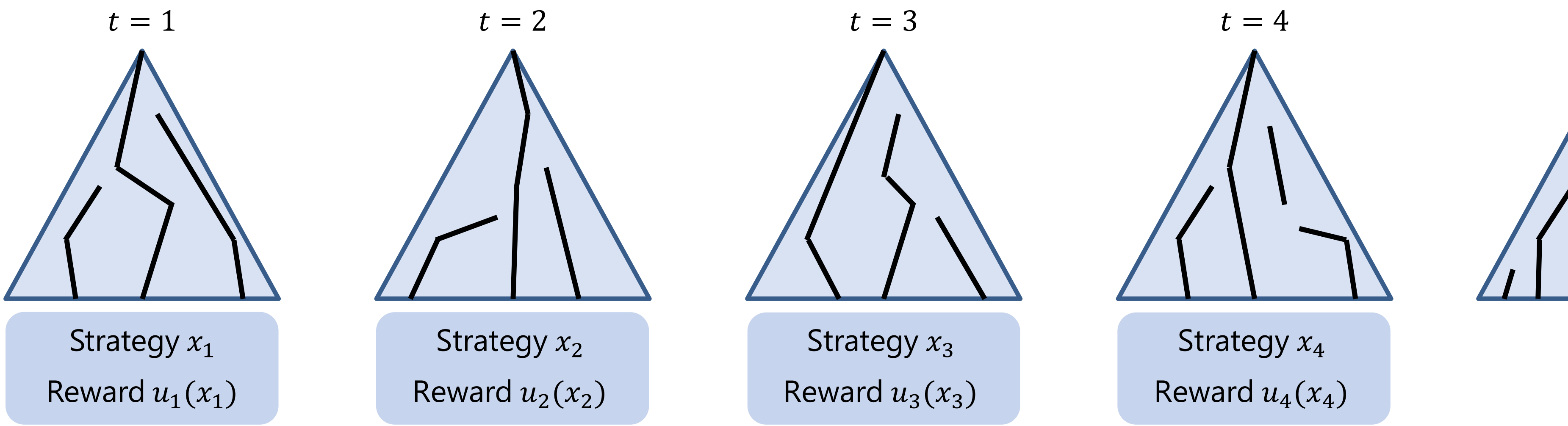
More regret minimization: Φ-Regret

What does it mean to learn?



