Games Embedded in Institutions

Game Theory · Module 10 — Strategic Systems Integration

Node 10.1

Position in Knowledge Graph

System: Game Theory **Structural Domain:** Strategic Systems Integration **Node:** 10.1

Up to this point in the course, a "game" has been an object specified in isolation: a triple (N, S, u), or an extensive form, or a mechanism. But no actual strategic situation is ever specified in isolation. Every real game sits inside an **institution** — a set of rules, norms, legal constraints, enforcement mechanisms, and custom — that determines which game is being played at all. The institution specifies who the players are, what actions are legally available, what the payoffs look like after taxes and transfers, which strategies are prohibited, and what happens when someone breaks a rule. In this sense, institutions are a **meta-game** whose outcome is the selection of an object-level game. 10.1 opens Module 10 by making this observation precise: the formalism of games within games, the institutional rules as a higher-order game-selector, and the way institutions both constrain and create strategic possibilities. Everything in Module 10 — layered mechanisms (10.2), contracts (10.3), commitment devices (10.4), collective choice (10.5), strategic ecosystems (10.6) — depends on this reframing.

Formal Definition

An **institution** I is a tuple

$$I = (\mathcal{G}, R, E, \phi)$$

where:

- $\mathcal{G}$ is a set of candidate object-level games over a common player set N. Each $G \in \mathcal{G}$ is a strategic-form or extensive-form game as defined in Modules 1–3.
- R is a set of **institutional rules**: constraints on player actions, admissibility conditions, and payoff transformations (taxes, rake, fines, transfers).
- E is an **enforcement mechanism** specifying which deviations from R are detectable, which are punished, and with what cost.
- $\phi : R \times E \to \mathcal{G}$ is the **game-selection function**: it takes the rules and the enforcement regime and returns the effective object-level game that the players face.

Players move in two phases:

1. **Institutional phase (meta-game).** A prior process — legislation, custom formation, casino licensing, civil-law development, market-platform design — selects R and E. This is itself a game (possibly a collective-choice game of the form studied in 10.5).
2. **Object-level phase.** Given $\phi(R, E) = G$, the players play G as an ordinary game.

The effective payoff function for the object-level game is

$$u_i^{\text{eff}}(s) = u_i^{\text{raw}}(s) - t_i(s) - \pi_i(s) \cdot \nVdash[s \notin R_i]$$

where $t_i(s)$ is the institutional transfer (taxes, fees, rake), R_i is the rule-admissible action set for player i, and $\pi_i(s)$ is the punishment if player i violates rules and is caught by E.

A **well-defined institution** is one where ϕ is deterministic and the enforcement mechanism E is credible (in the sense of 10.4): the prescribed punishments are actually carried out when violations occur. An **institutional equilibrium** is a pair (s^*, R^*) where s^* is a Nash equilibrium of $\phi(R^*, E)$ and R^* is stable under the meta-game process that generates rules.

Intuition

No one plays "poker" in the abstract. You play poker-at-the-Bellagio, or poker-on-PokerStars, or poker-at-a-home-game. These are structurally different games even though the card-ranking rules are identical, because the **institutions** are different: rake structures, player pools, enforcement of collusion norms, legal status of winnings, tax treatment, availability of real-time assistance, tournament structures, table selection rules. Each institution selects a different effective game.

The same is true of every other strategic situation. You do not play "oligopoly" — you play oligopoly-under-EU-competition-law, or oligopoly-under-US-antitrust, or oligopoly-in-an-unregulated-grey-market. You do not play "the dating game" — you play dating-in-1950s-America, or dating-on-Tinder, or dating-in-rural-Japan. The institutional context is not flavor. It determines the action set, the payoff function, the information structure, the credibility of threats, and the identity of the players.

