A State Is Its Relations: The Yoneda Lemma and Relational Identity in International Relations

Robert J. Carroll*

April 6, 2026

Abstract

International relations theory has been moving toward a relational ontology, but the commitment has remained difficult to formalize. Statistical tools model patterns in relational data without theorizing how relations compose; formal-theoretic tools capture relational dynamics only at small scale. This paper provides a general formal language, drawing on category theory. A theory is a category—actors and typed, composable relations. A model is a functor to the category of sets. A translation between models is a natural transformation. The central result is the Yoneda lemma: within any such theory, an actor's identity is its complete relational profile—the totality of relationships every other actor bears to it. This is a theorem, not a philosophical stance, available to rationalists, constructivists, and network analysts alike. Applied to two distinct bargaining models of war, the naturality condition derives a previously unnoticed equilibrium constraint linking their parameter spaces.

Keywords: category theory, Yoneda lemma, relational ontology, state identity, formal theory, metatheory

*Department of Political Science, University of Illinois at Urbana-Champaign. Email: rjc@illinois.edu. Working paper; comments welcome.

1 Introduction

International relations theory has been moving toward a relational ontology for decades. Constructivists argue that state identity is constituted by interaction, not given prior to it (Wendt, 1992; Adler, 1997; Hopf, 1998). Network analysts foreground the structure of ties over unit attributes (Hafner-Burton et al., 2009; Ward et al., 2011). Students of hierarchy treat authority itself as a relational contract (Lake, 2009; Nexon, 2009). Behind these programs lies a common ontological commitment: what Emirbayer (1997) calls *relationalism*—the view that entities are constituted by the transactions in which they participate, rather than entering those transactions with a pre-given nature. Ruggie (1998) names this commitment explicitly in IR: a "relational ontology" in which the identities and interests of actors are constructed by shared ideas rather than fixed by material endowment.

The commitment is widely shared but difficult to formalize. Statistical tools for relational data—network centrality measures, exponential random graph models, latent space models—have made important empirical progress (Cranmer et al., 2012; Hoff and Ward, 2004), but they model *patterns* in relational data rather than providing a general theoretical language for what follows from taking relations as primitive. They do not, for instance, tell you how relations *compose*, or what composition implies about identity. Meanwhile, formal-theoretic tools that *do* capture relational dynamics—bargaining models, contest models, mechanism design—work at small scale: two actors, one issue, a specific strategic structure. The result is a gap. The tools that are general are not theoretical; the tools that are theoretical are not general. We lack formal machinery for asking what it means, in full generality, to take relations as the foundation of a theory of international politics.

This paper provides such machinery and shows that the relational commitment has a deep consequence. The framework draws on category theory, a branch of mathematics developed in the 1940s to study relationships between mathematical objects (Mac Lane, 1998). In a category, the primitive data are not objects or properties but *morphisms*—typed, directed relationships that compose associatively. Objects are characterized entirely by how they relate to other objects.

The central result is the *Yoneda lemma*. It says: within any category, an object is completely determined, up to isomorphism, by the totality of relationships that all other objects bear to it. There are no hidden properties, no intrinsic essence, no residual identity beyond the relations. Applied to international relations: a state's identity, within any relational theory, is nothing more and nothing less than its complete relational profile. This is not a philosophical stance. It is a theorem—one that holds in any framework satisfying the minimal axioms of a category, regardless of whether the framework is rationalist, constructivist, or eclectic.

1.1 What the paper does

The paper builds the necessary vocabulary from scratch within the context of IR. A *theory* is formalized as a category—actors connected by typed, composable relations (Section 2). A *model* is a functor to the category of sets—a structure-preserving interpretation that assigns data to actors and functions to relations (Section 3). A *translation* between models is a natural transformation—a systematic conversion that respects the relational structure at every point (Section 4). Applied to two well-known bargaining models of war (Fearon and Bevíá–Corchón), the naturality condition identifies exactly which parameter configurations in one model can be microfounded by the other.

The *relational profile* of a state is the representable functor $\mathrm{Hom}(-, X)$: the complete record of how every other actor relates to X (Section 5). The *Yoneda lemma* proves that every model's assessment of X is a projection of this profile, and the *Yoneda embedding* proves that two states are structurally identical if and only if their profiles agree. Functors between categories formalize *theory comparison*: "does trade cause peace?" becomes a question about a specific functor from the trade category to a conflict category, subject to compositional constraints (Section 6).

1.2 What the paper contributes

The paper makes three contributions.

First, it provides a *formal framework for relational ontology* in IR. The claim that identity is relational has been advanced by constructivists (Wendt, 1992), network analysts (Duque, 2018), and students of hierarchy (Lake, 2009), but it has remained a philosophical commitment or an empirical finding rather than a formal result. Checkel (1998) argued that the constructivist research program's central challenge is to develop genuine theory; Finnemore and Sikkink (2001) characterized it as a "framework, not a substantive theory." The Yoneda lemma shows that any framework satisfying the minimal axioms of a category—any framework that takes relational structure seriously—will, as a matter of mathematical necessity, yield the conclusion that identity is relational profile.

Second, it provides a *precise metatheory* for comparing IR models. Whether two models are commensurable has a determinate answer in terms of natural transformations; whether two theories with different primitives can be related has a determinate answer in terms of functors between categories.

Third, it provides *concrete analytical tools*: a provable constraint on when the Fearon and Beví́á–Corchón bargaining models are commensurable, a precise account of what is lost when relational structure is flattened into attribute tables, and a formalization of "does trade cause peace?" as a functor between categories.

The paper is self-contained and written so that an IR scholar with no prior exposure to category theory can follow the argument on a first reading. Every concept is defined from scratch, motivated by IR examples, and illustrated with a running four-state trade network (Laos, Vietnam, China, the United States) that threads through every section. The mathematics is elementary—the Yoneda lemma's proof is a few lines—but its consequences are not. The paper is addressed to IR theorists across paradigmatic lines—constructivists, rationalists, and network analysts alike—who want a formal language for the relational claims their work already makes.

2 What Is a Theory of International Relations?

What does it mean to have a *theory* of international relations? At minimum, a theory specifies a domain of actors and a type of relationship between them. Realism foregrounds power relations among states. Liberalism foregrounds institutional and economic ties. Constructivism foregrounds the social relationships—recognition, enmity, friendship—through which actors constitute one another. In each case, the theory tells us who the relevant actors are and what kind of connection between them matters.

Category theory offers a way to make this intuition precise. A *category* is a mathematical structure that captures exactly this: a collection of objects and a specified type of relationship (called *morphisms*) between them, subject to minimal axioms that any reasonable notion of "relationship" should satisfy.

The claim of this paper is that a theory of international relations, understood at the appropriate level of abstraction, *is* a category.

2.1 Categories as theories

Definition 2.1 (Category)**.** A *category* $\mathcal{C}$ consists of:

(i) a collection of *objects*, $\mathrm{Ob}(\mathcal{C})$;

(ii) for each pair of objects A, B, a set of *morphisms* $\mathrm{Hom}_{\mathcal{C}}(A, B)$;

(iii) for each triple of objects A, B, C, a *composition* operation

$$\circ\colon \mathrm{Hom}(B, C) \times \mathrm{Hom}(A, B) \to \mathrm{Hom}(A, C);$$

(iv) for each object A, an *identity morphism* $\mathrm{id}_A \in \mathrm{Hom}(A, A)$;

subject to the axioms:

- **Associativity.** For morphisms $f\colon A \to B$, $g\colon B \to C$, $h\colon C \to D$, we have $h \circ (g \circ f) = (h \circ g) \circ f$.
- **Identity.** For any morphism $f\colon A \to B$, we have $f \circ \mathrm{id}_A = f = \mathrm{id}_B \circ f$.