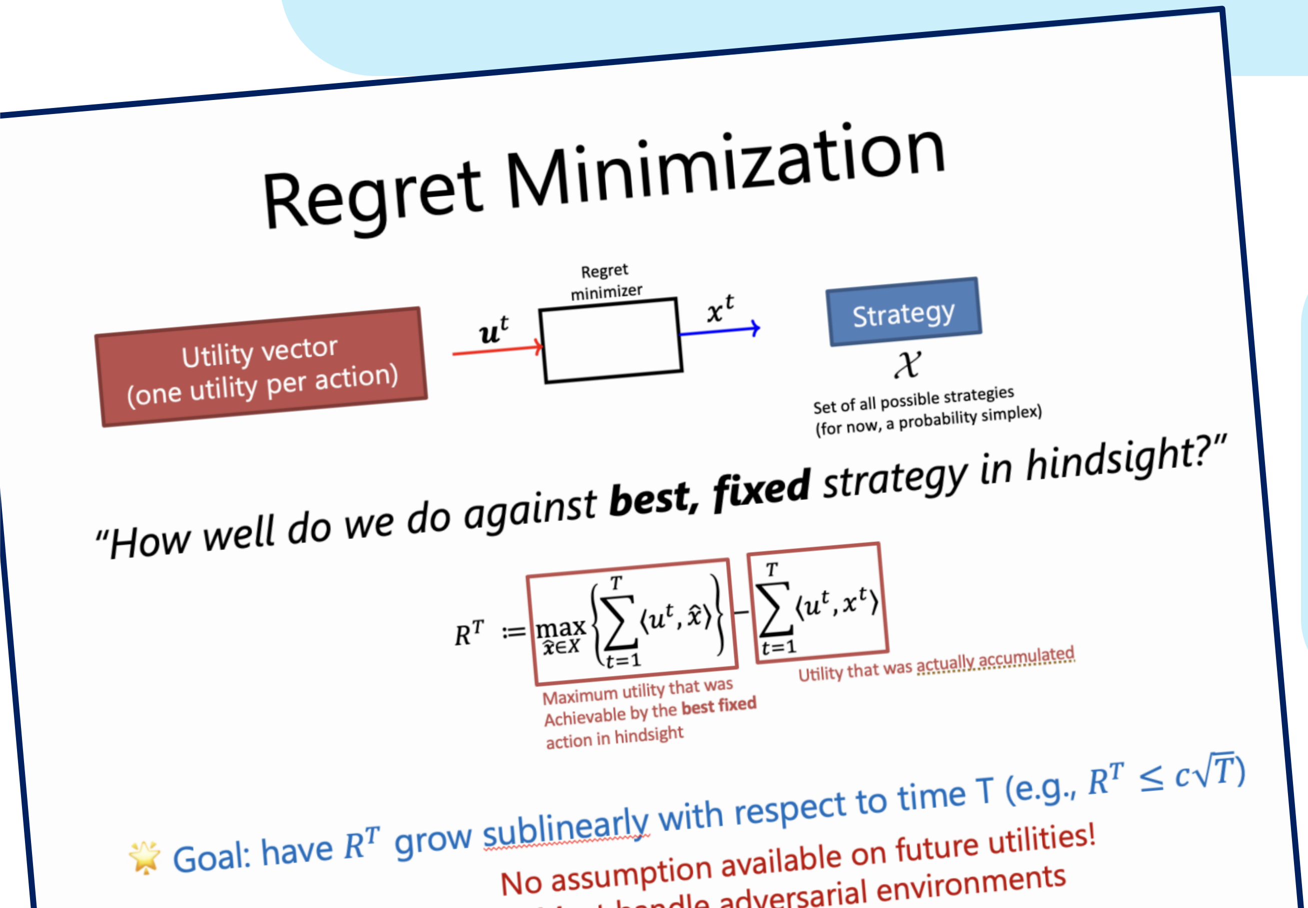
Idea: we have *learnt* when we do not strongly wish we had played differently (aka low **regret**)

In the past lecture, we always equated

“Playing differently in hindsight”

with

“Going back and using *one, fixed* strategy"



Isn’t that a weak notion of hindsight rationality?

A no-regret agent as defined for now can consistently observe that its average utility would have increased had it chosen action *a*′ **every time that it actually played *a***…

…And yet never switch to playing action *a*′ in these situations.