The structural point is that **treating the institution as fixed is a modeling choice, not a fact**. For short-run questions (how should I play this hand?) the institution is rightly taken as given. For medium-run questions (should I play at the Bellagio or on PokerStars?) the institution is a strategic choice made by the player. For long-run questions (should this kind of poker be legal, taxed, licensed?) the institution is the object of strategic design. Module 10 is the module where the boundary of "the game" becomes a strategic variable.

Three observations make this precise:

1. **Institutions create games.** Without the Bellagio's rake structure, dealer, security, and licensing, the game "poker at the Bellagio" would not exist. The institution is not a modifier on a preexisting game; it is the condition of possibility for the game.
2. **Institutions are equilibria of higher-order games.** Rules do not fall from the sky. They are selected by a collective-choice process (legislation, custom, market competition among platforms) that is itself a game. This nesting is potentially unbounded: the rules of the legislative process are themselves an institution, etc.
3. **Institutional rules transform payoffs, strategies, and players.** A tax changes payoffs. A licensing rule changes who can play. An enforcement mechanism changes which strategies are credible. A disclosure rule changes the information structure. Every dimension of the game triple (N, S, u) is institutionally contingent.

Structural Role

10.1 is the opening reframing of Module 10. Every subsequent node in the module is a specialization of the institutional embedding:

- **10.2 Layered mechanisms** treats institutions as nested mechanism-design problems: the object-level mechanism is embedded in a higher-level regulatory mechanism.
- **10.3 Contracts and principal-agent** treats contracts as private institutions that selectively constrain actions and transform payoffs for a bilateral relationship.

- **10.4 Commitment and credibility** asks when institutional rules are self-enforcing — when the meta-game actually produces the intended object-level game.
- **10.5 Political economy and collective choice** opens the black box of the institutional-selection meta-game and analyzes it with its own tools (voting, bargaining, constitutional design).
- **10.6 Strategic ecosystems** treats the whole stack — institutions, object-level games, players, strategies — as a coevolving dynamical system.

Looking backward, 10.1 is where many threads from earlier modules are retied. Repeated-game folk theorems (Module 5) gave us a large set of possible equilibria; institutions are what select among them. Mechanism design (Module 6) gave us tools to shape outcomes by choosing rules; institutions are the *context* in which a mechanism designer operates. Evolutionary dynamics (Module 7) gave us adjustment processes; institutional change is one such process operating on a longer timescale. Bayesian games (Module 4) gave us type-contingent strategies; institutional rules often work precisely by constraining how types can be revealed or concealed.

Mathematical Development

Institutional transformation of games

Let $G^{\text{raw}} = (N, S^{\text{raw}}, u^{\text{raw}})$ be a "bare" game before any institutional overlay. An institution $I = (R, E, \phi)$ produces a transformed game $G^{\text{eff}} = \phi(R, E)$ by three operations:

1. **Strategy restriction.** $S_i^{\text{eff}} = S_i^{\text{raw}} \cap R_i$. Institutional rules delete some strategies from the action set (prohibition: collusion in auctions, insider trading, use of solvers during live hands).

2. **Payoff transformation.** $u_i^{\text{eff}}(s) = u_i^{\text{raw}}(s) - t_i(s)$. Taxes, fees, rake, transfers are subtracted; subsidies are added.

3. **Enforcement overlay.** For strategies that violate rules but are not physically prevented, an enforcement term is added: $u_i^{\text{eff}}(s) = u_i^{\text{raw}}(s) - t_i(s) - p_i(s)\pi_i(s)$ where $p_i(s)$ is the detection probability and $\pi_i(s)$ is the punishment if caught.

The institutional transformation is generally **not a bijection on game space**: many raw games can map to the same effective game (informational indifference at the object level), and one raw game can map to many effective games (institutional heterogeneity). The transformation is also not invertible from observing play: you cannot in general recover u^{raw} from observing u^{eff} unless you separately observe the institution.

