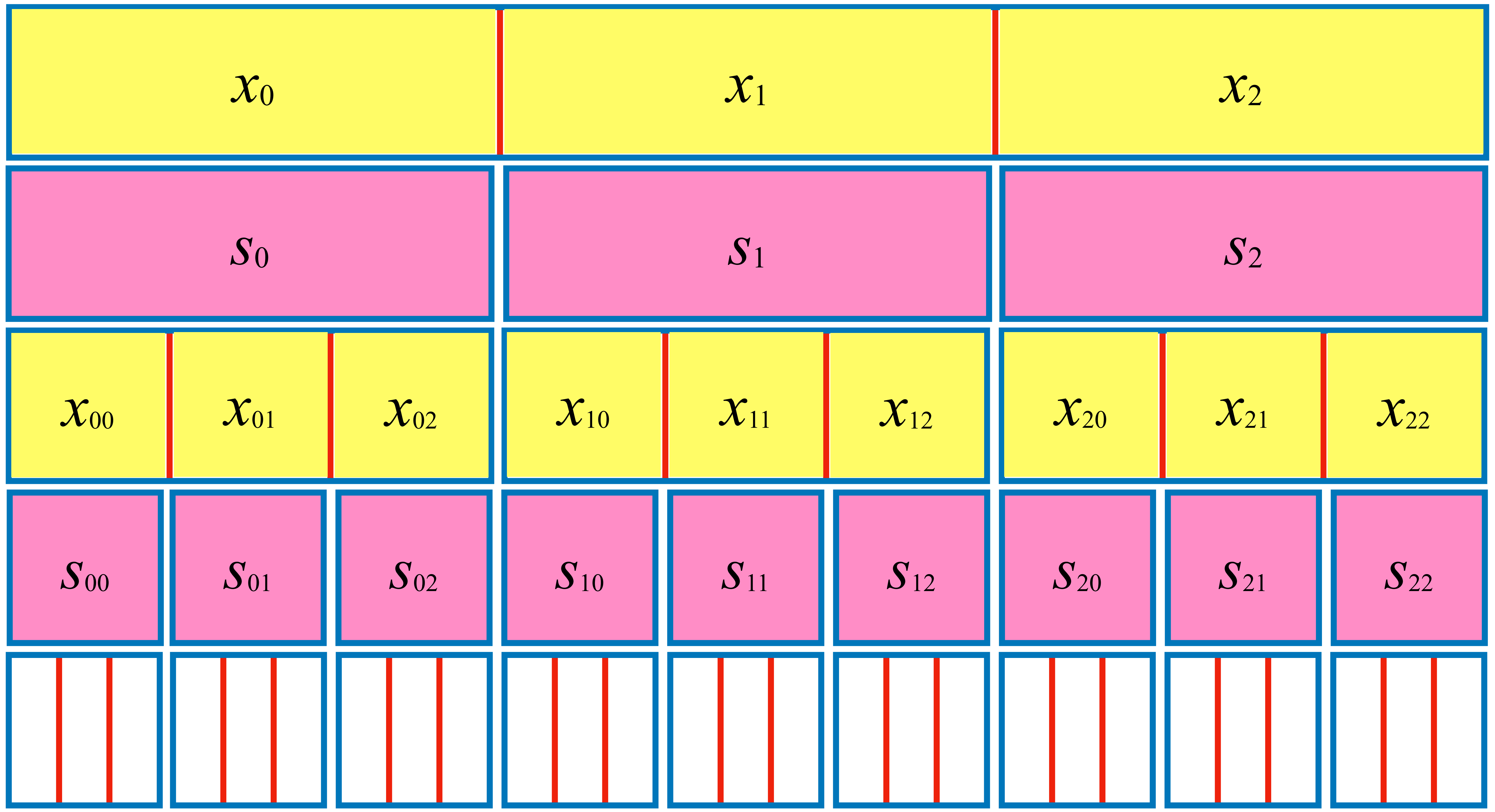
Tree-Swap Swap Regret



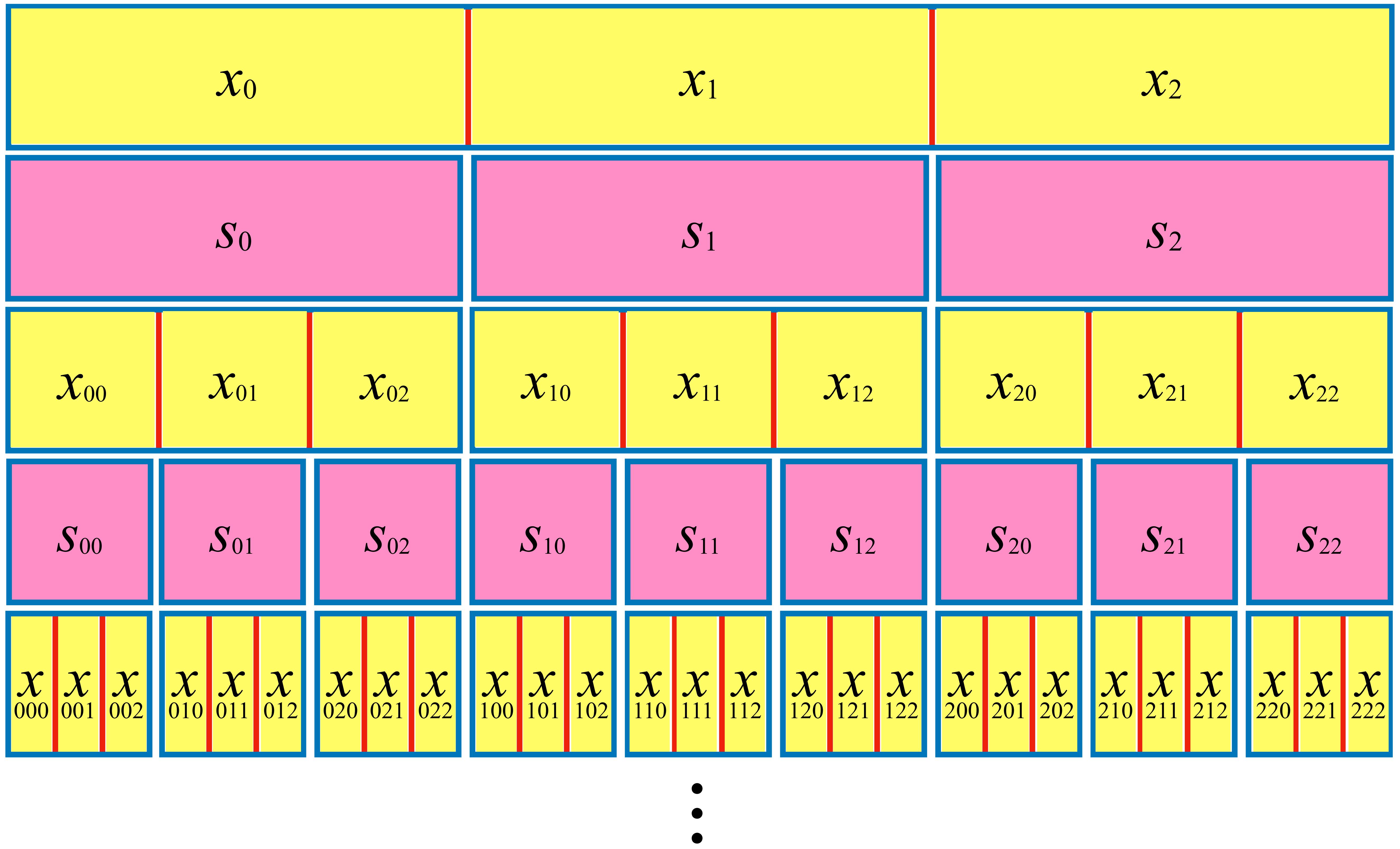
Tree-Swap Swap Regret



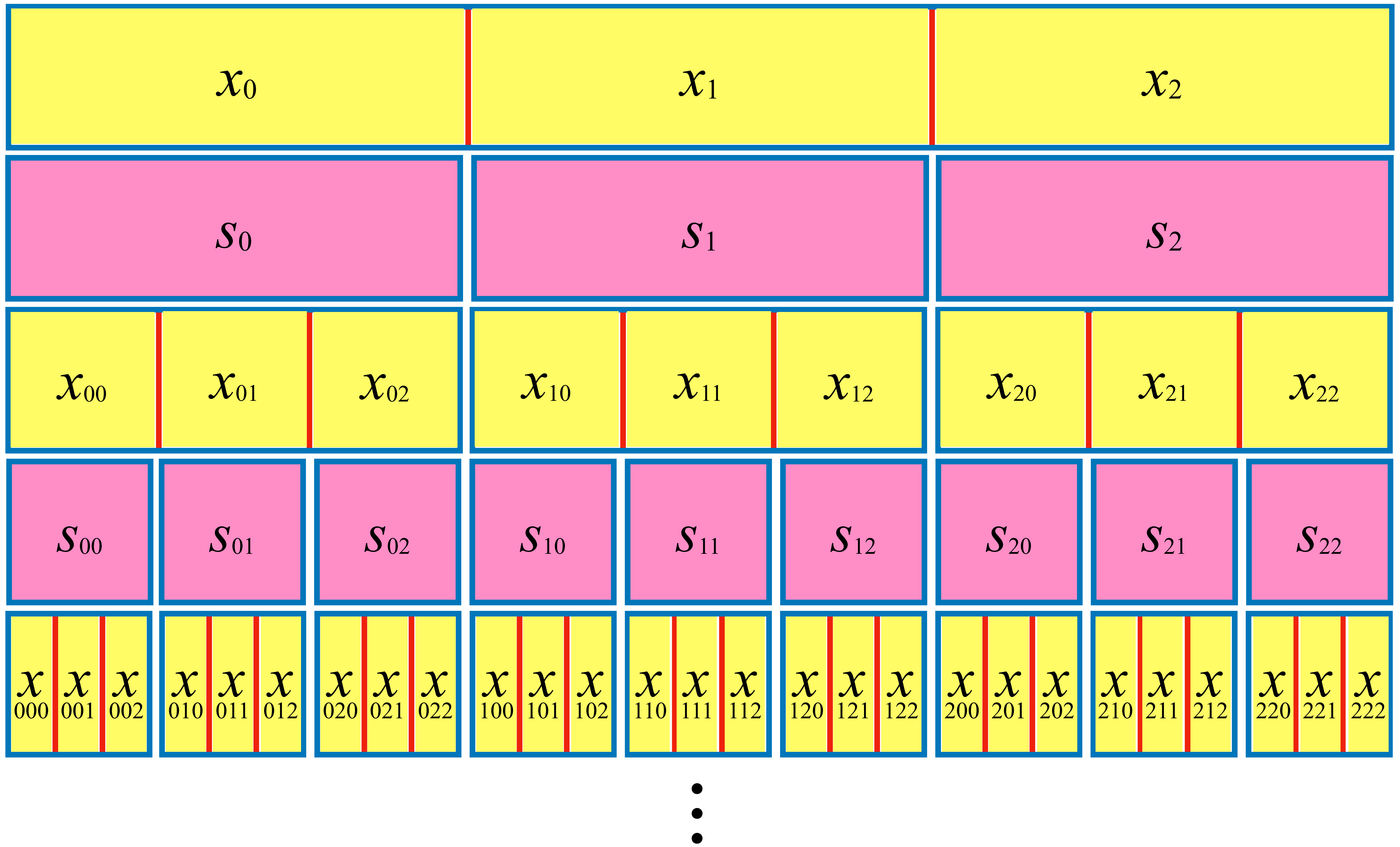
Tree-Swap Swap Regret



Tree-Swap Swap Regret

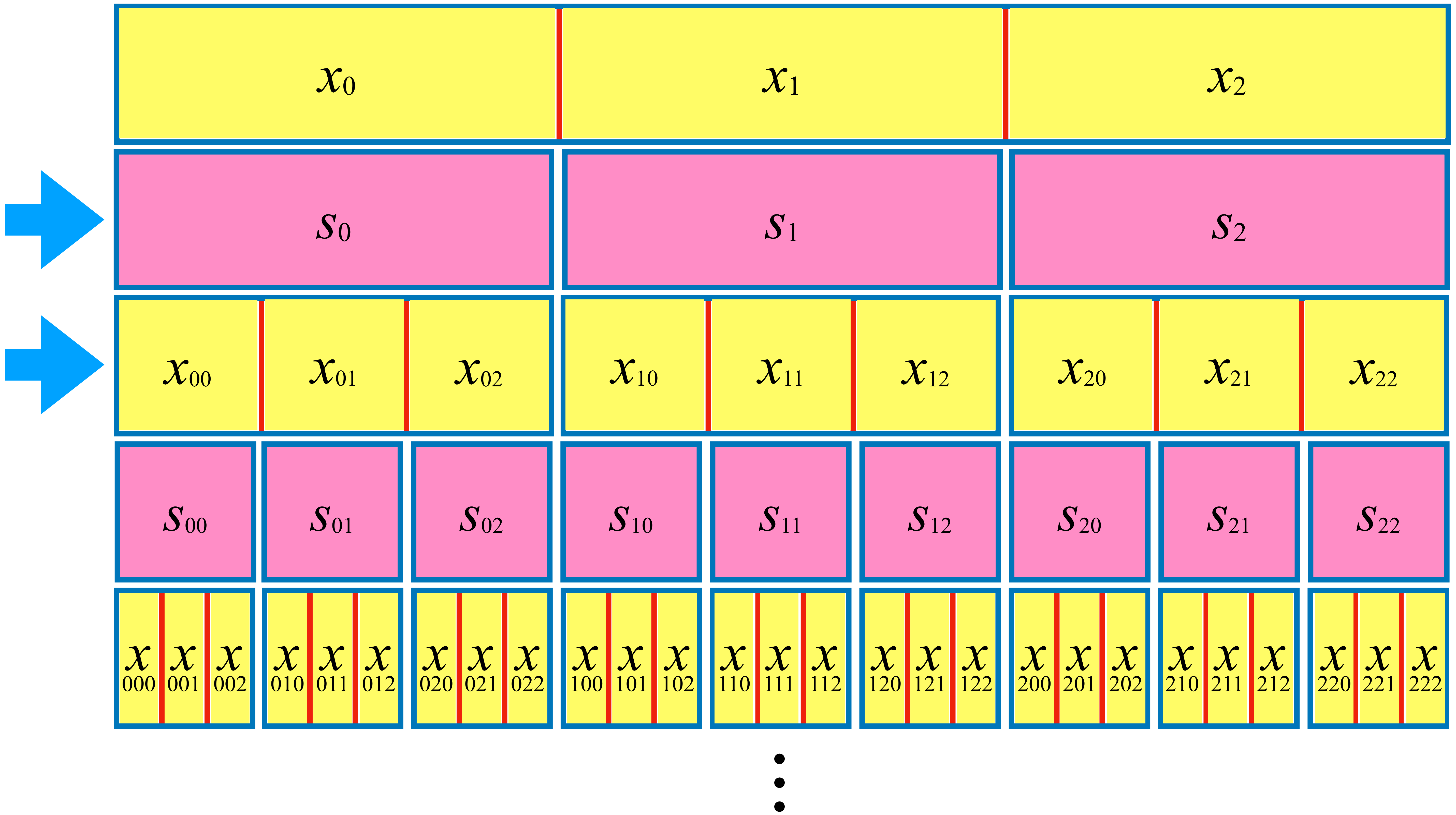


Tree-Swap Swap Regret



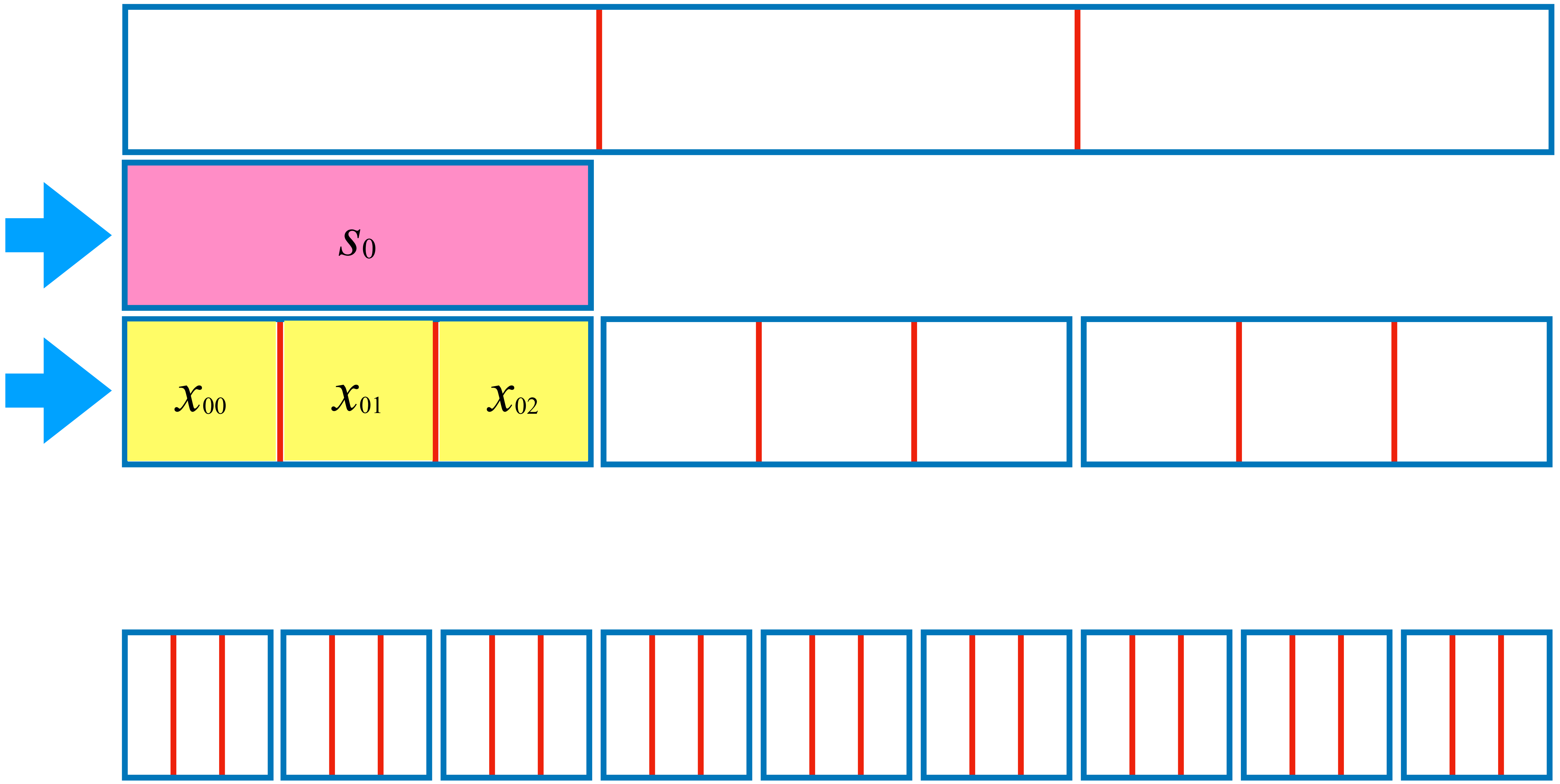
Swap Regret ≤ Swap Reward - Reward

Tree-Swap Swap Regret