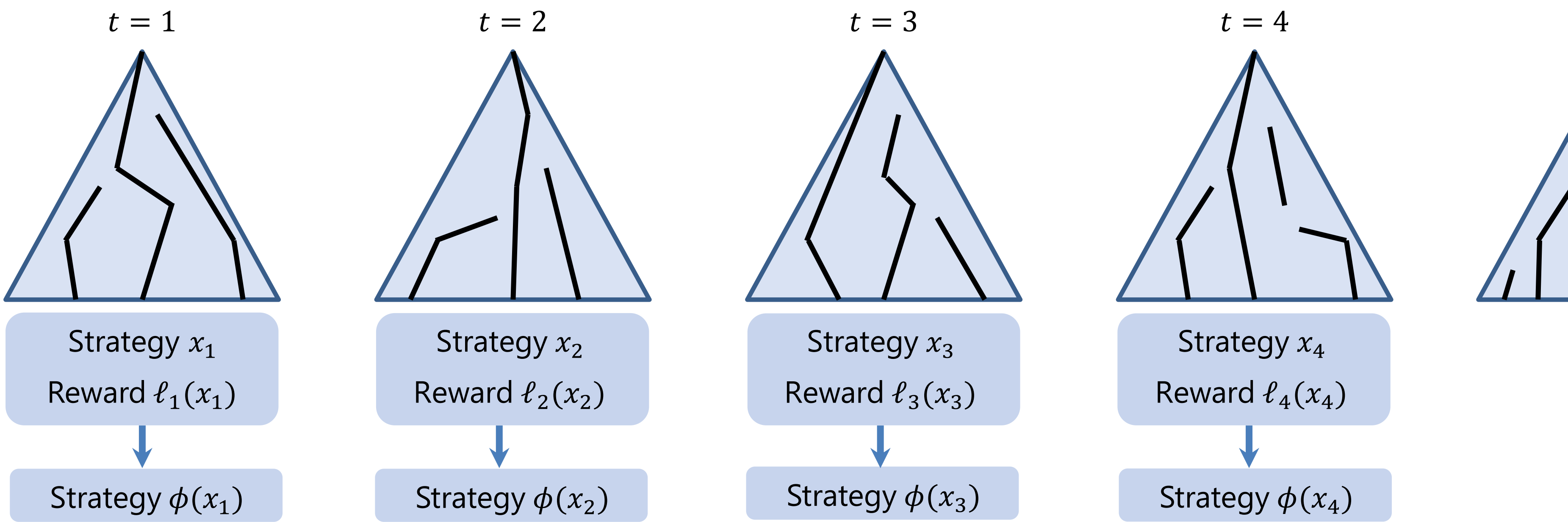
Φ-Regret

- Φ-regret is a generalization of the concept of regret as seen in the previous class

- It enables to specify what **transformations of strategies** the agent wants to learn to not regret

> Function ϕ from the strategy set to itself

What does it mean to learn?



$$\Phi\text{Regret}_T \coloneqq \max_{\phi \in \Phi} \sum_{t=1}^{T} \Big(u_t(\phi(x_t)) - u_t(x_t) \Big)$$

What does it mean to learn?

$$\Phi\text{Regret}_T \coloneqq \max_{\phi \in \Phi} \sum_{t=1}^{T} \Big(u_t\big(\phi(x_t)\big) - u_t(x_t) \Big)$$

The larger the set of transformations Φ we don't want to regret, the more rational the learner is

What sets Φ make the most sense?

Is there a connection with game-theoretic notions of rationality and equilibria?

Let's start from the notion we have already seen last time

When Φ is the set of **constant transformations**, then ΦRegret_T is called **external** regret

Constant transformations $\phi(x)$ ignore the input strategy, and always map to the same strategy

This notion applies to agents acting on either an extensive-form or a normal-form strategy space

External regret minimization

When Φ is the set of **constant transformations**, then $\Phi\text{-Regret}_T$ is called **external** regret

$$\Phi\text{Regret}_T \coloneqq \max_{\phi \in \Phi} \sum_{t=1}^{T} \Big(u_t\big(\phi(x_t)\big) - u_t(x_t) \Big)$$

$$= \max_{\hat{x} \in X} \sum_{t=1}^{T} \big(u_t(\hat{x}) - u_t(x_t)\big)$$

This was the formula we used in the previous lecture!

Guarantees of external-regret minimization

Folklore result: if all players play in a way that guarantees sublinear external regret over time, then:

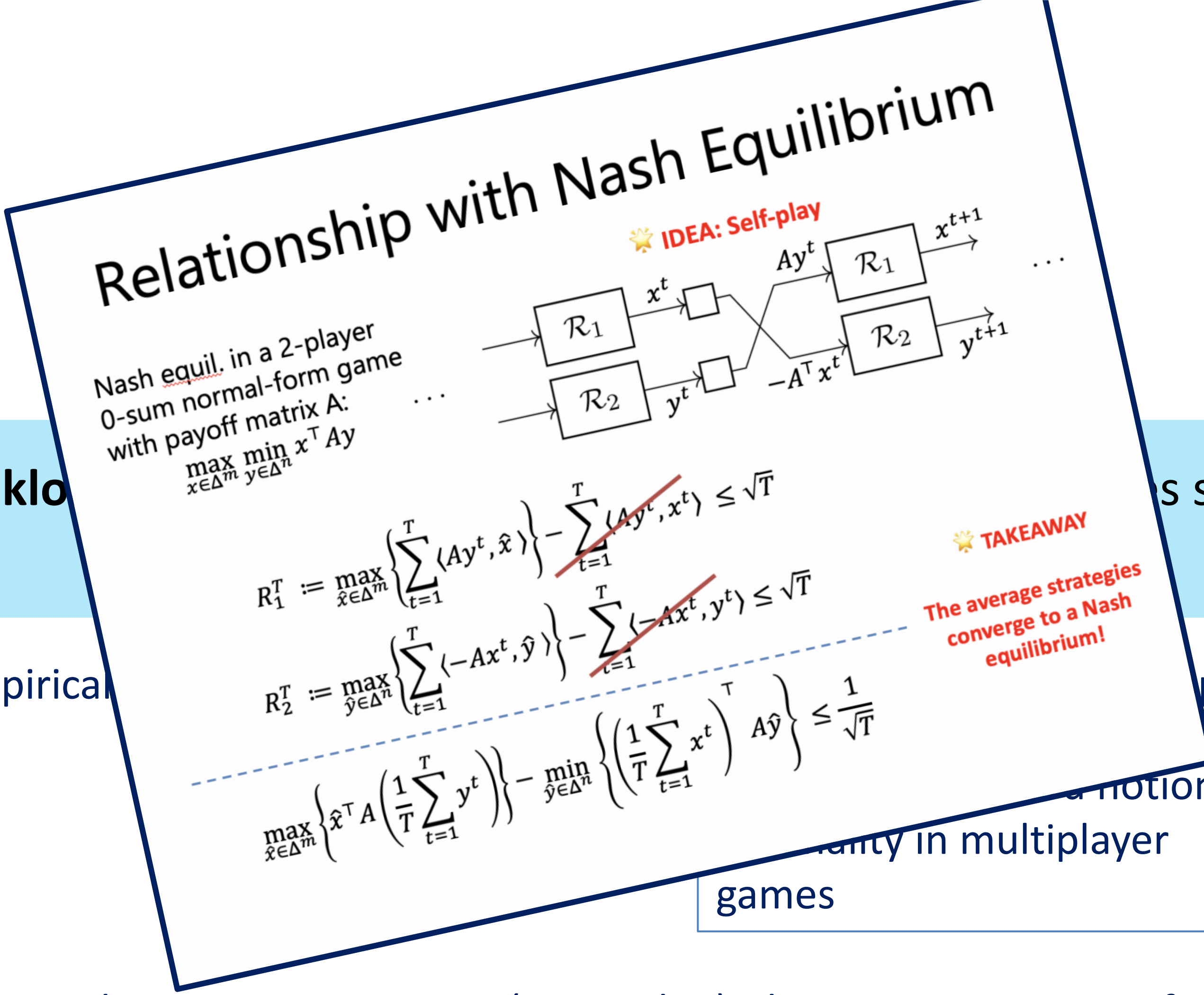
1. Empirical frequency of play is a ***coarse correlated equilibrium*** (multiplayer games)

One of the standard notions of rationality in multiplayer games

2. In *two-player zero-sum* games (e.g., heads-up poker), the average strategy of each player is a ***Nash equilibrium***

"Optimal" game-theoretic play, before exploiting specific weaknesses of the opponent. In poker, this is already superhuman