The nested-game formalism

Let $\mathcal{I}$ be a set of candidate institutions and $\mu : \mathcal{I} \to \Delta(\mathcal{G})$ be a **meta-game strategy** for the institutional phase: given a state of the meta-game, what distribution over institutions is the output? An **institutional equilibrium** is a fixed point

$$(I^*, s^*) \quad \text{with} \quad I^* \in \arg\max_{\text{meta}} \sum_i u_i^{\text{meta}}(I, s^*(I)), \qquad s^*(I) \in \text{NE}(\phi(I))$$

where the meta-game players (legislators, platform designers, custom-makers) anticipate the object-level equilibrium $s^*(I)$ when choosing their institution I. This is a two-level fixed-point problem: the institution is chosen taking the object-level equilibrium as given, and the object-level equilibrium is computed taking the institution as given. Solving both levels simultaneously is the conceptual heart of institutional analysis.

Institutional stability: self-enforcement

An institution is **self-enforcing** if, once in place, no coalition of players has both the incentive and the ability to change it. Formally, let $C \subseteq N$ be a coalition and u_C^{meta} be their meta-game payoff. I^* is self-enforcing if

$$\forall C \subseteq N, \quad u_C^{\text{meta}}(I^*) \geq \max_{I' \in \text{reach}(I^*, C)} u_C^{\text{meta}}(I')$$

where $\text{reach}(I^*, C)$ is the set of institutions that coalition C could unilaterally bring about given the current state. This condition ties 10.1 to the folk-theorem machinery of 5.4 and to the core of cooperative game theory: a self-enforcing institution is analogous to a core allocation in cooperative games.

The constitutional layer

Above any specific institution sits the **constitutional** layer that governs how institutions can be changed. This is a meta-meta-game. The sequence can in principle continue indefinitely, but in practice it terminates at a layer that is **evolutionarily stable** (7.4) rather than strictly self-enforcing: the constitutional layer is held in place not by incentive compatibility at its own level but by the cost of coordinated change. This is why constitutions are so rarely rewritten even when no one argues they are optimal.

Worked Example

Casino-rake-as-institution

Consider heads-up no-limit hold'em cash as the object game. Raw game: zero-sum two-player poker, payoff $u_i^{\text{raw}}(s)$ is expected winnings.

Institution: a Las Vegas casino with a 10% rake capped at \$5 per hand, a minimum buy-in, a maximum buy-in, a prohibition on real-time assistance (no solvers, no phones at the table), a prohibition on collusion, a dealer providing neutral administration, and an enforcement mechanism consisting of floor staff, surveillance cameras, and ejection / banning / prosecution for violations.

Effective game transformation:

1. **Strategy restriction.** S_i^{eff} excludes strategies involving phone consultation, explicit communication with other players in a multiway game, and marked cards. The action space is narrower.
2. **Payoff transformation.** $u_i^{\text{eff}}(s) = u_i^{\text{raw}}(s) - \text{rake}(s)$. The game is no longer zero-sum between the two players — the house extracts a third-party payoff. The effective game is **constant-negative-sum for the players**.
3. **Enforcement overlay.** Phone use during a hand is physically not prevented (you could pull out your phone), but the enforcement term $-p(s)\pi(s)$ where p is the detection probability (essentially 1, cameras and floor staff) and π is the punishment (ejection, loss of comp, possible forfeiture of winnings) is large. So the effective payoff is large-negative for phone-use strategies even though they are technically in S^{raw}.

Strategic consequences. The rake means that weak players lose faster than in a zero-sum game; the edge required for a winning player rises. The prohibition on real-time assistance means that the relevant notion of "strategy" is a human's memorized strategy, not a solver's dynamically computed strategy — the game is about human implementation, not pure game-theoretic optimality. The buy-in cap limits the effective stack size and the scope of deep-stack strategy. The collusion prohibition rules out multiway coordinated play that would be possible in an unregulated game.