The definition is spare, and deliberately so. Each component carries a substantive interpretation in international relations:

Objects as actors. The objects of the category are the actors that the theory takes as primary. These are typically states, but nothing in the formalism prevents us from taking non-state actors, international organizations, firms, or individuals as objects when the theory demands it. What matters is that the theory specifies a determinate domain.

Morphisms as relations. For each pair of actors A and B, the set hom-set $\mathrm{Hom}(A, B)$ collects the relationships of the relevant type that run from A to B. Crucially, morphisms are *typed* and *directed*. A trade relationship is not an alliance relationship; an export flow from Germany to Poland is not the same as an export flow from Poland to Germany. The choice of what counts as a morphism is the fundamental theoretical commitment: it is what distinguishes one theory from another.

A single category captures one type of relationship at a time. Multiple categories on the same set of actors can coexist, and *functors* between them (Section 6) formalize how one relational structure maps onto another.

Composition as transitivity. If A bears a relationship to B and B bears a relationship to C, then there exists a composite relationship from A to C. This says that the relational type under study is *transitive* in a structured way: indirect connections, mediated through third parties, are themselves relationships of the same type. This is a modeling choice, not a fact about the world. Some relational types compose naturally (trade flows, supply chains); others compose only under idealization (alliance commitments, authority relations). The choice of whether composition holds—and what the composite means—is part of the theoretical commitment encoded in the category. Morphisms the theorist considers irrelevant can always be sent to trivial maps by the model (the functor), but their formal existence must be acknowledged; this is part of the discipline the framework imposes.

Identity as self-relation. Every actor bears a distinguished relationship to itself. In a trade category, this is the domestic economy—the baseline of self-exchange against which foreign trade is measured. In an alliance category, it is the commitment to self-defense that exists prior to any alliance. These are not trivial glosses: any theory of international politics must announce what self-relation consists in, and different theories give different answers. The identity morphism is the formal expression of this commitment.[1]

Associativity as coherence. When chaining three or more relationships, the order of composition does not matter: there is no ambiguity in what it means to trace a path through the network of relationships.

Remark 2.2. A category looks like a directed graph with extra structure, and this is instructive. Network analysis has made important inroads into IR by foregrounding relational structure over unit attributes (Hafner-Burton et al., 2009; Maoz, 2012), but a network says nothing about how ties *compose*. Centrality measures and brokerage scores describe patterns of ties; they do not theorize the algebra of how ties combine. Category theory adds exactly

[1]The word "identity" is doing double duty in this paper. The *identity morphism* id_A is a structural element of the category—it says that A stands in a self-relation and that this self-relation is neutral under composition. The *identity of a state*—the question of what makes A the actor it is—is the subject of Section 5, where it will turn out to be determined by the totality of A's relationships to every other actor, not by the self-loop alone. The two notions are related but distinct: the identity morphism tells us what a theory takes self-relation to be; the Yoneda lemma tells us that a state's identity is constituted by all its relations, including but not limited to the self-relation.

this: composition is an axiom, and the identity and associativity conditions ensure coherence.[2]

2.2 Example: The trade category

Example 2.3 (The trade category $\mathcal{T}$)**.** Let the objects of $\mathcal{T}$ be states. For states A and B, a morphism $f \in \mathrm{Hom}_{\mathcal{T}}(A, B)$ represents a trade relationship from A to B—an export flow, a trade agreement, a supply-chain dependency. There may be multiple morphisms between the same pair of states, reflecting different trade channels or agreements.

Composition captures transitive linkage: if $f\colon A \to B$ represents A's export relationship to B and $g\colon B \to C$ represents B's export relationship to C, then the composite $g \circ f\colon A \to C$ represents the indirect trade relationship between A and C mediated by B. This is the formal counterpart of the intuition that global supply chains create relationships between states that do not trade directly.

The identity morphism id_A represents A's domestic economy—the self-exchange that is always available and that leaves other trade relationships unchanged when composed with them.

To make this concrete, consider a small fragment of the trade category with four objects: the United States, China, Vietnam, and Laos. The morphisms include, among others: a massive bilateral export flow from China to the United States; an export flow from Vietnam to China driven by electronics assembly; a smaller flow from Laos to Vietnam in agricultural goods; and so on. Composition gives us the indirect trade linkages: the composite of the Laos→Vietnam and Vietnam→China morphisms is an indirect trade relationship from Laos to China mediated by Vietnam—a relationship that is real (Laotian raw materials enter Chinese supply chains via Vietnamese intermediaries) but invisible in any dataset that records only direct bilateral flows.

We will return to this four-state example throughout the paper.

[2]As Zhukov and Stewart (2013) demonstrate, network-based IR results are acutely sensitive to the choice of network specification. In our framework, each specification is a different category (same objects, different morphisms), and the sensitivity is the expected consequence of working in different relational structures. The question of how results in one network translate to another becomes a question about functors between categories (Section 6).

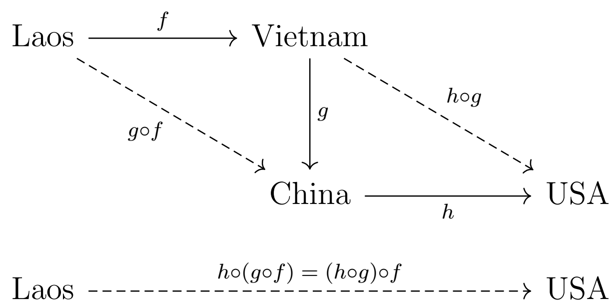

Figure 1: A fragment of the trade category $\mathcal{T}$. *Top:* solid arrows are direct trade morphisms; dashed arrows are composites. The composite $g \circ f$ (Laos to China, via Vietnam) and $h \circ g$ (Vietnam to the USA, via China) represent different intermediate mediations. *Bottom:* the fully mediated relationship from Laos to the USA. Associativity guarantees that the two ways of computing it agree: $h \circ (g \circ f) = (h \circ g) \circ f$.

The trade category specifies actors and trade relationships but does not say what trade *means*—whether it produces welfare gains, shifts power balances, or generates interdependence. The category is the structural skeleton; the model is the flesh. That interpretive step requires a *functor* (Section 3).

2.3 Example: The alliance category

Example 2.4 (The alliance category $\mathcal{A}$)**.** Let the objects of $\mathcal{A}$ be states. A morphism $f \in \mathrm{Hom}_{\mathcal{A}}(A, B)$ represents an alliance commitment from A to B—a mutual defense pact, a security guarantee, or an informal alignment. Composition captures *extended deterrence*: if A is allied to B and B is allied to C, the composite $g \circ f$ represents the chain of commitments linking A to C. Whether such chains are credible is an empirical and strategic question; the category records their formal existence as part of the alliance structure.

The identity morphism captures a state's commitment to its own defense—the baseline security relationship that exists independently of any alliance.

The alliance category illustrates an important feature of the framework: the same set of states can be organized into different categories by choosing different morphism types. The United States, Japan, and Australia are the same objects in both $\mathcal{T}$ and $\mathcal{A}$, but their morphisms—trade flows versus security commitments—are entirely different. This is precisely the sense in which the choice of morphisms constitutes a theoretical commitment.