Folklo… …s sublinear



1. Empirical … player games)

… notions of … ality in multiplayer games

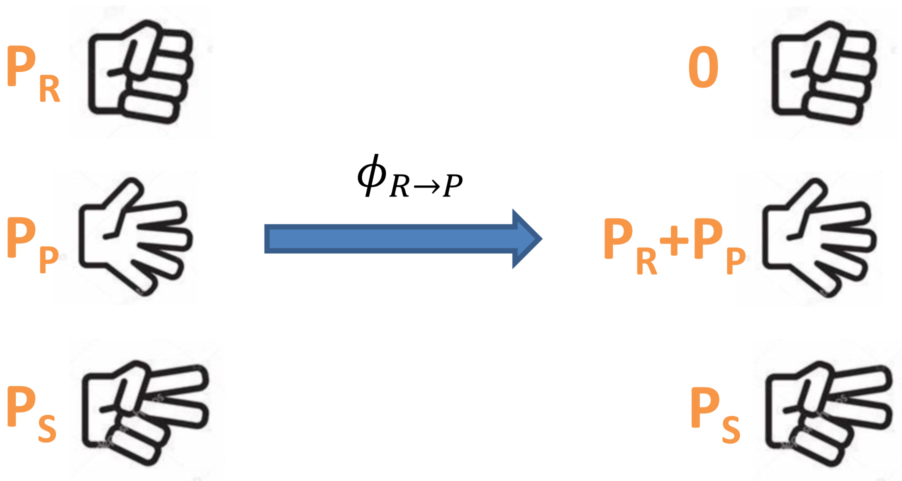
2. In *two-player zero-sum* games (e.g., poker), the average strategy of each player is a ***Nash equilibrium***

"Optimal" game-theoretic play, before exploiting specific weaknesses of the opponent. In poker, this is already superhuman

Internal regret minimization

Normal-form strategy spaces only

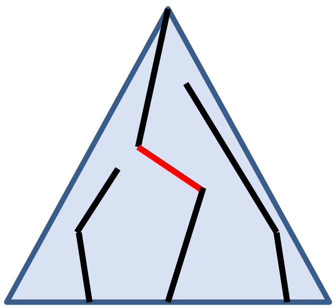
Internal regret minimization: be robust to **“probability mass transferring”** transformations



“Every time I played rock I should have played paper”

Trigger regret

Generalization of internal regret to extensive-form games



Trigger regret: considers **conditional** strategy transformations:

"If action A is played at decision point D, transform the strategy from D onwards by setting it equal to S"
(for all possible A, D, and S)

Many other notions of hindsight rationality exist for extensive form games
[Morrill, D'Orazio, Lanctot, Wright, Bowling, Greenwald; Efficient Deviation Types and Learning for Hindsight Rationality in Extensive-Form Games, ICML 2021]

[Celli, Marchesi, Farina, Gatti; No-Regret Learning Dynamics for Extensive-Form Correlated Equilibrium, NeurIPS 2020]
[Farina, Celli, Marchesi, Gatti; Simple Uncoupled No-Regret Learning Dynamics for Extensive-Form Correlated Equilibrium, J.ACM 2022]

Swap regret minimization

Normal-form strategy spaces only

Swap regret minimization: be robust to any linear transformation from probabilities to probabilities

P_R

P_P

P_S

$$\phi = \begin{pmatrix} 0.5 & 0.2 & 0.9 \\ 0.3 & 0.4 & 0.0 \\ 0.2 & 0.4 & 0.1 \end{pmatrix}$$

$0.5P_R + 0.2P_P + 0.9P_S$

$0.3P_R + 0.4P_P + 0.0P_S$

$0.2P_R + 0.4P_P + 0.1P_S$

Swap regret minimization

Swap regret subsumes both internal and external regret minimization in terms of "power" (that is, hindsight rationality)

Consequences: internal regret $\leq$ swap regret
external regret $\leq$ swap regret

Lemma: swap regret $\leq$ num actions $\times$ internal regret

$\Rightarrow$ Minimizing swap regret and internal regret are "equivalent" problems (up to game-dependent factors)

Cool fact: we have learning algorithms that guarantee sublinear swap regret

(In fact, we will see how one can construct one later)

What do internal and swap regret achieve?