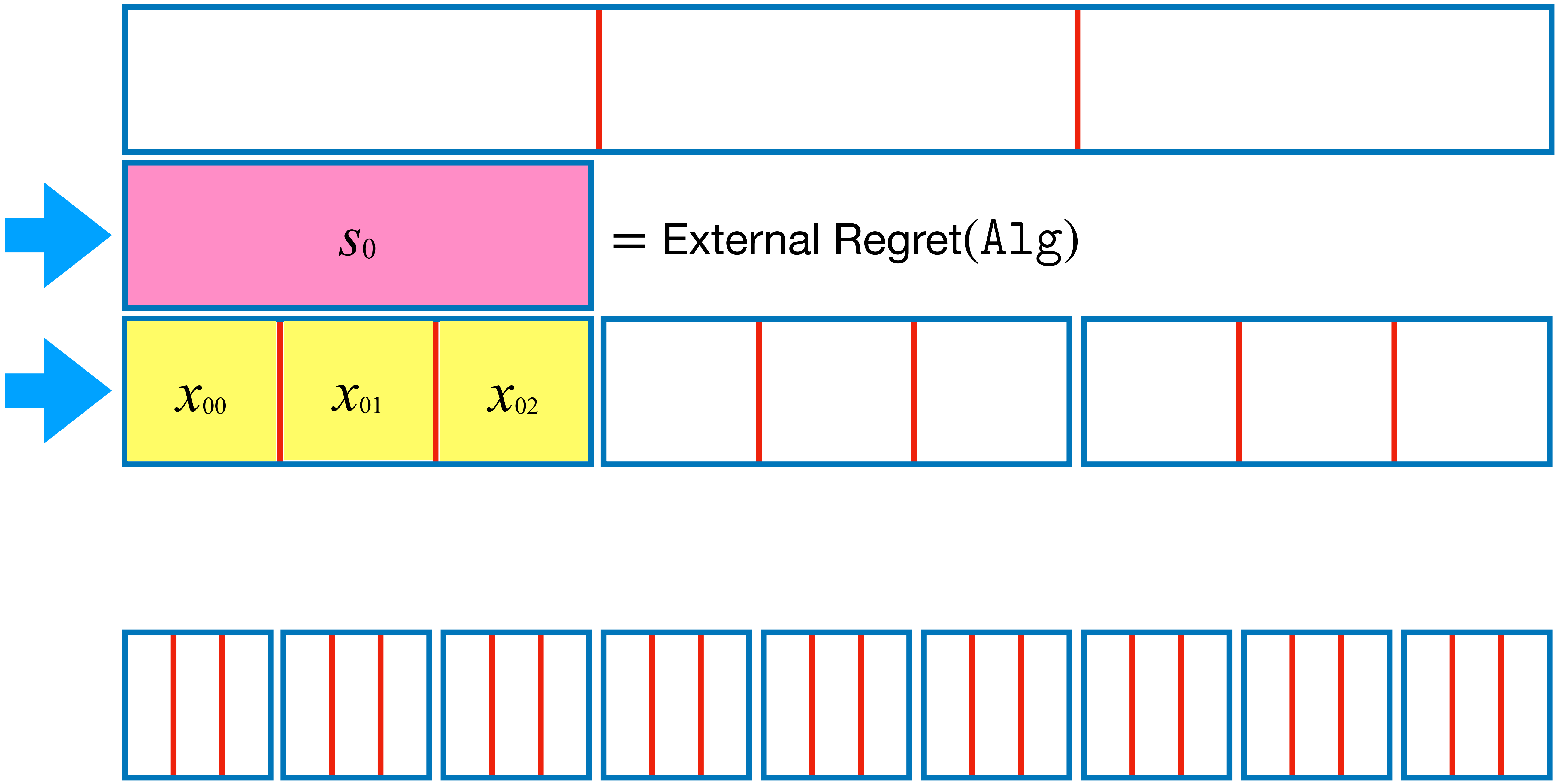
Swap Regret ≤ Swap Reward - Reward

Tree-Swap Swap Regret



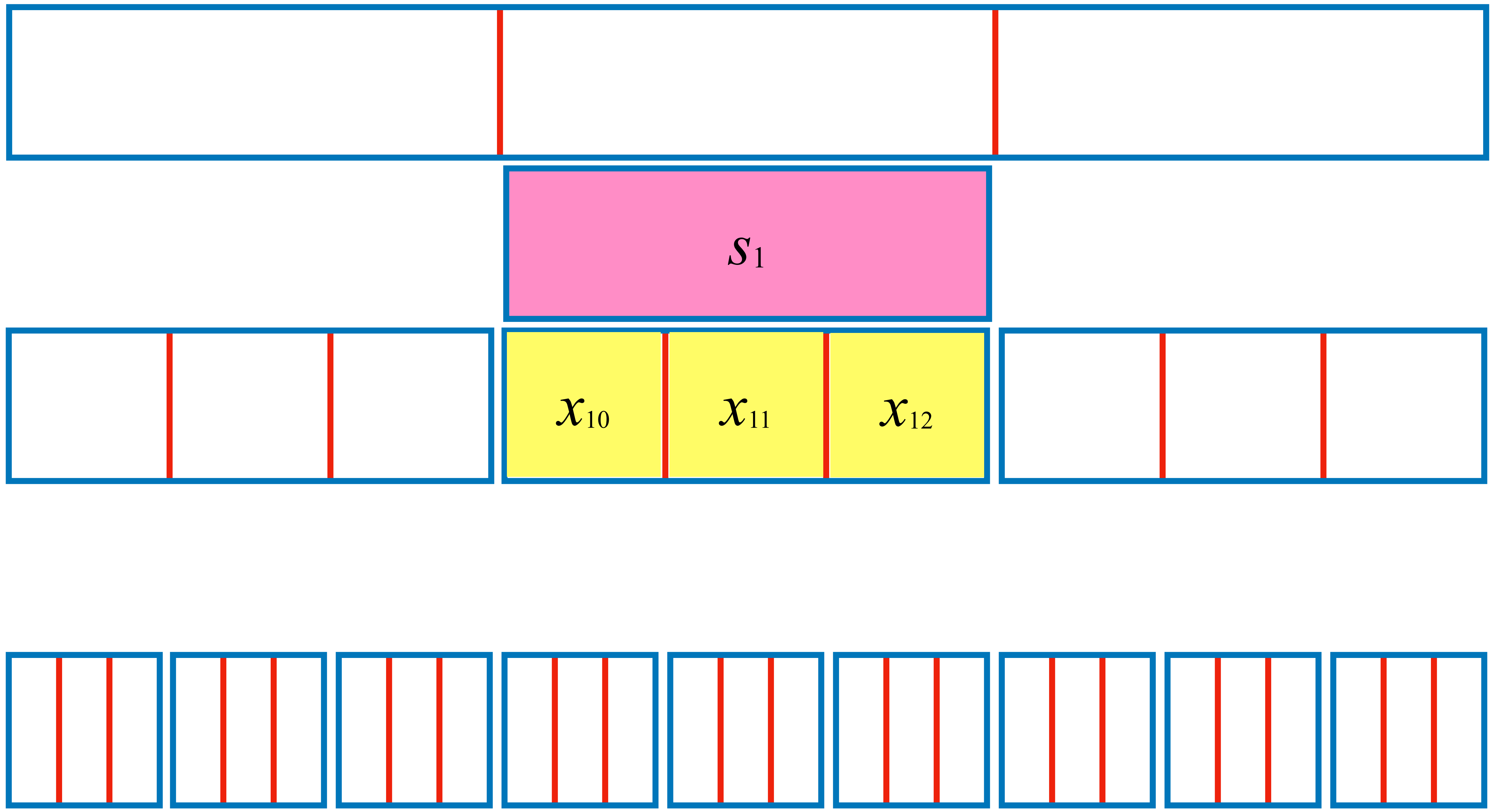
Swap Regret ≤ Swap Reward - Reward

Tree-Swap Swap Regret



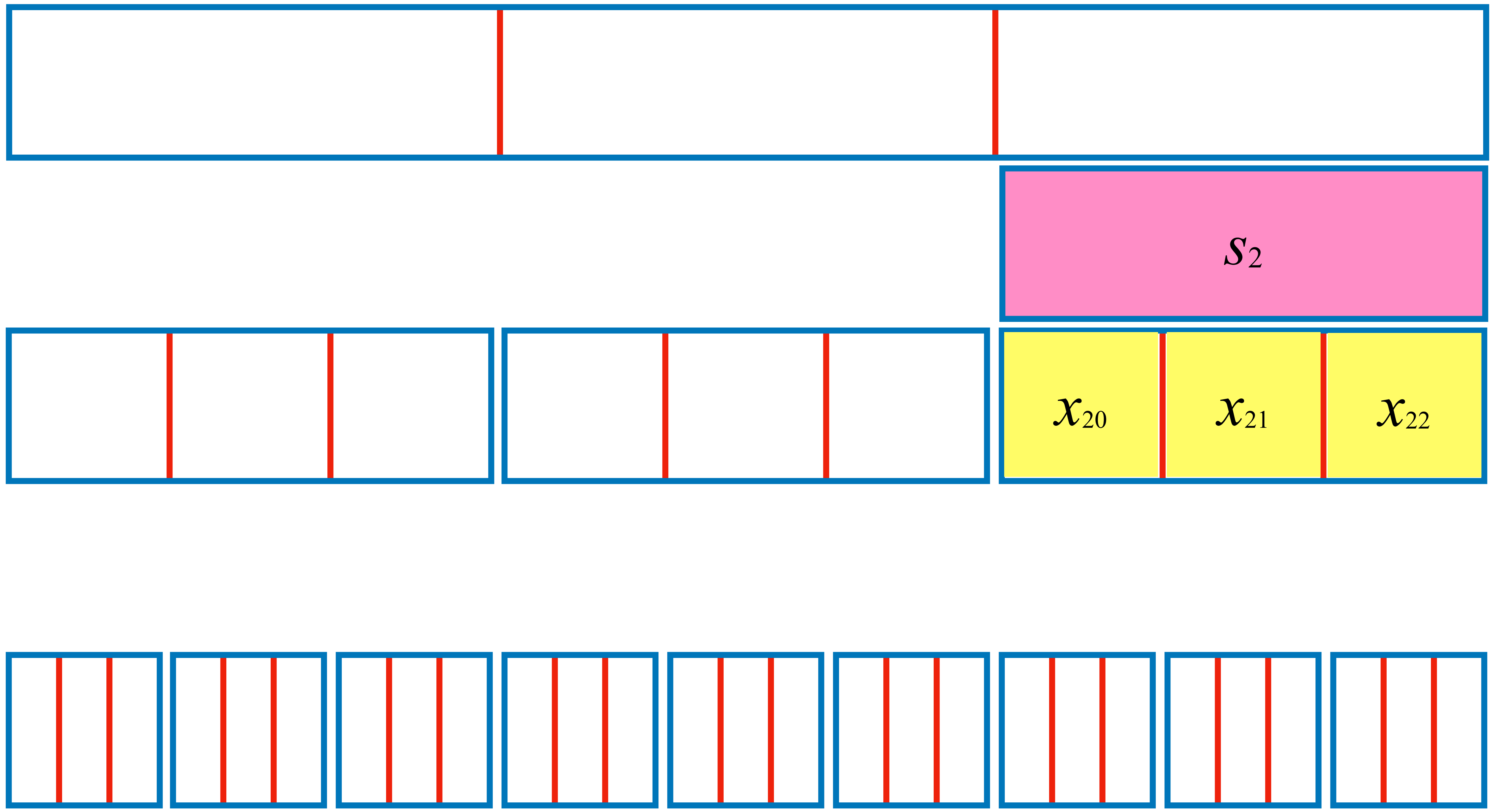
Swap Regret ≤ Swap Reward - Reward

Tree-Swap Swap Regret



Swap Regret ≤ Swap Reward - Reward

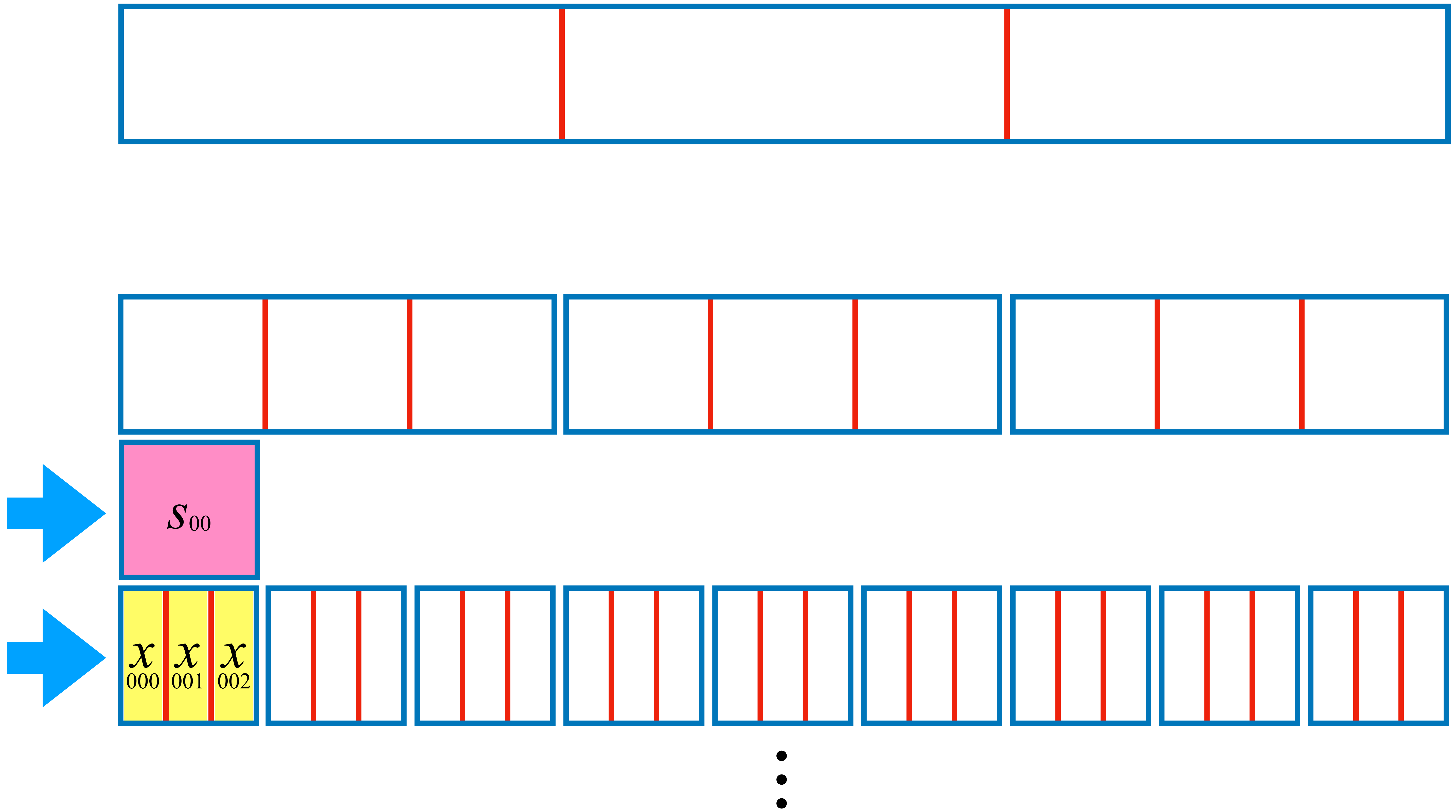
Tree-Swap Swap Regret