2.4 What the framing already makes visible

A theory, in the sense intended here, is a *signature*—a specification of the basic vocabulary (actors, relations, composition) without an interpretation. To give it content requires a *functor* (Section 3). Cross-theory questions like "does trade cause peace?" require functors *between* categories (Section 6).

But even before introducing models, the category-theoretic framing makes something visible. The dominant data infrastructure of quantitative IR—attribute tables with rows indexed by states and columns indexed by properties—encodes an ontology in which states are atoms characterized by intrinsic properties, with relations absent or flattened into dyadic variables that discard composition (Hoff and Ward, 2004; Beckfield, 2010). Ward et al. (2011) document the pattern: most quantitative work reduces relational structure to node-level statistics (centrality, betweenness), then feeds those statistics into independence-assuming regression models—precisely the compression from a relational category to a discrete one. The consequences are not merely aesthetic. Cranmer et al. (2012) show that when alliance data are modeled as a network (via exponential random graph models) rather than as independent dyads, the effect of joint democracy on alliance formation *reverses sign*—a standard finding in the literature is an artifact of discarding relational structure. Duque (2018) demonstrates the same principle for international status: when the full network of diplomatic recognition is modeled rather than compressed into a count of embassies received, endogenous relational structure—not material attributes—drives status. Network analysis recovers the relational structure that attribute tables discard; category theory adds the compositional structure that networks themselves lack—how relations chain, what identity means, what the Yoneda lemma implies—and provides a theory of databases that makes each dataset's ontological commitments explicit (Section 3.6).

3 What Is a Model?

A theory, in the sense of Section 2, specifies actors and relations but says nothing about what those relations produce, measure, or mean. The trade category $\mathcal{T}$ tells us that states are connected by trade, but not whether trade enriches, empowers, or endangers them. To interpret a theory—to give it empirical or normative content—we need to assign concrete data to its abstract structure. The mathematical tool for this is the *functor*.

3.1 Functors as structure-preserving maps

Definition 3.1 (Functor)**.** Let $\mathcal{C}$ and $\mathcal{D}$ be categories. A *functor* $F\colon \mathcal{C} \to \mathcal{D}$ consists of:

(i) an assignment of an object $F(A) \in \mathcal{D}$ for each object $A \in \mathcal{C}$;

(ii) an assignment of a morphism $F(f)\colon F(A) \to F(B)$ in $\mathcal{D}$ for each morphism $f\colon A \to B$ in $\mathcal{C}$;

subject to:

- $F(\mathrm{id}_A) = \mathrm{id}_{F(A)}$ for all objects A;
- $F(g \circ f) = F(g) \circ F(f)$ for all composable morphisms f, g.

The definition says: a functor is a map from one category to another that preserves the relational structure. It sends objects to objects, morphisms to morphisms, and respects both composition and identity. Nothing is lost in translation—every relationship in the source category has an image in the target, and the way relationships compose is faithfully tracked.

3.2 Models as functors to Set

The general definition allows functors between any two categories, but one target category is of special importance.

Definition 3.2 (Model of a theory)**.** A *model* of a theory $\mathcal{C}$ is a functor $F\colon \mathcal{C} \to \mathbf{Set}$, where **Set** is the category whose objects are sets and whose morphisms are functions.

What does this mean concretely? A model assigns to each actor A a set $F(A)$—the "data" associated with A under this interpretation. These might be welfare levels, power rankings, policy positions, or observable behaviors; the choice is the model's to make. To each relationship $f\colon A \to B$, the model assigns a function $F(f)\colon F(A) \to F(B)$ that specifies how the relationship transforms data—how A's attributes are carried, filtered, or distorted as they pass through the relation to B.

The two functorial axioms now carry substantive meaning. The condition $F(\mathrm{id}_A) = \mathrm{id}_{F(A)}$ says that the self-relation leaves a state's data unchanged—the identity morphism, whatever substantive content it carries at the level of the theory (Section 2.1), acts trivially on the model's data. The condition $F(g \circ f) = F(g) \circ F(f)$ says that the model respects mediated relationships: the data transformation induced by the indirect link from A to C (via B) is the same whether we compute it in one step or two.[3]

This framing inverts the usual relationship between theory and data. In standard practice, we begin with a dataset—a table of attributes—and then ask which theory best explains the patterns. Here, the theory comes first as a relational structure, and the dataset is a *consequence* of the choice of model: the functor F determines what gets measured, for whom, and how measurements at different nodes are related. Different functors applied to the same theory yield different datasets—not because the world has changed, but because the interpretive lens has.

3.3 Two models of the trade category

Example 3.3 (Liberal and realist models of trade)**.** Consider two models of the trade category $\mathcal{T}$ from Example 2.3:

- The *liberal model* $F\colon \mathcal{T} \to \mathbf{Set}$ assigns to each state A the set $F(A)$ of possible welfare outcomes for A's economy—real income levels, consumer surplus, gains from specialization. To each trade morphism $f\colon A \to B$, it assigns a function $F(f)\colon F(A) \to F(B)$ that tracks how A's welfare position translates into welfare effects on B through the trade link. The liberal model reads trade as a mechanism of mutual enrichment.

- The *realist model* $G\colon \mathcal{T} \to \mathbf{Set}$ assigns to each state A the set $G(A)$ of relative power positions—military capacity, economic leverage, technological advantage relative to competitors. To each trade morphism $f\colon A \to B$, it assigns a function $G(f)\colon G(A) \to G(B)$ that maps A's power position to the power shift induced in B by the trade relationship. The realist model reads trade as a vector of strategic dependence.

[3] This is Lawvere's insight: the classical notion of a "model of a theory" in mathematical logic is exactly a functor from the theory (understood as a category) to **Set** (Lawvere, 1963). We are applying the same idea to theories of international relations.

Both F and G are models of the *same* underlying theory $\mathcal{T}$. They agree on who the actors are and what relationships exist between them. They disagree on what those relationships *mean*—on the sets assigned to each state and the functions assigned to each trade link.

This is the precise sense in which the liberal–realist debate about trade is not a disagreement about structure but about interpretation. The two camps share a theory (states linked by trade) but operate with different models (welfare versus power). The question of whether they can be systematically compared—whether there is a coherent translation from welfare data to power data that respects the trade structure—is a question about *natural transformations*, the subject of Section 4.

3.4 The running example, modeled

Return to the four-state fragment of $\mathcal{T}$: Laos, Vietnam, China, and the United States (Figure 1). Under the liberal model F, each state is assigned a set of welfare outcomes:

$F(\text{Laos})$	welfare outcomes for Laos (agricultural export revenues, etc.)
$F(\text{Vietnam})$	welfare outcomes for Vietnam (manufacturing wages, FDI effects, etc.)
$F(\text{China})$	welfare outcomes for China (supply-chain value-added, etc.)
$F(\text{USA})$	welfare outcomes for the USA (consumer prices, import competition, etc.)

The trade morphism $f\colon \text{Laos} \to \text{Vietnam}$ maps to $F(f)\colon F(\text{Laos}) \to F(\text{Vietnam})$, tracking how Laotian welfare translates into Vietnamese welfare effects. The composite $g \circ f\colon \text{Laos} \to \text{China}$ maps to $F(g) \circ F(f)$—the indirect welfare linkage mediated by Vietnam. Under the realist model G, the same composite maps to $G(g) \circ G(f)$—not welfare effects but strategic dependencies. The structure is identical; the interpretation is not.