- **Recall**: Internal regret cares about transformations of the form "every time I did X I should have done Y instead"

- So, the following might not be surprising:

No-internal and no-swap regret agents in self play converge to ***correlated*** equilibria in general-sum multi-player normal-form games

In extensive-form games, trigger regret leads to ***extensive-form correlated equilibrium***

Zooming out

One can try to make a pretty strong case that no-external-regret as notion of rationality is very weak and does not deserve so much attention…

Instead, as it turns out, external regret minimization is so important that most people refer to it as just "**regret minimization**"

Why is that?

First answer: it leads to Nash equilibrium in two-player zero-sum games. So, perhaps it is less weak than it appears

Second answer: it is easy to define for any strategy set, and we know of very powerful algorithms that minimize external regret for any convex set

Third answer: In general, it is possible to reduce general Φ regret to external regret minimization, provided the set of transformations considered satisfies certain *general properties*

1. We have a regret minimizer for the **set of transformations** Φ

2. Given any transformation $\phi \in \Phi$, we have an oracle to compute a fixed-point strategy of ϕ

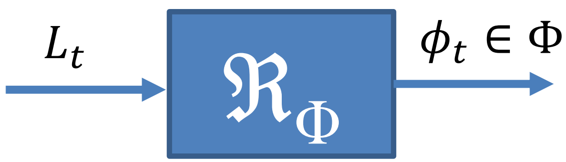
1. We have a regret minimizer for the **set of transformations** Φ
2. Given any transformation $\phi \in \Phi$, we have an oracle to compute a fixed-point strategy of ϕ

We ***want***:

Φ-regret minimizer for strategy set X

$$\max_{\hat{\phi} \in \Phi} \sum_{t=1}^{T} \left(u_t(\hat{\phi}(x_t)) - u_t(x_t)\right) = o(T)$$

We ***have***:



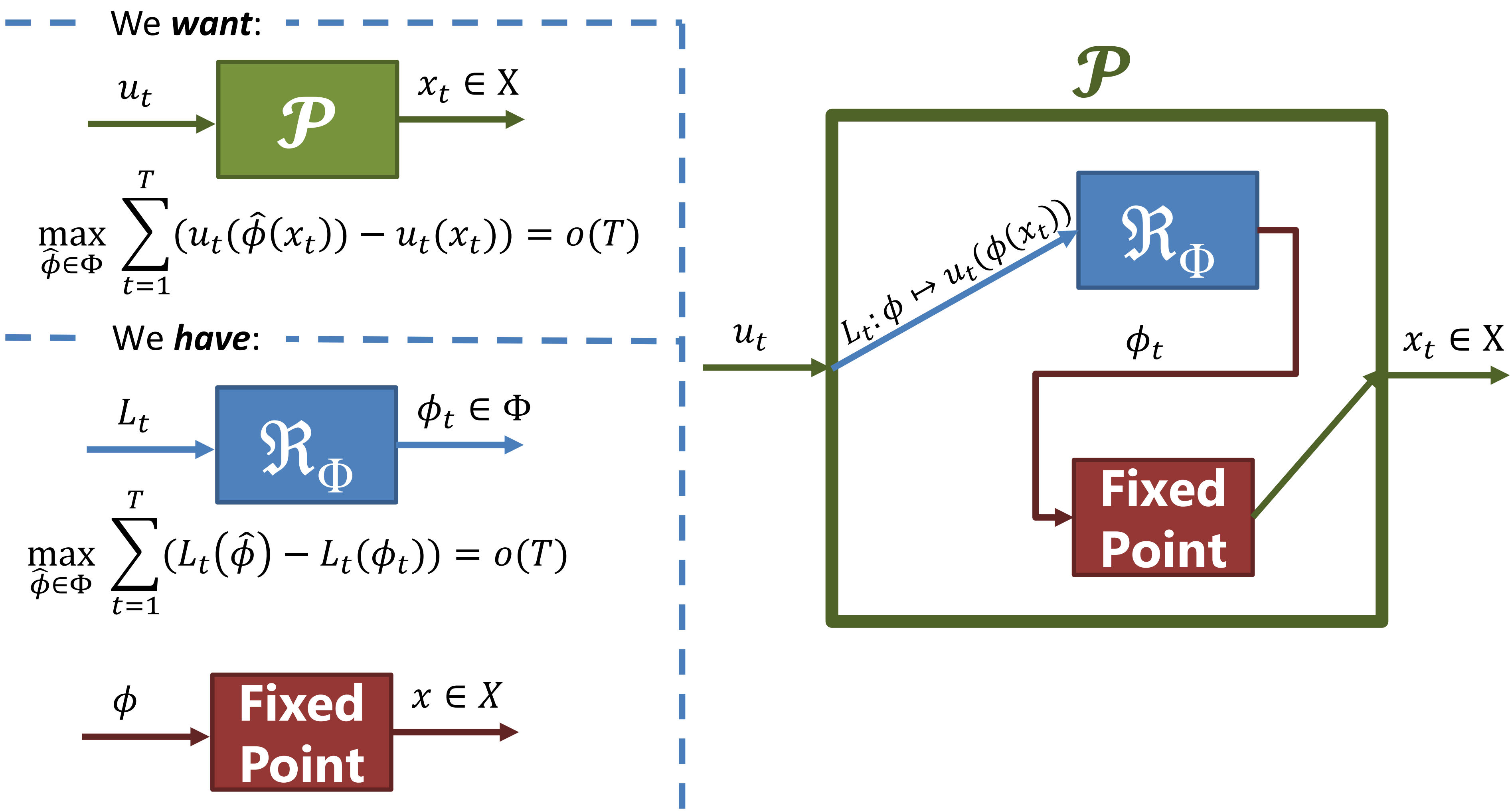
External regret minimizer for the set of transformations Φ