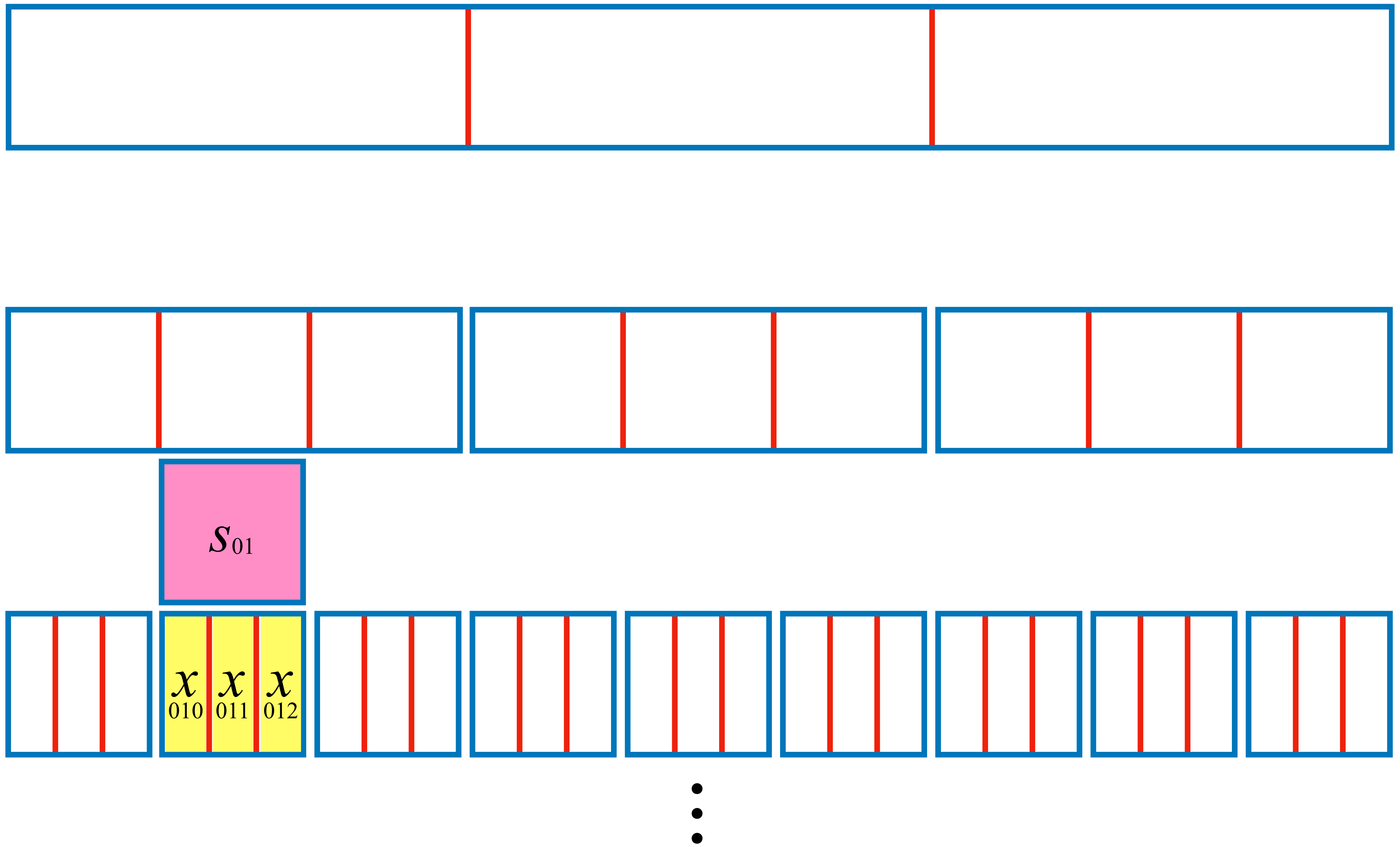
Swap Regret ≤ Swap Reward - Reward

Tree-Swap Swap Regret

Swap Regret ≤ Swap Reward - Reward

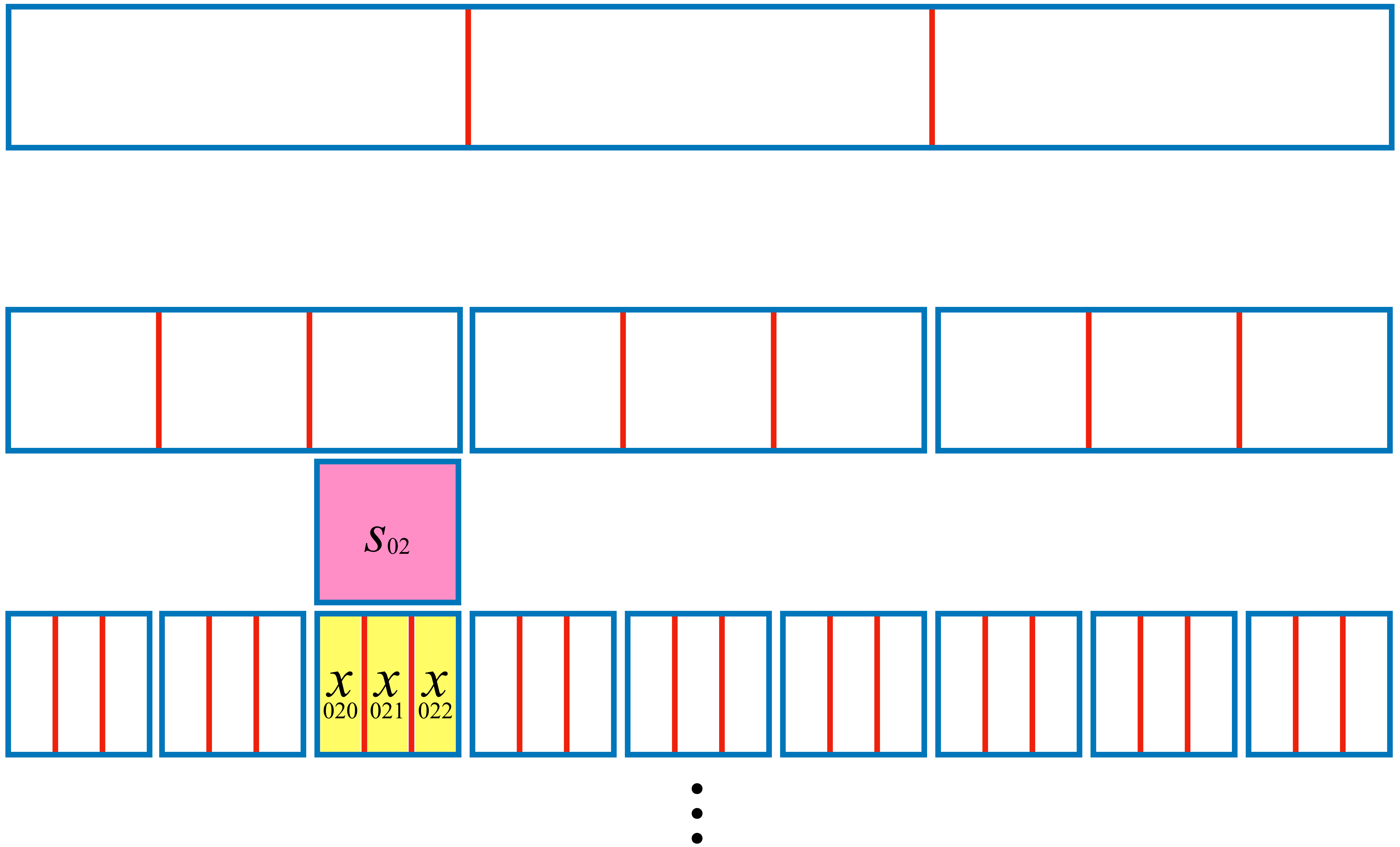


Tree-Swap Swap Regret



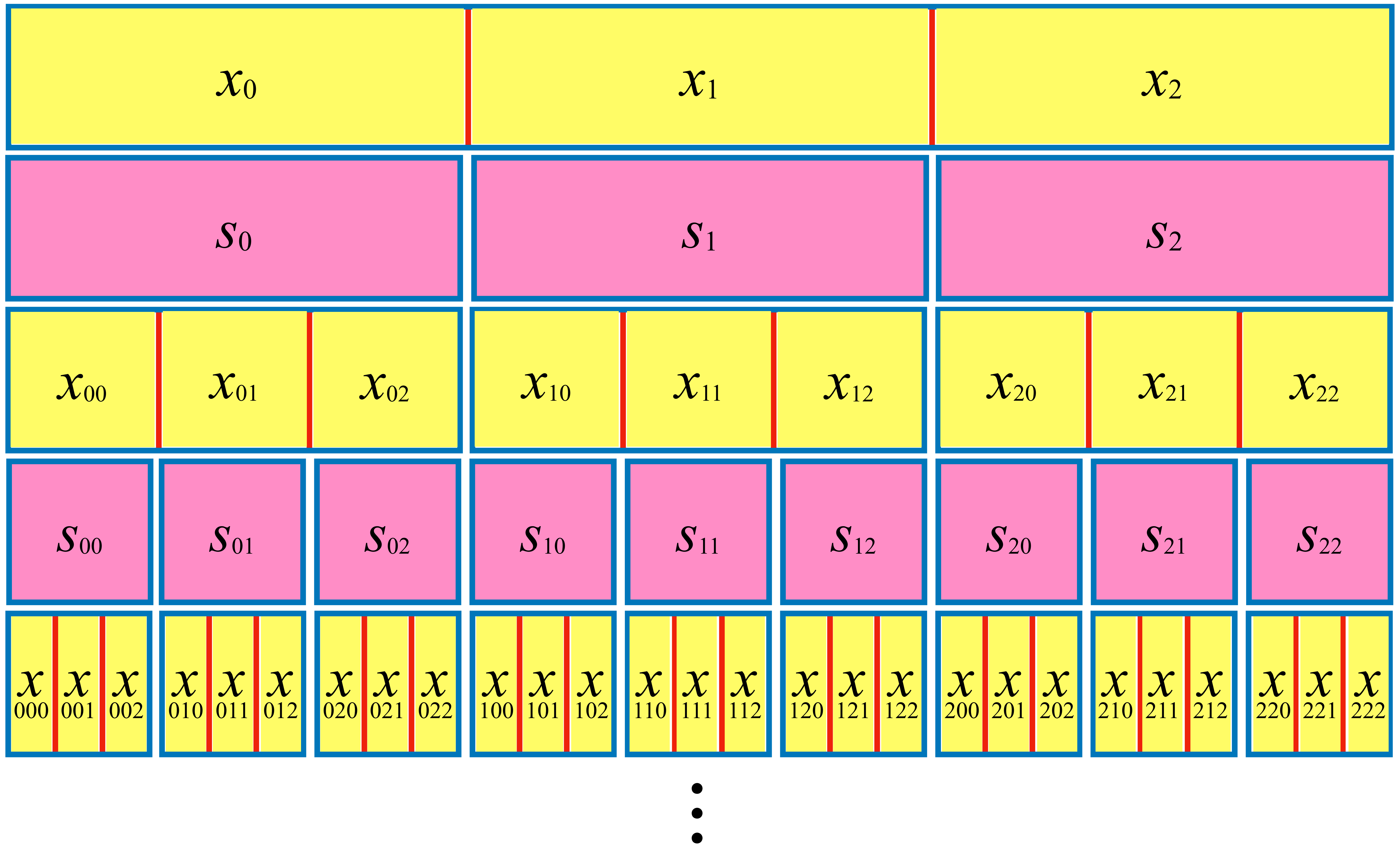
Swap Regret ≤ Swap Reward - Reward

Tree-Swap Swap Regret



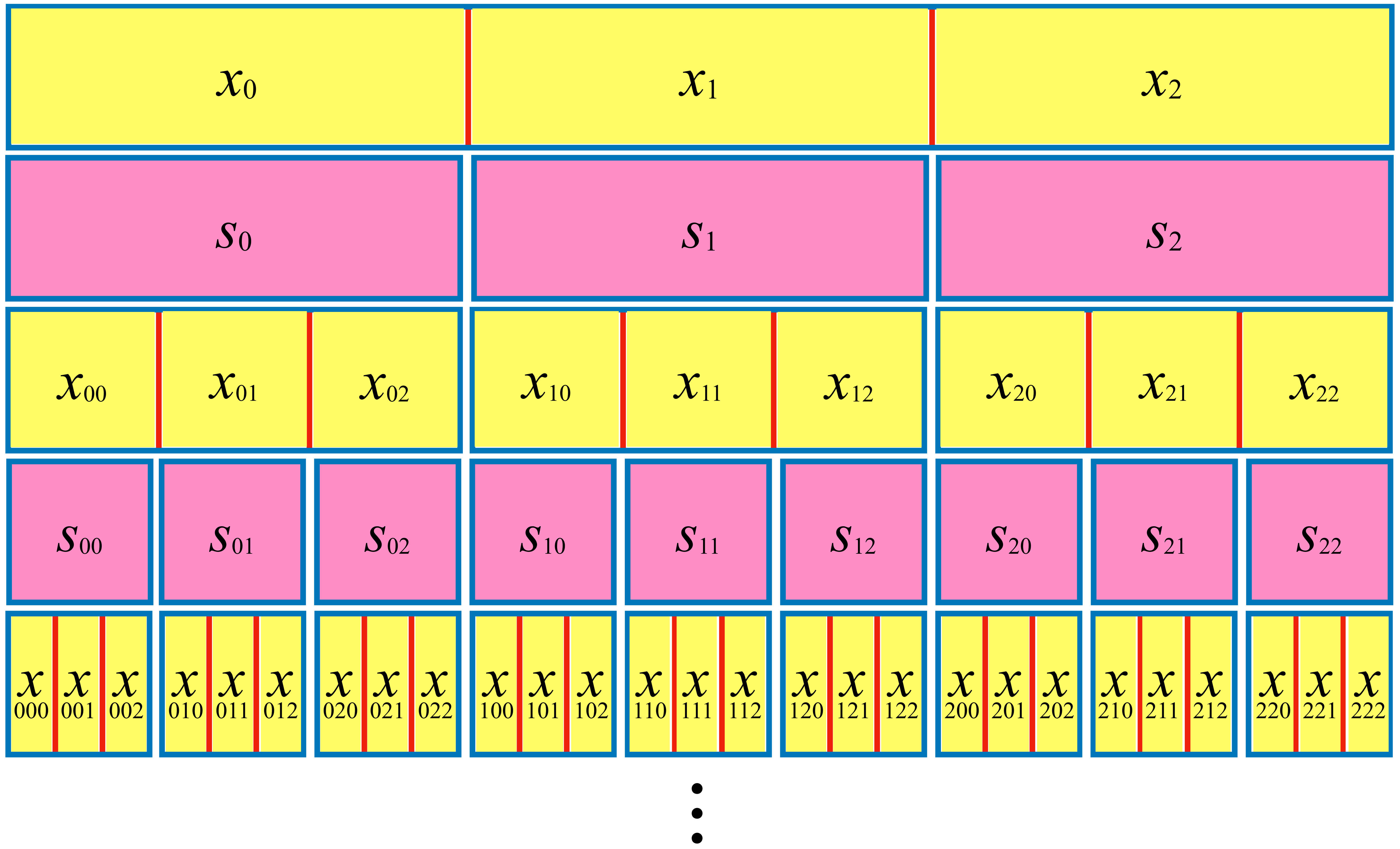
Swap Regret ≤ Swap Reward - Reward

Tree-Swap Swap Regret



Swap Regret ≤ Swap Reward - Reward