3.5 Strategic interaction as a multi-sorted theory

The trade and alliance categories of Section 2 have a single sort of object: states. But many theories in IR involve objects of fundamentally different kinds. Game-theoretic models are a natural example.

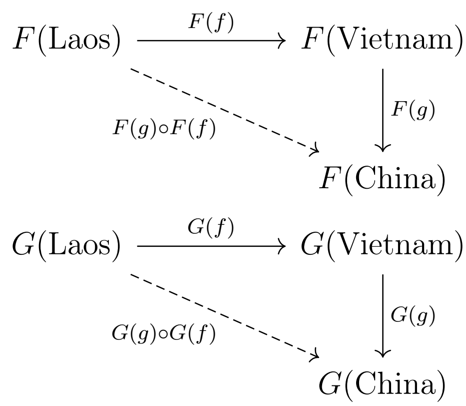

Figure 2: The liberal model F (above) and the realist model G (below) applied to the same fragment of the trade category. Both functors preserve the compositional structure: the indirect link from Laos to China decomposes identically. What differs is the content of the sets and the meaning of the functions.

Example 3.4 (The bargaining skeleton)**.** Consider the class of bargaining models of war. At the highest level of abstraction, these models share a common structure—a *bargaining skeleton* $\mathcal{B}$ whose objects fall into three sorts:

- *Players*—the states engaged in the dispute;
- *The good*—the object of contention (territory, policy, status);
- *Outcomes*—the possible resolutions (a peaceful division, or war with its attendant costs and probabilities).

The morphisms encode the structure of strategic interaction: there are morphisms from players to the good (demand functions), from players to outcomes (payoff maps), and a contest morphism that determines who prevails in war.

Different bargaining models are different *functors* on this shared skeleton. Two of the cleanest examples—chosen because they share the same primitives but differ sharply in their assumptions—are the following.

Example 3.5 (Fearon's model as a functor)**.** Fearon's (1995) model sends $\mathcal{B}$ to **Set** as follows. The good maps to $[0, 1]$. Each player's action space is a set of offers and accept/reject decisions. The contest morphism maps to an exogenous probability $p \in (0, 1)$, and costs of war are exogenous parameters $c_1, c_2 > 0$. The payoff morphisms send player 1 to x (if peace at division x) or to $p - c_1$ (if war), and player 2 to $1 - x$ or $(1 - p) - c_2$. Call this functor F.

Example 3.6 (Bevíá and Corchón's model as a functor)**.** Bevíá and Corchón (2010) model the same skeleton differently. The good maps to total resources $V = V_1 + V_2 \in \mathbb{R}_+$. Each player's action space now includes transfers and effort choices e_i—richer than Fearon's accept/reject. The contest morphism maps not to an exogenous p but to an endogenous contest success function $p_1 = \lambda e_1^\gamma / (\lambda e_1^\gamma + e_2^\gamma)$, where λ measures military proficiency and γ measures the sensitivity of winning to effort. Costs are endogenous: war effort costs ke_i, a fixed proportion of resources committed. Call this functor G.

The two models share the same primitives—players, a contested good, a contest, payoffs—but their assumptions differ in a structured way. Fearon treats the probability of winning and the costs of war as exogenous parameters; Bevíá and Corchón endogenize both through the effort choices and the contest success function. In our language: F and G are two functors on the same category $\mathcal{B}$, differing in what they assign to the contest and cost morphisms.

Notice the role of equilibrium. In Bevíá and Corchón's model, the contest success function depends on effort levels, so the functor G is well-defined only once the game has been solved: equilibrium is part of the specification of G's assignments, not a separate layer. Fearon's functor F sidesteps this—its assignments are exogenous parameters. The difference between the two functors is not just a difference in parameter values but in *what kind of theoretical work is required to specify the functor at all.*

This clarifies a distinction that game theorists often leave implicit: the *primitives* of a model (players, good, payoffs) are the objects and morphisms of a category, while the *assumptions* (that utility is linear, that costs are exogenous) are the functor's assignments. Two models sharing primitives but differing in assumptions are two functors on the same category; models differing in primitives are functors on different categories (Section 6).

The question of whether Fearon's model and Bevíá and Corchón's model can be systematically related—whether there is a coherent translation from one to the other—is a question about *natural transformations*, the subject of Section 4.

3.6 Existing datasets as implicit functors

The functor framework makes it possible to say precisely what existing IR datasets are doing—and what they are missing. The idea that a database schema is a category and a database instance is a functor to **Set** is well established in applied category theory; Fong and Spivak (2019, ch. 3) give a thorough introduction. Our contribution is to apply this lens to the data infrastructure of IR, where the implicit categorical commitments have gone unexamined.

The Penn World Tables as a functor. The Penn World Tables (PWT) assign to each state a vector of macroeconomic indicators: real GDP, capital stock, productivity levels, terms of trade. In our framework, this is a functor—but a very particular one. The PWT treats the international system as a *discrete category*: a collection of states with no morphisms other than identities. The functor F_{PWT} assigns to each state A the set $F_{\text{PWT}}(A)$ of its macroeconomic data, and to each identity morphism id_A the identity function on that set. There is nothing else to assign, because the theory underlying the PWT has no relations.

A dataset organized as an attribute table, with rows indexed by states and no relational structure, is a model of a theory in which states are atoms: unconnected, self-contained, characterized entirely by intrinsic properties.

The COW bilateral trade dataset as a functor. The Correlates of War bilateral trade data is richer. It records, for each pair of states, a trade volume—making it a model of something closer to the trade category $\mathcal{T}$. But the COW data does not record composition. The trade volume between the United States and China is a direct measurement; the mediated trade relationship from Laos to the United States via Vietnam and China (the composite $h \circ g \circ f$ in Figure 1) is not a row in the dataset. In functorial terms, the COW data is a model of the *underlying graph* of $\mathcal{T}$—the directed graph that records which morphisms exist—but not of the full category, because it does not track how trade relationships compose.[4]

[4] More precisely, the COW bilateral trade data can be understood as a functor on the *free category* generated by the trade graph, modulo the additional assumption that only direct edges carry data. The composite morphisms exist in the free category, but the functor assigns them no independent empirical content.

The COW National Material Capabilities index (CINC score) is similarly a functor on a discrete category—a realist model in miniature, assigning to each state a measure of material capability. The liberal economist's objection that GDP per capita is a better measure is not a dispute about data but about which functor to apply.

In each case, the shape of the dataset encodes theoretical commitments: the PWT discards all morphisms; the COW trade data discards composition. The cost of these choices becomes precise once we have the Yoneda lemma (Section 5).

4 Translating Between Models

We now have the vocabulary to say what a theory is (a category) and what a model is (a functor to **Set**). Section 3.3 showed that the liberal and realist models of the trade category interpret the same relational structure in different ways—welfare versus power. The natural question is: can we systematically translate between them?