Comparison with online poker. The same raw game under PokerStars' institution has a different rake structure (smaller per hand, percentage of pot up to a cap), different enforcement (software-detected collusion, but no floor staff for angle shooting), different strategy restrictions (HUDs allowed on some sites, banned on others), and different player-pool dynamics. The effective game is structurally distinct even though the card-ranking rules are the same. A player who is winning at PokerStars may be losing at the Bellagio or vice versa, not because they are "worse at poker" but because they are playing a different effective game.

Meta-game. The casino selects its rake structure, rule set, and enforcement regime to maximize its own payoff (casino profit) subject to the constraint that players must voluntarily participate. This is an institutional design problem — a mechanism-design problem in the sense of Module 6, but with the meta-game feature that the casino must anticipate the object-level equilibrium to know whether players will actually show up. Too much rake and the game dies; too little and the casino loses money. The observed rake structures across major card rooms are a revealed equilibrium of this meta-game.

Poker Anchor

Concept mapping

Formal object	Poker instantiation
Raw game G^{raw}	Abstract poker: card dealing, betting rounds, hand rankings, zero-sum transfer
Institution I	The venue: Bellagio, PokerStars, a home game, Macau high-stakes, a bar tournament
Rules R	Rake, buy-in caps, no phones, no collusion, tournament structure, seating
Enforcement E	Dealer, floor staff, surveillance, software detection, community norms
Transfer $t_i(s)$	The rake — casino's cut per hand
Effective game G^{eff}	The actual game that sits in front of you when you sit down
Meta-game	Casino licensing, legislation, platform competition, venue selection by players

Concrete situation: rake structure and winning threshold

At a \$2/\$5 NLH cash game with 10% rake capped at \$5, the house's per-hand extraction is significant in small-to-medium pots and negligible in large pots. A break-even player in the zero-sum abstract game loses at the rake rate; a small winner becomes a break-even player; a large winner becomes a small winner. The winning threshold — how good you must be to profit — is shifted by the institution. Move the same game to an online site with 5% rake capped at \$3 and the threshold drops. Move it to Macau where the rake is nominally 2.5% and the threshold drops further. Move it to a home game with no rake and the threshold is exactly the zero-sum threshold.

The same abstract skill (ability to find the Nash equilibrium strategy and execute it) translates to very different bankrolls under different institutions. This is why serious players select venues as carefully as they select spots within a hand.

Concrete situation: online vs live as different games

Live poker and online poker are structurally different games not because the cards are different but because the institutions are different. Live poker has physical tells, table-captain dynamics, live-ratio pacing (slower play, fewer hands per hour), and physical presence norms. Online poker has multi-tabling, HUD availability, faster pacing, and weaker enforcement of some norms. A live specialist who is crushing \$5/\$10 live can easily be losing at \$5/\$10 online because the effective games are different even though the raw game is the same. The transition between the two is a change of institution, not a change of difficulty.

What this understanding buys you

The choice of where to play is a strategic choice on par with how to play. Many serious poker players treat venue selection as a tactical question (availability, comfort, convenience). Recognizing that venues are institutions that select different effective games promotes this to a strategic question. The right venue for a given player is the one whose institutional transformation is most favorable to that player's actual skill profile. A player who reads physical tells well should play live; a player who exploits HUD data well should play online; a player who specializes in exploiting weak recreational players should play in a venue with a high recreational-player ratio.

Complaining about the institution is a category error. When a player says "the rake is too high" or "PokerStars is rigged" they are complaining about the institutional terms of the game. These terms are not a feature of the raw game; they are the institution. The correct response is not to complain but to either (a) accept the institution and recalibrate your expected value, or (b) switch institutions, or (c) participate in the meta-game that sets institutional terms (venue competition, community pressure, regulation). These are three different strategic responses to the same observation.

Institutional arbitrage is a real edge. Differences in rake structure, player-pool quality, and enforcement across venues create arbitrage opportunities for players who are willing to travel, play online, or specialize in a particular venue type. The very best professionals know the institutional landscape of poker across continents and select venues accordingly. This is 10.1 in action.