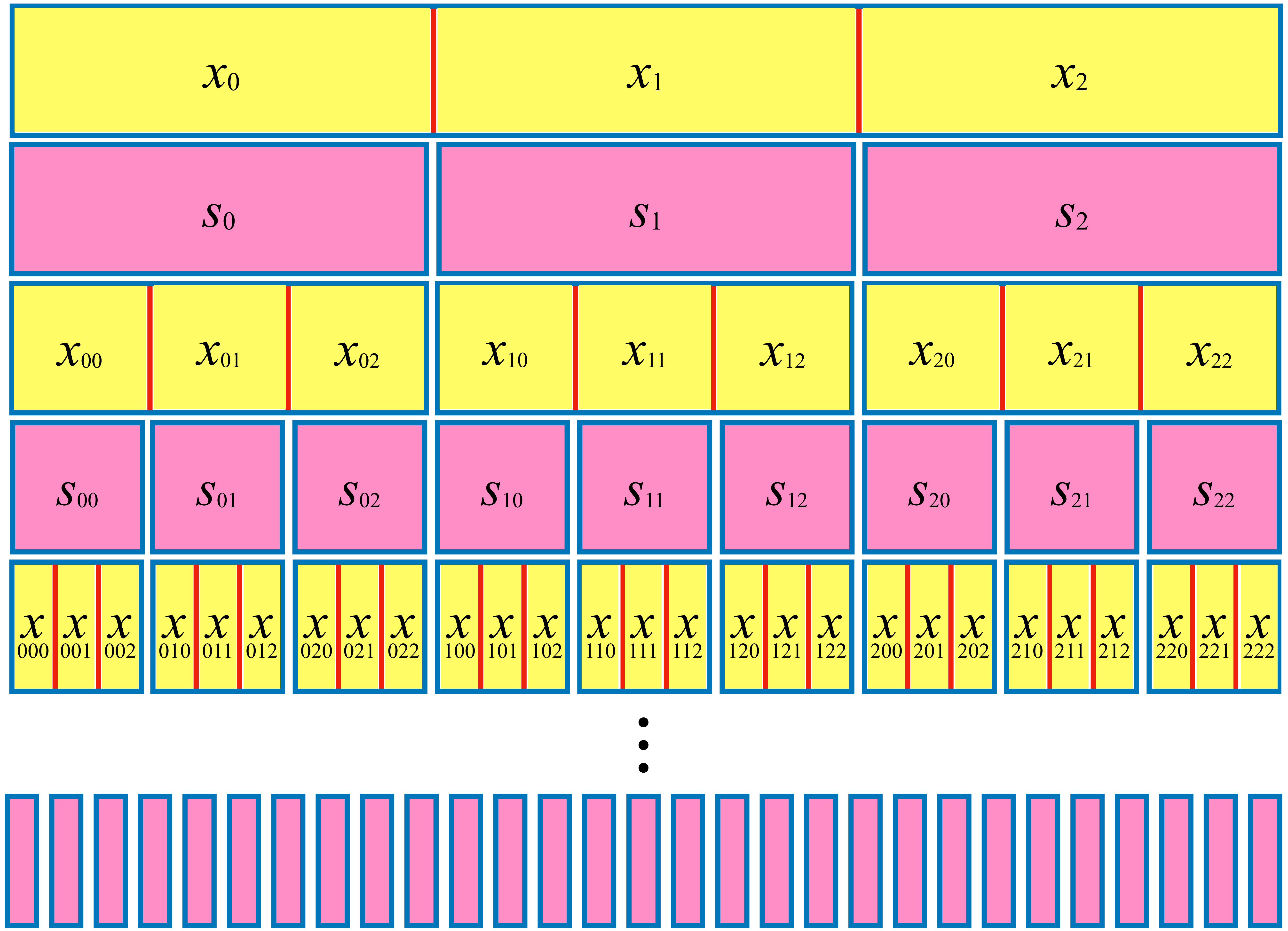
Tree-Swap Swap Regret



Swap Regret ≤ Swap Reward - Reward

$$= \sum_{\text{layer } 1}^{d-1} \text{External Regret}$$

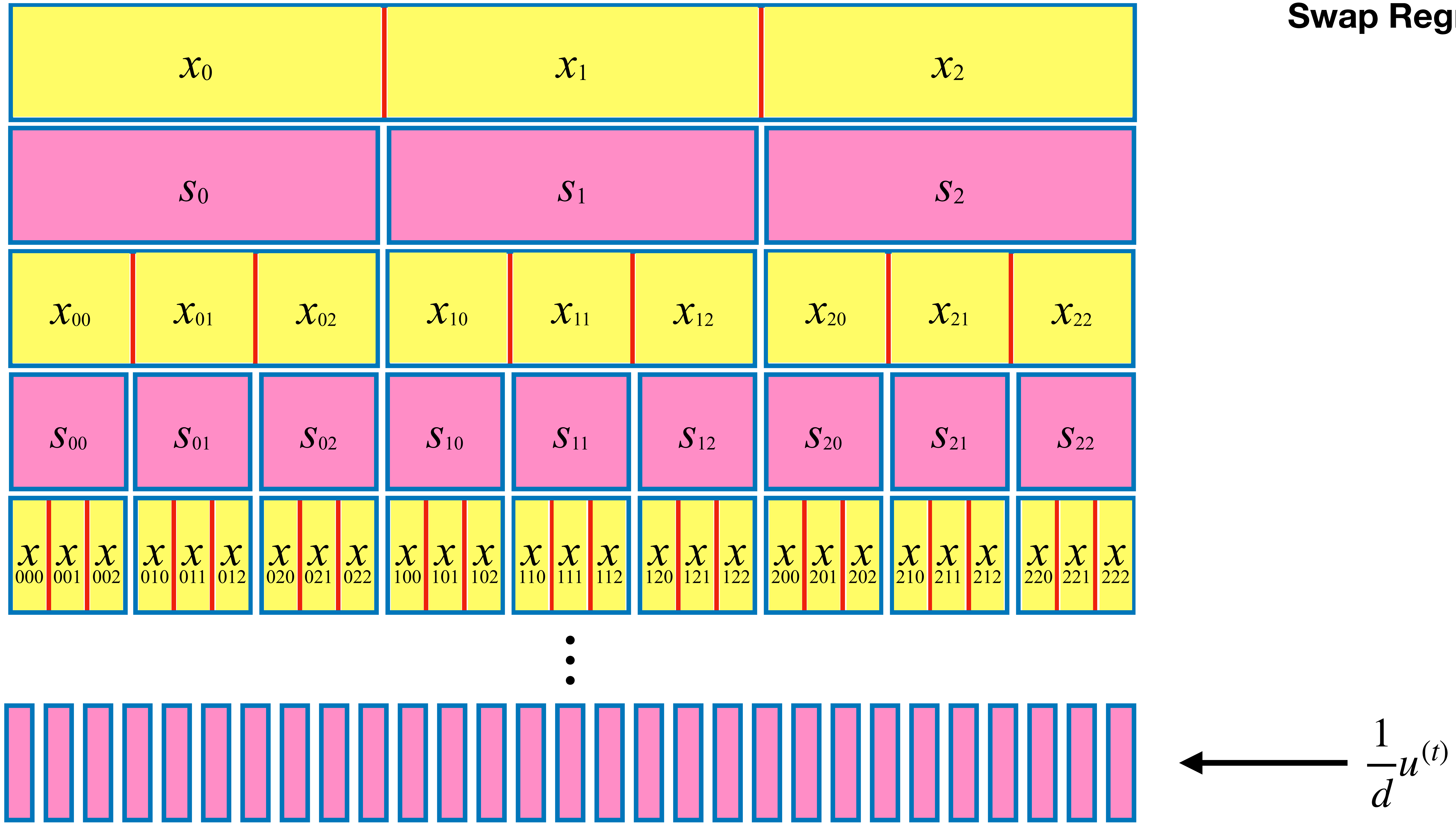
Tree-Swap Swap Regret



Swap Regret ≤ Swap Reward - Reward

$$= \sum_{\text{layer } 1}^{d-1} \text{External Regret}$$

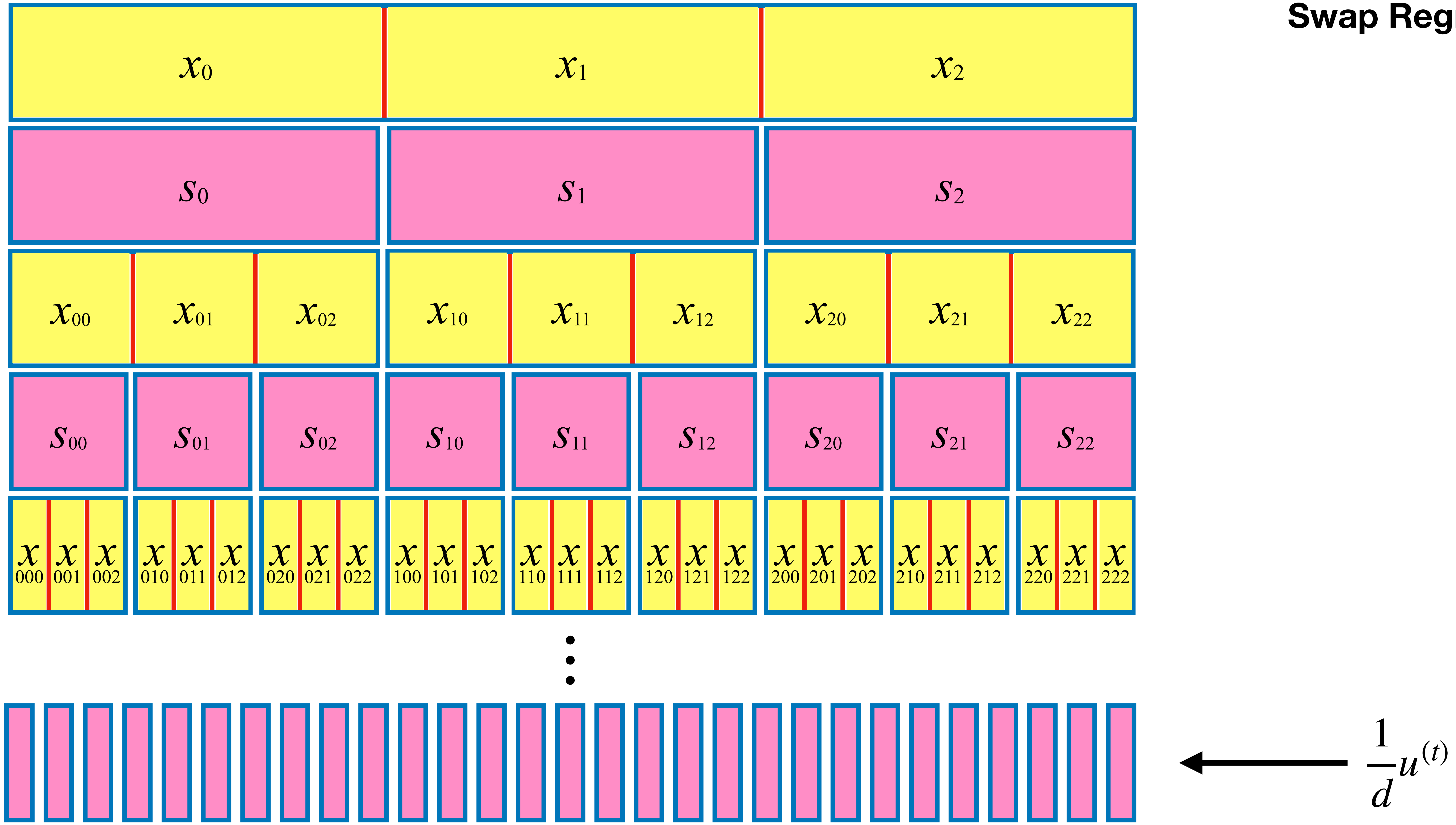
Tree-Swap Swap Regret