Not just for one state, but for *every* state, and in a way that respects the relational structure? The answer lies in the concept of a *natural transformation*—the central notion of coherent translation in category theory, and arguably the concept that motivated the entire subject.[5]

4.1 Definition and the naturality condition

Definition 4.1 (Natural transformation)**.** Let $F, G\colon \mathcal{C} \to \mathbf{Set}$ be two models of a theory $\mathcal{C}$. A *natural transformation* $\eta\colon F \Rightarrow G$ consists of, for each object A in $\mathcal{C}$, a function $\eta_A\colon F(A) \to G(A)$, such that for every morphism $f\colon A \to B$ in $\mathcal{C}$, the following diagram commutes:

$$\begin{array}{ccc} F(A) & \xrightarrow{\eta_A} & G(A) \\ {\scriptstyle F(f)}\big\downarrow & & \big\downarrow{\scriptstyle G(f)} \\ F(B) & \xrightarrow[\eta_B]{} & G(B) \end{array}$$

[5] Eilenberg and Mac Lane introduced categories and functors in 1945 primarily as scaffolding for defining natural transformations. As Mac Lane later wrote, "I didn't invent categories to study functors; I invented them to study natural transformations" (Mac Lane, 1998, p. 29).

That is, $G(f) \circ \eta_A = \eta_B \circ F(f)$.

The diagram is called the *naturality square*, and the condition it expresses is called *naturality*. In words: translating A's data and then pushing it through G's version of the relationship gives the same result as pushing it through F's version of the relationship and then translating B's data. Translation commutes with the relational structure.

This is a strong requirement. A natural transformation is *not* merely a collection of functions $\eta_A \colon F(A) \to G(A)$, one per state, chosen independently. The naturality condition ties these functions together: they must be compatible with every morphism in the theory. Any translation that works for isolated states but breaks down when you account for their relationships is not natural—and in this framework, it is not a legitimate translation at all.

4.2 The naturality condition in IR

Return to the liberal model F and the realist model G of the trade category $\mathcal{T}$ (Example 3.3). A natural transformation $\eta \colon F \Rightarrow G$ would be a systematic way to convert welfare data into power data for every state, such that the conversion respects trade linkages.

Concretely, consider the trade morphism $f \colon \text{Laos} \to \text{Vietnam}$ (Laotian agricultural exports to Vietnam). The naturality square for f says:

$$\begin{array}{ccc} F(\text{Laos}) & \xrightarrow{\eta_{\text{Laos}}} & G(\text{Laos}) \\ {\scriptstyle F(f)}\downarrow & & \downarrow{\scriptstyle G(f)} \\ F(\text{Vietnam}) & \xrightarrow[\eta_{\text{Vietnam}}]{} & G(\text{Vietnam}) \end{array}$$

Reading around the square:

- *Top then right*: Take Laos's welfare data. Convert it to power data (via η_{Laos}). Then apply the realist trade function $G(f)$ to see how Laos's power position affects Vietnam.

- *Left then bottom*: Take Laos's welfare data. Apply the liberal trade function $F(f)$ to see how Laos's welfare affects Vietnam's welfare. Then convert Vietnam's welfare data to power data (via η_{Vietnam}).

Naturality demands that these two paths give the same answer. Converting welfare to power *before* tracing a trade link must agree with tracing the trade link *first* and then converting. The translation cannot distort the relational structure.

If the translation breaks for any state or any trade morphism, the naturality condition fails and no natural transformation exists between the two models.

4.3 When translation fails: incommensurability

The absence of a natural transformation between two models is not a deficiency of the framework; it is an informative result. It means that the two models assign meaning to the relational structure in ways that cannot be coherently reconciled.

Consider what this implies for the liberal–realist debate about trade. If no natural transformation $F \Rightarrow G$ exists, then there is no systematic, structure-respecting way to translate welfare data into power data across the international system. The two interpretations are not merely different; they are *incommensurable* in a precise sense: the relational structure of trade imposes constraints on translation that cannot be simultaneously satisfied.

This gives formal content to a claim that has long circulated in IR metatheory: that paradigms can be genuinely incommensurable, not just different in emphasis. Jackson (2011) argues that different approaches to IR are grounded in distinct philosophical ontologies that resist straightforward comparison. Monteiro and Ruby (2009) push back, arguing that the appearance of incommensurability dissolves under sufficiently careful philosophical analysis. In the present framework, the question is neither philosophical nor verbal—it is structural. Two models of the same theory are commensurable if and only if a natural transformation exists between them. Whether one does is a mathematical question with a determinate answer, not a matter of interpretation.

Remark 4.2. A natural transformation $\eta\colon F \Rightarrow G$ need not be invertible: one can translate from welfare to power without being able to translate back. When each η_A is a bijection, the two models are structurally indistinguishable (a *natural isomorphism*). More commonly, a natural transformation exists in one direction but not the other.

4.4 The running example

In the four-state fragment (Figure 1), a natural transformation $\eta\colon F \Rightarrow G$ must supply translation functions $\eta_A\colon F(A) \to G(A)$ at each state, satisfying:

$$\begin{aligned} G(f) \circ \eta_{\text{Laos}} &= \eta_{\text{Viet}} \circ F(f), \\ G(g) \circ \eta_{\text{Viet}} &= \eta_{\text{China}} \circ F(g), \\ G(h) \circ \eta_{\text{China}} &= \eta_{\text{USA}} \circ F(h). \end{aligned}$$

Composites generate no additional constraints (by functoriality), but these three equations are already demanding. Figure 3 displays the condition as a ladder of commuting squares: it is not enough to convert welfare to power state by state; the conversions must interlock across every trade relationship.

$$\begin{array}{ccc} F(\text{Laos}) & \xrightarrow{\eta_{\text{L}}} & G(\text{Laos}) \\ {\scriptstyle F(f)}\downarrow & & \downarrow{\scriptstyle G(f)} \\ F(\text{Viet}) & \xrightarrow{\eta_{\text{V}}} & G(\text{Viet}) \\ {\scriptstyle F(g)}\downarrow & & \downarrow{\scriptstyle G(g)} \\ F(\text{China}) & \xrightarrow{\eta_{\text{C}}} & G(\text{China}) \\ {\scriptstyle F(h)}\downarrow & & \downarrow{\scriptstyle G(h)} \\ F(\text{USA}) & \xrightarrow[\eta_{\text{U}}]{} & G(\text{USA}) \end{array}$$

Figure 3: The naturality condition for $\eta\colon F \Rightarrow G$ across the four-state trade chain. Each square must commute simultaneously.

4.5 Naturality and the bargaining model

The same logic applies to the game-theoretic setting of Section 3.5, and here it yields a particularly concrete payoff. Recall the two models of the bargaining skeleton $\mathcal{B}$: Fearon's functor F (Example 3.5), which treats the probability of winning p and the costs of war c_i as exogenous parameters, and Beviá and

Corchón's functor G (Example 3.6), which endogenizes both through effort choices and a contest success function.

A natural transformation $\eta\colon F \Rightarrow G$ must supply a function for each object in $\mathcal{B}$:

- $\eta_{\text{Good}}\colon [0,1] \to \mathbb{R}_+$, mapping Fearon's normalized pie to Bevíá and Corchón's resource space;
- η_{Player_i}, translating each player's data from one model to the other;
- η_{Outcome}, mapping Fearon's payoff space to Bevíá and Corchón's.

The naturality condition demands that these translations commute with the morphisms of $\mathcal{B}$—the demand functions, the contest morphism, and the payoff maps.