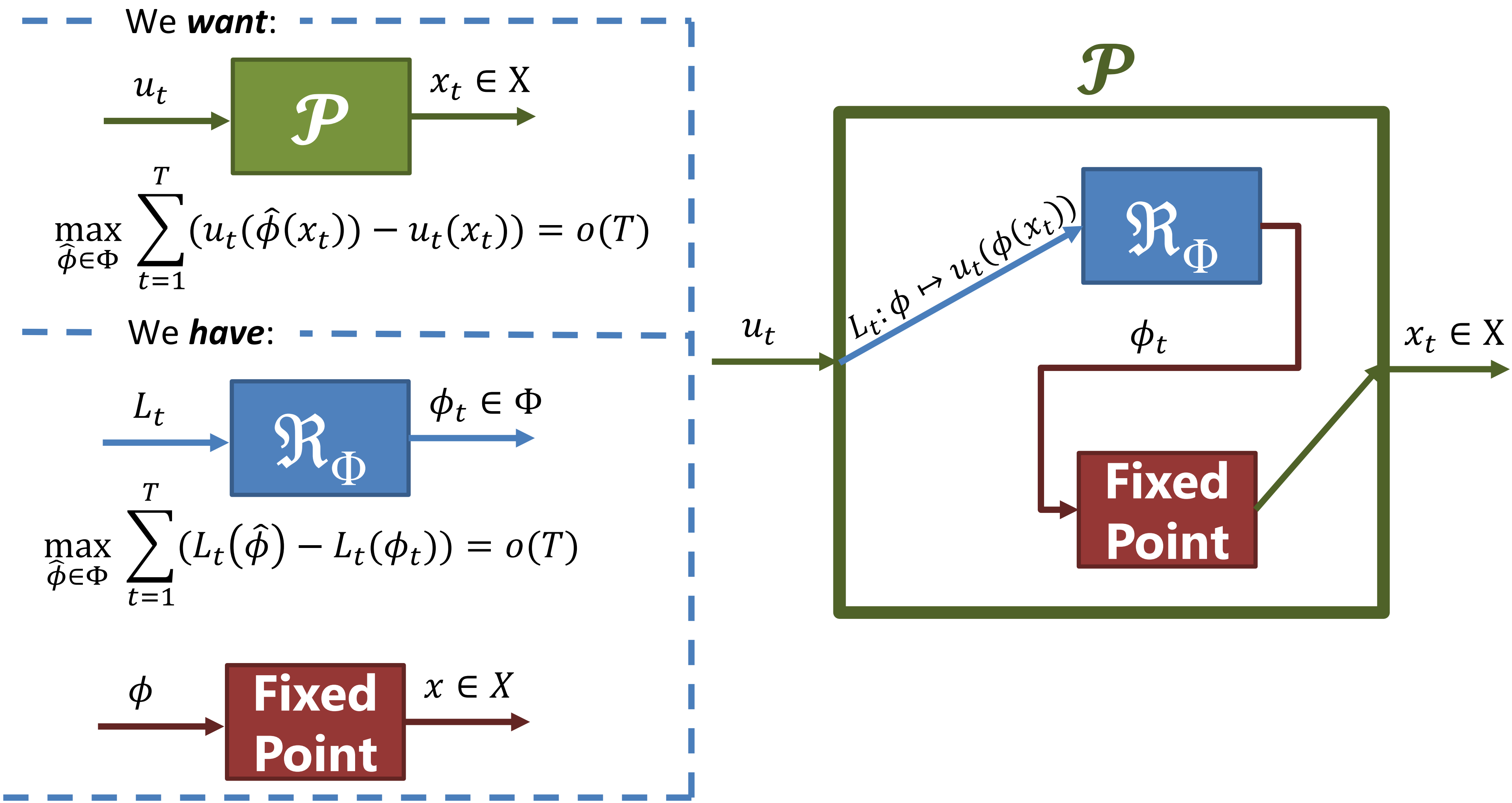
$$\max_{\hat{\phi} \in \Phi} \sum_{t=1}^{T} \left(L_t(\hat{\phi}) - L_t(\phi_t)\right) = o(T)$$