Swap Regret ≤ Swap Reward - Reward

$$= \sum_{\text{layer } 1}^{d-1} \text{External Regret}$$

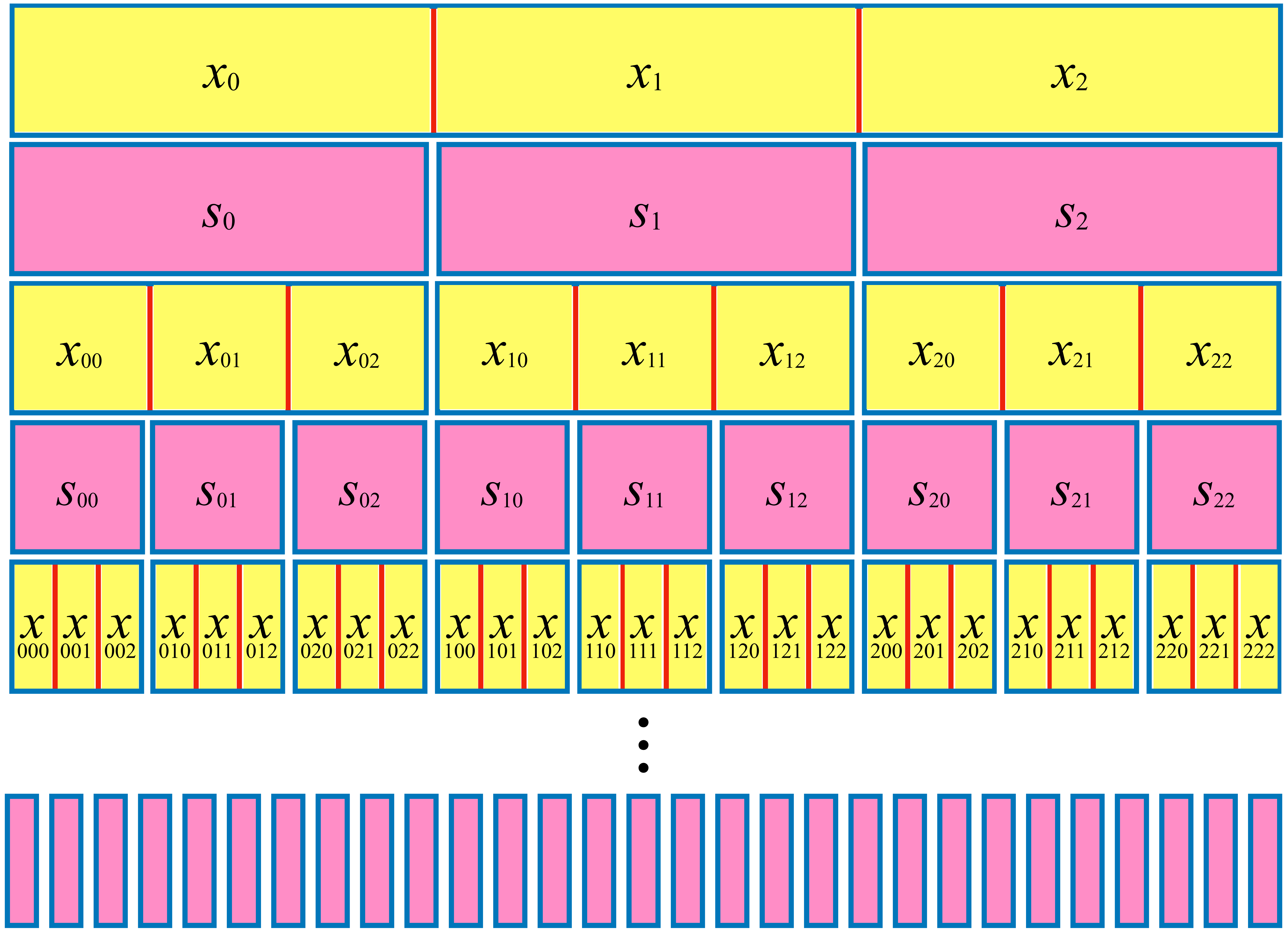
Tree-Swap Swap Regret



Swap Regret ≤ Swap Reward - Reward

$$= \sum_{\text{layer } 1}^{d-1} \text{External Regret} + \frac{T}{d}$$

Tree-Swap Swap Regret

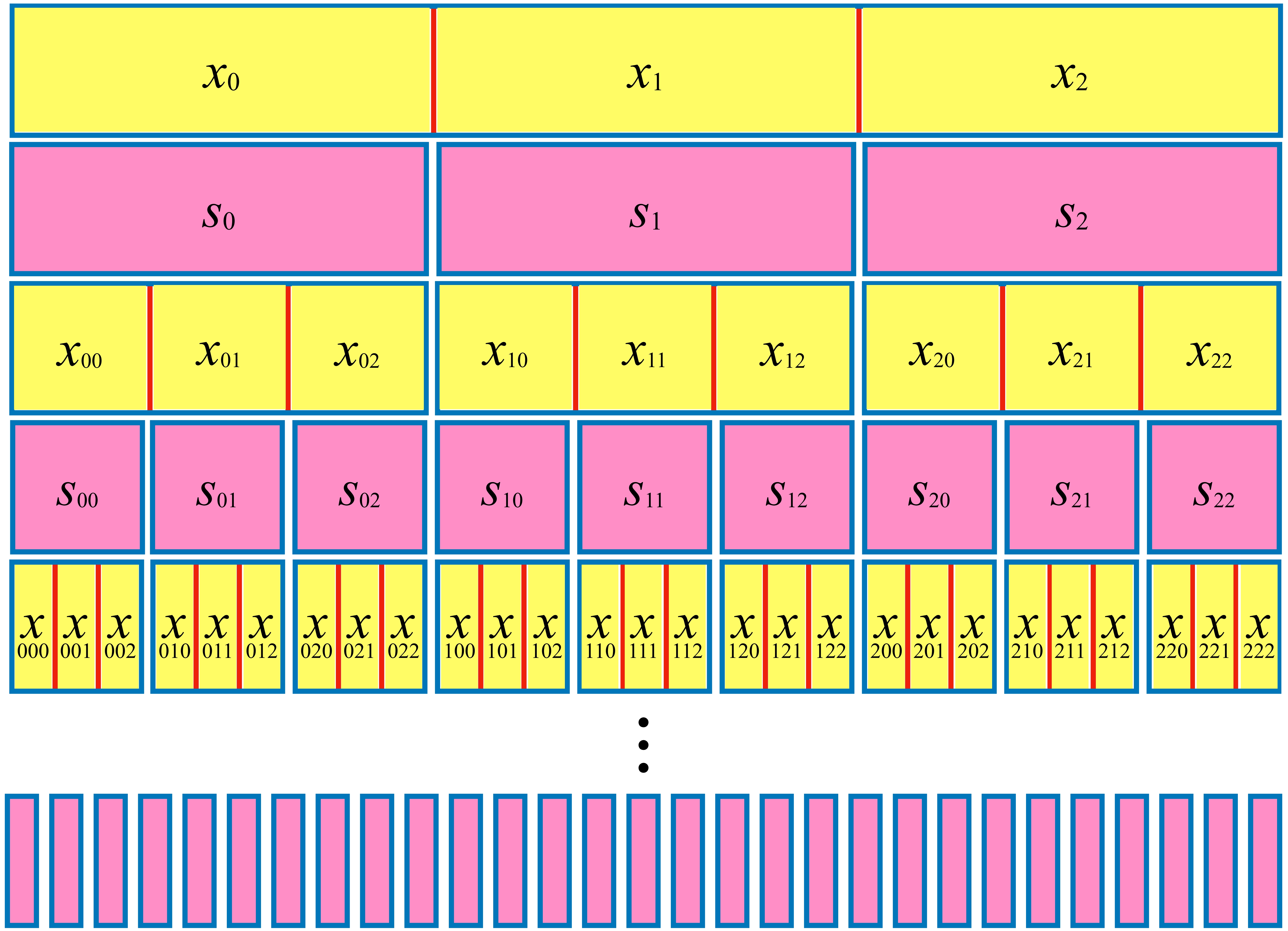


Swap Regret ≤ Swap Reward - Reward

$$= \sum_{\text{layer } 1}^{d-1} \text{External Regret} + \frac{T}{d}$$