The critical square is the one involving the contest morphism. In Fearon's model, the contest morphism maps to an exogenous probability p. In Bevíá and Corchón's, it maps to the equilibrium probability $p^*(V_1, V_2, k, \lambda, \gamma) = \lambda e_1^{*\gamma}/(\lambda e_1^{*\gamma} + e_2^{*\gamma})$ determined by the players' optimal effort choices. The naturality square says: translating Fearon's player data into Bevíá and Corchón's and *then* applying the contest success function must give the same result as applying Fearon's exogenous p and *then* translating the outcome.

$$\begin{array}{ccc} F(\text{Players}) & \xrightarrow{\eta_{\text{Players}}} & G(\text{Players}) \\ {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle p^*} \\ F(\text{Outcome}) & \xrightarrow[\eta_{\text{Outcome}}]{} & G(\text{Outcome}) \end{array}$$

This is a substantive constraint, and we can make it completely explicit. In Bevíá and Corchón's model, the first-order conditions for optimal effort are:

$$\frac{\gamma\lambda e_1^{\gamma-1}e_2^{\gamma}}{(\lambda e_1^{\gamma}+e_2^{\gamma})^2}\cdot V = k \qquad \text{and} \qquad \frac{\gamma\lambda e_1^{\gamma}e_2^{\gamma-1}}{(\lambda e_1^{\gamma}+e_2^{\gamma})^2}\cdot V = k.$$

Dividing the first by the second gives $e_2^*/e_1^* = 1$: at equilibrium, both players exert the same effort. It follows that $p^* = \lambda/(\lambda+1)$ and, crucially, that the equilibrium costs of war satisfy $c_1^* = ke_1^* = ke_2^* = c_2^*$. Substituting back, $e^* = \gamma\lambda V/(k(\lambda+1)^2)$, so $c^* = \gamma p^*(1-p^*)V$.

The constraint $c_1^* = c_2^*$ is not an artifact of the symmetric cost parameter k. Even if we allow different marginal costs k_1, k_2 for the two players, the first-order conditions give $e_2^*/e_1^* = k_1/k_2$, and so $c_1^* = k_1 e_1^* = k_2 e_2^* = c_2^*$: equilibrium war costs are equal regardless.

In Fearon's model, c_1 and c_2 are free parameters—there is no reason for them to coincide. A natural transformation $\eta\colon F \Rightarrow G$ can therefore exist only for Fearon triples (p, c_1, c_2) with $c_1 = c_2$. Every triple with $c_1 \neq c_2$ lies outside the image of Beviá and Corchón's equilibrium map, and no structure-preserving translation can reach it.

This constraint does not appear in either Fearon (1995) or Beviá and Corchón (2010); it is derived here for the first time as a consequence of the naturality condition. Beviá and Corchón's model is *not* simply a generalization of Fearon's with endogenous parameters; it is a model that constrains the feasible (p, c) space to a proper subset—specifically, the locus $c_1 = c_2 = \gamma p(1-p)V$. The naturality condition identifies exactly where the two models are compatible and where they are not. The Fearon configurations outside that locus are the ones for which the exogenous-parameter assumption does real theoretical work—they describe scenarios that *cannot* arise from any contest game of the Beviá–Corchón type.

Remark 4.3 (What the constraint tells a modeler). The locus $c = \gamma p(1-p)V$ makes visible which theoretical ingredients are needed to microfound a given Fearon configuration. Fix the simplest contest: $\lambda = \gamma = 1$ (symmetric, proportional) and $V = 1$. Then $p^* = \lambda/(\lambda+1) = 1/2$ and $c^* = \gamma p^*(1-p^*)V = 1/4$, regardless of k. (The cost parameter k scales effort—$e^* = 1/(4k)$—but cancels out of equilibrium costs: $c^* = ke^* = 1/4$.) The entire image in Fearon's parameter space is a single point: $(p, c_1, c_2) = (1/2, 1/4, 1/4)$.

To reach other Fearon configurations, specific theoretical commitments are required. Generating $p \neq 1/2$ requires $\lambda \neq 1$: asymmetric military proficiency is the ingredient that moves the probability of winning. Generating different cost levels requires varying γ (non-proportional contest technology) or V (the size of the contested good). So a modeler studying a Fearon scenario with $p = 0.7$ and $c_1 = c_2 = 0.1$ who wants Beviá–Corchón microfoundations can read off the required parameters: $\lambda = p^*/(1-p^*) = 7/3$ and $\gamma V = c^*/[p^*(1-p^*)] \approx 0.476$. The naturality condition does not merely say that a translation exists or fails; it identifies, for each point in Fearon's parameter space, exactly what theoretical resources are needed to reach it from Beviá and Corchón's equilibrium map—and for every point with $c_1 \neq c_2$, it says that no contest game of this type will suffice.

4.6 What we can now say

With categories, functors, and natural transformations in hand, we have the first three layers of the framework:

Category theory	IR interpretation	Section
Category $\mathcal{C}$	A theory (actors + relations)	2
Functor $F\colon \mathcal{C} \to \mathbf{Set}$	A model (interpretation of the theory)	3
Natural transformation $\eta\colon F \Rightarrow G$	A translation between models	4

This is the basic vocabulary of Lawvere's functorial semantics (Lawvere, 1963), transplanted into international relations.

But we have not yet addressed the question that motivates this paper: what determines the identity of a state within a given theory? To answer this, we need to look at a special family of models—those that arise canonically from the theory itself, without any external interpretive choice. These are the *representable functors*, and they are the subject of the next section.

5 The Yoneda Lemma: Identity Is Relational

So far, every model we have considered has been an external interpretation: a functor that assigns data to the theory's actors and relations from the outside. The liberal model reads trade as welfare; the realist model reads trade as power; the Penn World Tables read the system as a collection of disconnected attribute vectors. In each case, the interpretive choice—the functor—comes from the theorist.

But there is a family of models that arise canonically from the theory itself, without any external choice at all. These are the *representable functors*, and they encode the idea that a state's relational profile—the totality of its relationships to every other state—is itself a model. This section introduces the relational profile, states the Yoneda lemma, and draws out its consequences. The Yoneda lemma is the central result of the paper, and its content is this: every datum that any model assigns to a state is not a local fact about that state but a global structural commitment—a complete, coherent reading of that state's position in the system.

5.1 The relational profile

Definition 5.1 (Representable functor)**.** For any object X in a category $\mathcal{C}$, the *representable functor* $\mathrm{Hom}_{\mathcal{C}}(-, X) \colon \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ is defined by:

(i) On objects: $\mathrm{Hom}(-, X)$ sends each object A to the set $\mathrm{Hom}(A, X)$ of all morphisms from A to X.

(ii) On morphisms: given $g \colon B \to A$ in $\mathcal{C}$, the function $\mathrm{Hom}(g, X) \colon \mathrm{Hom}(A, X) \to \mathrm{Hom}(B, X)$ sends $f \mapsto f \circ g$ (precomposition with g).

What does this say? For a fixed state X, the functor $\mathrm{Hom}(-, X)$ assigns to every other state A the set of all relationships that A bears to X. It collects, for the entire system, the answer to the question: "how does the world relate to X?"

This is X's *relational profile*—not an intrinsic property of X, but the complete record of how every other actor in the system is connected to it.[6]

The relational profile is itself a functor to **Set**—a model in the sense of Definition 3.2. But it is a model of a very particular kind. Unlike the liberal or realist models, it makes no interpretive choice about what trade *means*. It simply records, for every state in the system, the full inventory of relationships pointing toward X. It is the most theoretically innocent model available: the model that says "to understand X, look at how the entire world relates to it."