Cross-Links

- **6.1 Mechanism design** — institutions as large-scale mechanism-design problems.
- **5.4 Folk theorem** — which equilibria of the raw game survive institutional selection.
- **10.2 Layered mechanisms** — the nested-mechanism structure of institutional design.
- **10.4 Commitment and credibility** — why enforcement mechanisms work (or fail).
- **10.5 Political economy and collective choice** — the meta-game that produces institutions.
- **7.4 Evolutionarily stable strategies** — the analogous notion of institutional stability.
- **Economics / institutional economics** — North, Ostrom, Acemoglu: institutions as rules of the game.

Structural Compression

Invariant: Every real strategic situation is the object-level component of an institutional embedding; the institution determines which game is being played by restricting strategies, transforming payoffs, and enforcing rules; analysis must account for both the object-level game and the meta-game that produces the institution.

Mechanism: The institution transforms a raw game $G^{\text{raw}} = (N, S^{\text{raw}}, u^{\text{raw}})$ into an effective game G^{eff} by deleting prohibited strategies, subtracting institutional transfers, and adding enforcement-cost terms for rule-violating strategies; the institution itself is selected by a meta-game whose equilibrium is a self-enforcing institutional fixed point.

Cross-System Echo: This is structurally identical to the **layered protocol stacks** of computer networking (each layer's "game" is defined by the rules of the layer below it) and to the **nested gauge groups** of theoretical physics (each level of symmetry is a constraint inherited from a higher level). In both cases, the object at a given level is only well-defined given a choice of the level above. Institutions are the gauge-fixing of strategic situations: they select a particular representative from an equivalence class of formally indistinguishable object-level games. This theme will recur in every remaining node of Module 10.

Exercises

1. Give three examples of economic situations where the same raw game is embedded in two different institutions, and describe how the effective game differs under each. At least one example should be from poker.

2. Let G^{raw} be the prisoner's dilemma with payoffs $(C, C) = (3, 3), (D, C) = (5, 0), (C, D) = (0, 5), (D, D) = (1, 1)$. Define an institution that imposes a fine of f on any player who plays D, detected with probability p. Write the effective payoff matrix. Find the smallest pf such that (C, C) is a Nash equilibrium of the effective game.
3. Suppose the meta-game of institutional choice has two candidate institutions I_1 and I_2. The object-level equilibrium under I_1 gives players payoffs $(4, 4)$ and under I_2 gives $(5, 2)$. Under simple majority voting among the two players, which institution is chosen? Under unanimity? What does this tell you about the institutional-selection meta-game's sensitivity to voting rules?
4. Consider online poker with a HUD (heads-up display) allowed versus banned. Write down the effective strategy space in each case and argue that the allowed-HUD institution is not Pareto-dominant even though it gives each player more information.
5. Define "self-enforcing institution" precisely and show that grim-trigger-based enforcement of a rule against collusion can fail to be self-enforcing if the detection probability is below a threshold. Compute the threshold in a simple two-player setting.
6. A casino charges rake r per pot. Show that there is a critical r^* above which the highest-skill player at the table prefers to leave (because their post-rake edge goes negative). What does r^* depend on, and why does the casino want to set $r < r^*$?
7. **Argue, structurally,** why the distinction between "the game" and "the institution in which the game is embedded" is not merely a labeling convenience but a genuine ontological feature of strategic situations. Give one example from outside game theory where failing to distinguish the two leads to a substantive mistake.

Game Theory · Module 10 — Strategic Systems Integration · Node 10.1
Concepts: institutions, mechanism-design, nash-equilibrium
Prerequisites: 1.1, 2.1, 5.3, 6.1
Enables: 10.2, 10.3, 10.4, 10.5, 10.6

hbar.university — own.your.cognition