Tree-Swap: Swap Regret $= \left(\sqrt{\frac{\log N}{M}} + \frac{1}{d}\right) T$

Tree-Swap Swap Regret



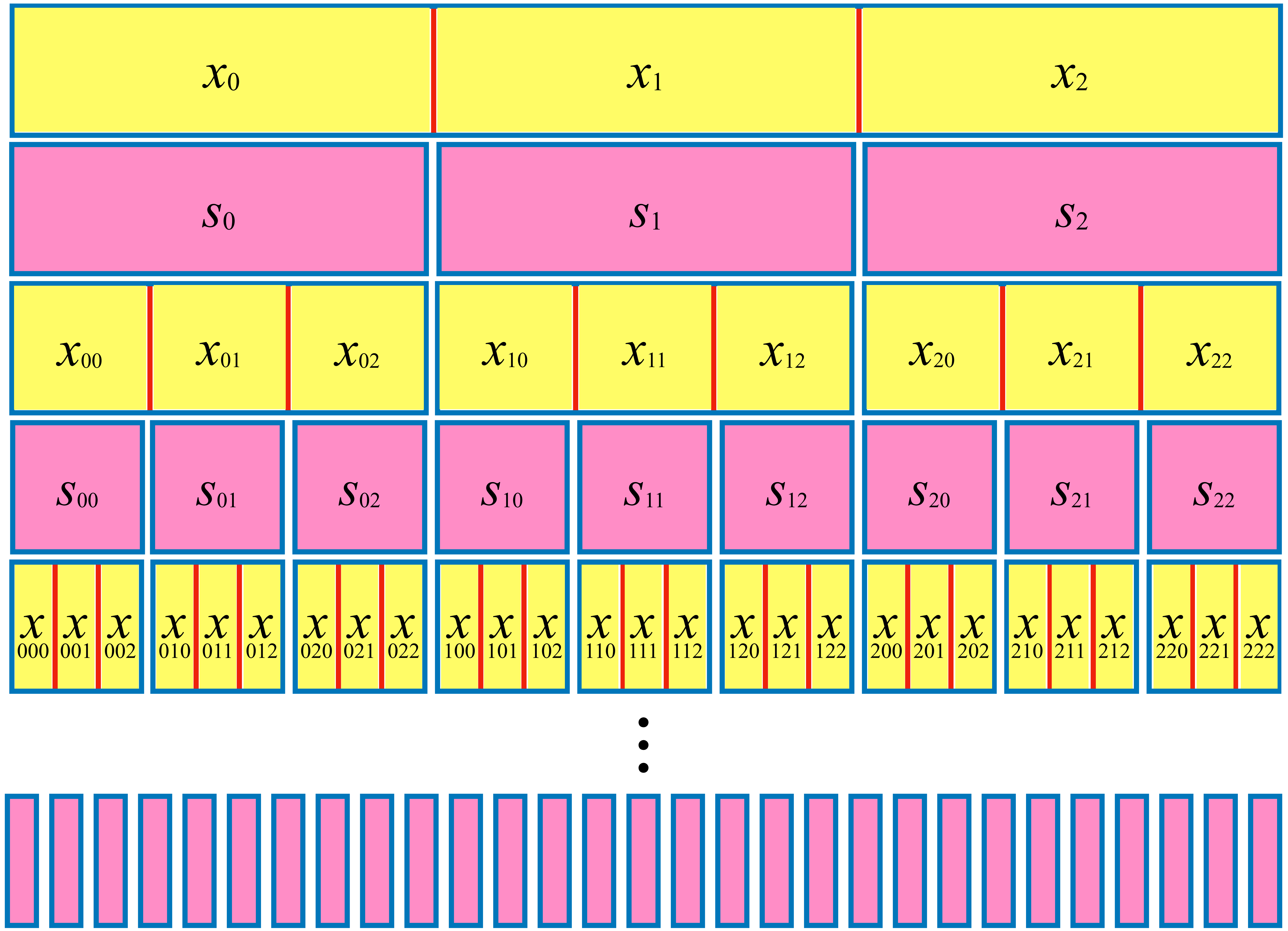
Swap Regret ≤ Swap Reward - Reward

$$= \sum_{\text{layer } 1}^{d-1} \text{External Regret} + \frac{T}{d}$$

Tree-Swap: Swap Regret $= \left(\sqrt{\frac{\log N}{M}} + \frac{1}{d}\right) T$

$$\boxed{M \geq \frac{\log N}{\epsilon^2}}$$