The profile as a working functor. The definition says that $\mathrm{Hom}(-, X)$ acts not only on objects but on morphisms, and this is where the profile acquires its structural force. Consider the trade morphism $g \colon \mathrm{Vietnam} \to \mathrm{China}$ in $\mathcal{T}$. The profile $\mathrm{Hom}(-, \mathrm{USA})$ sends g to the function

$$\mathrm{Hom}(g, \mathrm{USA}) \colon \mathrm{Hom}(\mathrm{China}, \mathrm{USA}) \to \mathrm{Hom}(\mathrm{Vietnam}, \mathrm{USA})$$

that takes each trade channel from China to the USA and extends it backward through g: the direct channel $h \colon \mathrm{China} \to \mathrm{USA}$ maps to the composite $h \circ g \colon \mathrm{Vietnam} \to \mathrm{USA}$. The profile registers that any trade relationship

[6] The notation $\mathcal{C}^{\mathrm{op}}$ indicates the *opposite category*, in which all morphisms are formally reversed. This technical detail ensures that the Hom-functor is covariant (a morphism $g \colon B \to A$ in $\mathcal{C}$ induces a function $\mathrm{Hom}(A, X) \to \mathrm{Hom}(B, X)$ going in the "right" direction). The reader who finds this confusing may safely ignore it; the intuition—$\mathrm{Hom}(-, X)$ collects all relationships pointing toward X—is what matters.

arriving at the USA from China can be pulled back, via the Vietnam–China link, to a trade relationship arriving from Vietnam. This is precomposition, and it is what makes the profile a functor rather than a mere list: it tracks not just which relationships exist but how they propagate through the network.

5.2 Two profiles in the running example

Example 5.2 (The trade profile of the United States)**.** In the four-state fragment of $\mathcal{T}$ (Figure 1), the representable functor $\mathrm{Hom}_{\mathcal{T}}(-, \mathrm{USA})$ assigns:

$$
\begin{aligned}
\mathrm{Hom}(\mathrm{Laos}, \mathrm{USA}) &= \{h \circ g \circ f\}, \\
\mathrm{Hom}(\mathrm{Viet}, \mathrm{USA}) &= \{h \circ g\}, \\
\mathrm{Hom}(\mathrm{China}, \mathrm{USA}) &= \{h\}, \\
\mathrm{Hom}(\mathrm{USA}, \mathrm{USA}) &= \{\mathrm{id}_{\mathrm{USA}}\}.
\end{aligned}
$$

Every state in the fragment has a nonempty set of trade channels pointing toward the USA. The profile is full: the USA is an actor to whom the entire network converges.

Example 5.3 (The trade profile of Laos)**.** The profile $\mathrm{Hom}_{\mathcal{T}}(-, \mathrm{Laos})$ tells a different story:

$$
\begin{aligned}
\mathrm{Hom}(\mathrm{Laos}, \mathrm{Laos}) &= \{\mathrm{id}_{\mathrm{Laos}}\}, \\
\mathrm{Hom}(\mathrm{Viet}, \mathrm{Laos}) &= \varnothing, \\
\mathrm{Hom}(\mathrm{China}, \mathrm{Laos}) &= \varnothing, \\
\mathrm{Hom}(\mathrm{USA}, \mathrm{Laos}) &= \varnothing.
\end{aligned}
$$

No other state in the fragment has a trade channel pointing toward Laos. The only relationship Laos bears to itself is the identity—the domestic economy. The profile is nearly empty: Laos is a source in the trade network, not a sink.

The asymmetry between these two profiles is structural. The USA is an actor to whom many relationships converge; Laos is an actor from whom relationships emanate but to whom few return. The profiles are different because the states *are* different—not in their intrinsic properties (the category says nothing about intrinsic properties) but in how they sit within the relational structure. And the difference is visible at a glance: $\mathrm{Hom}(\mathrm{China}, \mathrm{USA}) = \{h\}$ while $\mathrm{Hom}(\mathrm{China}, \mathrm{Laos}) = \varnothing$.

Example 5.4 (Relational profiles in the alliance category)**.** In the alliance category $\mathcal{A}$, the profile $\mathrm{Hom}_{\mathcal{A}}(-, \mathrm{USA})$ collects all alliance commitments pointing toward the United States: NATO members' commitments, bilateral defense treaties with Japan, South Korea, Australia, the Philippines, and so on. The profile $\mathrm{Hom}_{\mathcal{A}}(-, \mathrm{Switzerland})$ is nearly empty—Switzerland's defining foreign-policy feature is the paucity of alliance morphisms converging on it. In the alliance category, Switzerland and the United States are distinguished precisely by the difference in their relational profiles.

5.3 Reading a profile through a model

A natural transformation $\eta\colon \mathrm{Hom}(-, X) \Rightarrow F$ is the act of *reading* X's relational profile through a model F. For each state A, the component $\eta_A\colon \mathrm{Hom}(A, X) \to F(A)$ translates every relationship that A bears to X into a datum in F's language—coherently, because the naturality condition ties the translations together across the relational structure.

The Yoneda lemma says that every such reading is completely determined by a single element: what F assigns to X itself. One datum at X, and the entire interpretation is fixed.

5.4 The Yoneda lemma

Theorem 5.5 (The Yoneda Lemma)**.** *Let $\mathcal{C}$ be a category, $F\colon \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ a functor, and X an object of $\mathcal{C}$. There is a bijection, natural in both X and F:*

$$\mathrm{Nat}(\mathrm{Hom}_{\mathcal{C}}(-, X),\ F) \ \cong\ F(X).$$

In words: the set of all natural transformations from X's relational profile into any model F is in exact, one-to-one correspondence with what F assigns to X.

Proof sketch. Given a natural transformation $\eta\colon \mathrm{Hom}(-, X) \Rightarrow F$, evaluate its component at X on the identity morphism: $\eta_X(\mathrm{id}_X) \in F(X)$. This defines a map $\mathrm{Nat}(\mathrm{Hom}(-, X), F) \to F(X)$. Conversely, given an element $a \in F(X)$, define $\eta_A(f) = F(f)(a)$ for each $f \in \mathrm{Hom}(A, X)$. Naturality of η follows from functoriality of F, and the two constructions are inverse to each other. □

The proof is short—a few lines of diagram chasing—but the content is deep. The identity morphism id_X acts as a "universal probe": every natural transformation out of $\mathrm{Hom}(-, X)$ is completely determined by what it does to id_X. This is the formal expression of the idea that X's self-relation, combined with the relational structure, pins down everything.

What the bijection means. Every element $a \in F(X)$ is not a local fact about X but a *global structural commitment*: the rule $\eta^a_A(f) = F(f)(a)$ propagates a outward through every relationship converging on X. To say "the USA has welfare level a" is a statement about the *entire trade network as seen from the USA's position*. China's welfare assessment is not a separate parameter; it is $F(h)(a)$, forced by a and the liberal model's transmission function. Vietnam's is $F(g)(F(h)(a))$; Laos's is $F(f)(F(g)(F(h)(a)))$. Each is a consequence of the single datum a, propagated through the relational structure. There are no free parameters left.

Conversely, every coherent reading of X's relational profile through F is generated by some element of $F(X)$, and no two elements generate the same reading. The local datum and the global interpretation are the same object, viewed from different ends.

5.5 The Yoneda lemma at work

Let us see the bijection concretely in the four-state trade category, taking $X = \mathrm{USA}$.