Fixed-point oracle

$\phi(x) = x \in X$

1. One can construct a regret minimizer for the **set of transformations** Φ
2. Given any transformation $\phi \in \Phi$, we have an oracle to compute a fixed-point strategy of ϕ





Proof:

$$o(T) = \max_{\hat{\phi}\in\Phi} \sum_{t=1}^{T} (L_t(\hat{\phi}) - L_t(\phi_t)) = \max_{\hat{\phi}\in\Phi} \sum_{t=1}^{T} (u_t(\hat{\phi}(x_t)) - u_t(\phi_t(x_t)))$$

$$= \max_{\hat{\phi}\in\Phi} \sum_{t=1}^{T} u_t(\hat{\phi}(x_t)) - u_t(x_t))$$

Summary of this lecture

- The notion of regret can be generalized beyond external regret to Φ-regret
- While external regret leads to Nash equilibrium (in two-player zero-sum settings) and coarse correlated equilibria, more powerful Φ-regret notions lead to correlated equilibria in general-sum multi-player games
- In many cases, Φ-regret minimizers can be constructed from external regret minimizers

Summary of this lecture

- The notion of regret can be generalized beyond external regret to Φ-regret
- While external regret leads to Nash equilibrium (in two-player zero-sum settings) and coarse correlated equilibria, more powerful Φ-regret notions lead to correlated equilibria in general-sum multi-player games
- In many cases, Φ-regret minimizers can be constructed from external regret minimizers

Summary of this lecture

- The notion of regret can be generalized beyond external regret to Φ-regret
- While external regret leads to Nash equilibrium (in two-player zero-sum settings) and coarse correlated equilibria, more powerful Φ-regret notions lead to correlated equilibria in general-sum multi-player games
- In many cases, Φ-regret minimizers can be constructed from external regret minimizers