Tree-Swap Swap Regret



Swap Regret ≤ Swap Reward - Reward

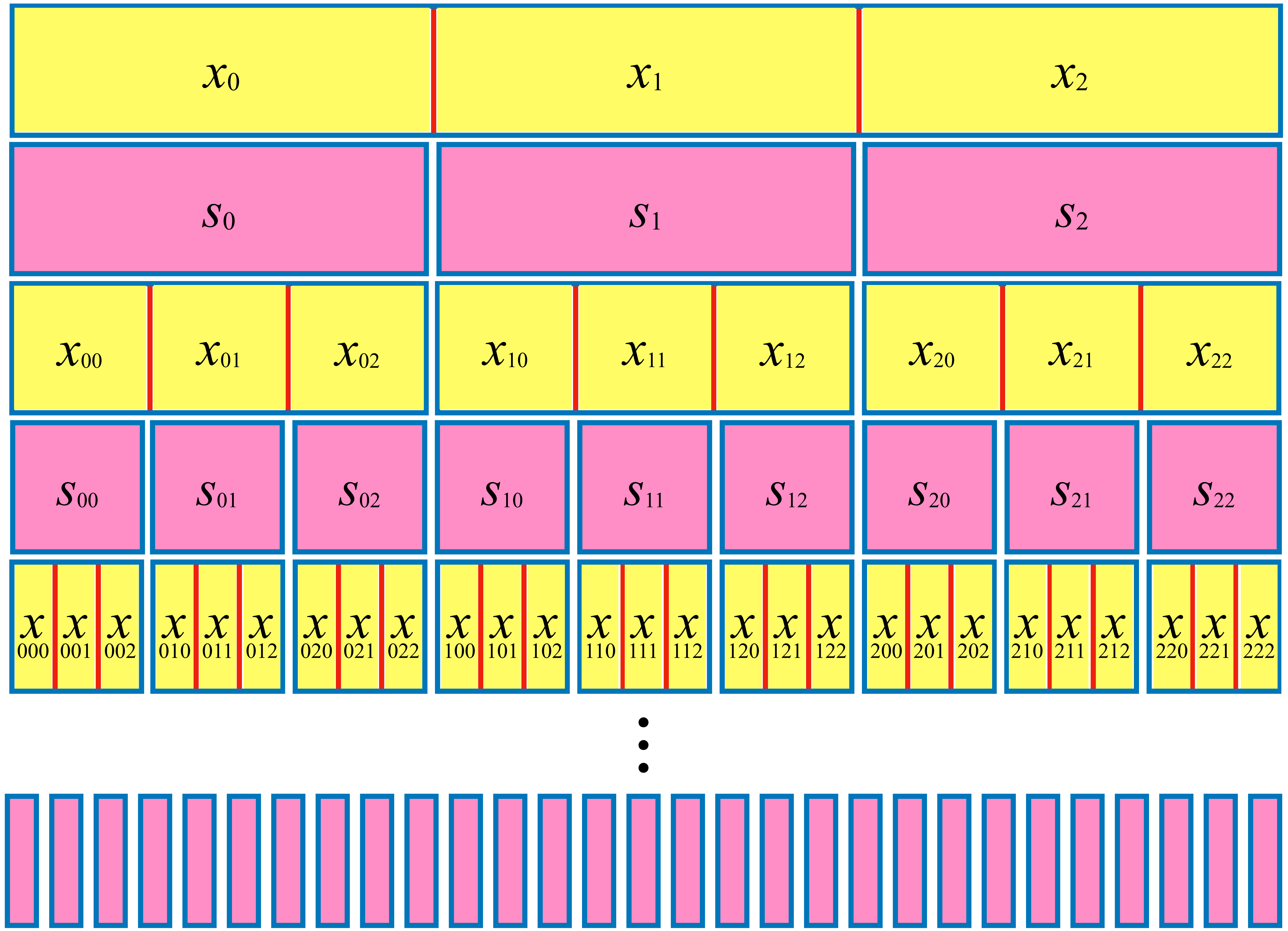
$$= \sum_{\text{layer } 1}^{d-1} \text{External Regret} + \frac{T}{d}$$

Tree-Swap: Swap Regret $= \left(\sqrt{\frac{\log N}{M}} + \frac{1}{d}\right) T$

$$M \geq \frac{\log N}{\epsilon^2}$$

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

Tree-Swap Swap Regret



Swap Regret ≤ Swap Reward - Reward

$$= \sum_{\text{layer } 1}^{d-1} \text{External Regret} + \frac{T}{d}$$

Tree-Swap: Swap Regret $= \left(\sqrt{\frac{\log N}{M}} + \frac{1}{d}\right) T$

$$M \geq \frac{\log N}{\epsilon^2} \qquad d \geq \frac{1}{\epsilon}$$

Tree-Swap: ϵ-Swap Regret

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

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**
[DSS19] **https://arxiv.org/pdf/1909.13861**