The liberal model. Take F to be the liberal model of Example 3.3. A natural transformation $\eta\colon \mathrm{Hom}(-, \mathrm{USA}) \Rightarrow F$ must assign, for each state A, a function $\eta_A\colon \mathrm{Hom}(A, \mathrm{USA}) \to F(A)$ that is compatible with trade morphisms. The Yoneda lemma says all such transformations are determined by a single element: $\eta_{\mathrm{USA}}(\mathrm{id}_{\mathrm{USA}}) \in F(\mathrm{USA})$.

Pick any welfare outcome $a \in F(\mathrm{USA})$—say, a particular level of consumer welfare. This determines a unique natural transformation η^a by the rule:

$$\eta^a_A(f) = F(f)(a) \quad \text{for each } f \in \mathrm{Hom}(A, \mathrm{USA}).$$

Tracing the consequences:

- $\eta^a_{\mathrm{China}}(h) = F(h)(a)$: the welfare effect on China of its direct trade link to the USA, starting from the USA's welfare level a.

- $\eta^a_{\text{Viet}}(h \circ g) = F(g)(F(h)(a))$: the welfare effect on Vietnam, obtained by first tracing a through China and then through Vietnam.

- $\eta^a_{\text{Laos}}(h \circ g \circ f) = F(f)(F(g)(F(h)(a)))$: the welfare effect on Laos, at the end of the fully mediated chain.

A single welfare datum for the USA generates a complete welfare distribution across the system. A different datum $a' \neq a$ generates a different distribution—different at every state where the transmission functions $F(h)$, $F(g)$, $F(f)$ are not constant. The bijection $\text{Nat}(\text{Hom}(-, \text{USA}), F) \cong F(\text{USA})$ is now visible: the USA's welfare level is not an independent assignment but a compressed summary of the entire system's welfare structure, as seen from the USA's position.

The realist model. The same bijection applies to the realist model G: each power position $b \in G(\text{USA})$ propagates through the trade network via G's transmission functions, distributing power assessments where F distributes welfare—from the same relational profile, through different lenses.

Why the profile of Laos works differently. The Yoneda lemma applies equally to Laos, but the bijection looks very different. Since $\text{Hom}(\text{Viet}, \text{Laos}) = \text{Hom}(\text{China}, \text{Laos}) = \text{Hom}(\text{USA}, \text{Laos}) = \varnothing$, a natural transformation $\eta \colon \text{Hom}(-, \text{Laos}) \Rightarrow F$ has nothing to assign at Vietnam, China, or the USA—the empty set admits only the empty function. The entire natural transformation is determined by $\eta_{\text{Laos}}(\text{id}_{\text{Laos}}) \in F(\text{Laos})$, but it carries no information to any other state. The liberal model's welfare assessment of Laos does not propagate through the system, because no trade channels converge on Laos through which consequences could propagate.

This is the formal expression of a structural fact: in this fragment of the trade network, Laos is a peripheral actor whose relational profile gives models little to work with. The USA's welfare level is a global commitment; Laos's welfare level, in this fragment, is a local island. The difference is not in the states' intrinsic properties but in the richness of their relational profiles—and the Yoneda lemma makes the difference precise.

What datasets lose. The Laos example is a local thinning of the relational profile. The datasets of Section 3.6 represent a global thinning. In the PWT's discrete category (no morphisms other than identities), the Yoneda bijection is vacuous: every state's datum is a free parameter, unconstrained by any other. "The USA has GDP a" really is a local fact, because the theory contains no relational structure through which consequences could propagate. Embed the same states in the trade category, and the same datum becomes a global structural commitment. The COW bilateral trade data sits between these extremes: it records direct morphisms but discards composition, so the Yoneda lemma propagates consequences along direct links but not along chains.

The more morphisms a theory contains, the more each datum is constrained by every other. Attribute tables are the limiting case: maximum freedom, zero relational constraint, and a Yoneda lemma with nothing to say.

5.6 The Yoneda embedding

The Yoneda lemma has an immediate and powerful corollary.

Corollary 5.6 (The Yoneda Embedding)**.** *The functor* $\mathbf{y}\colon \mathcal{C} \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ *defined by* $X \mapsto \mathrm{Hom}(-, X)$ *is* fully faithful*. That is, for any objects* X *and* Y *in* $\mathcal{C}$*:*

$$\mathrm{Hom}_{\mathcal{C}}(X, Y) \;\cong\; \mathrm{Nat}(\mathrm{Hom}(-, X),\ \mathrm{Hom}(-, Y)).$$

In particular, $X \cong Y$ *if and only if* $\mathrm{Hom}(-, X) \cong \mathrm{Hom}(-, Y)$*.*

This is the punchline. Two states are isomorphic—structurally identical within the theory—if and only if they have identical relational profiles. There is no identity beyond relations.

The embedding in the running example. The profiles computed in Examples 5.2 and 5.3 make this concrete. The USA and Laos cannot be isomorphic in $\mathcal{T}$, because their profiles disagree: $\mathrm{Hom}(\mathrm{China}, \mathrm{USA}) = \{h\}$ while $\mathrm{Hom}(\mathrm{China}, \mathrm{Laos}) = \varnothing$. A natural isomorphism $\mathrm{Hom}(-, \mathrm{USA}) \cong \mathrm{Hom}(-, \mathrm{Laos})$ would require, at the component for China, a bijection between a one-element set and the empty set—which does not exist. The Yoneda embedding thus distinguishes the USA from Laos on purely relational grounds: not because they differ in GDP or population or regime type, but because the morphisms converging on them are different.

More subtly, the Yoneda embedding says that the *only* way two states can be isomorphic is if their profiles agree at *every* state in the system. Even a single discrepancy—a single state A for which $\mathrm{Hom}(A, X)$ and $\mathrm{Hom}(A, Y)$

differ—is enough to distinguish them. Identity is overdetermined by relational data: there are as many independent witnesses to the difference between two non-isomorphic states as there are states at which their profiles disagree.

5.7 Philosophical implications

Against essentialism. The Yoneda embedding is a theorem against essentialism. It says that there are no intrinsic properties of a state—no hidden essence—that are not already captured by its relational profile. If two states are relationally identical (in a given theory), they are identical, full stop. There is no room for a residual "nature" or "character" that might distinguish them beyond their relations.

Identity is model-dependent. There is no bare "identity of A"—only what the liberal model assigns to A, what the realist model assigns to A, and so on. $F(A) \neq G(A)$ in general: the welfare profile of the United States is not the same object as its power profile. This much is obvious from the definition of a functor. What is not obvious—and what the Yoneda lemma makes precise—is that every such model-specific identity is extracted from the *same* relational profile $\mathrm{Hom}(-, A)$ via a natural transformation. The liberal model's welfare assessment $F(A)$ is obtained by mapping A's relational profile into F; the realist model's power assessment $G(A)$ is obtained by mapping the same profile into G. Different models yield different identities, but none of them is arbitrary: each is a projection of A's relational position through an interpretive lens that must respect the relational structure at every point. Identity is not only relational but perspectival—and the perspectives are disciplined by naturality.

For constructivism. Wendt's central claim—that "identities are constituted by interaction" (Wendt, 1992)—is now a theorem, not a philosophical stance. The constructivist insight that "identity constitutes interests" (Hopf, 1998) receives a precise formalization: a state's model-specific data (which determines its behavior under that model) is a projection of its relational profile, extracted via a natural transformation. The Yoneda embedding proves that within any relational theory, the identity of an actor is constituted by its relationships to all other actors. This does not settle the broader philosophical debate (the Yoneda lemma is a mathematical result, not a metaphysical one), but it shows that the constructivist claim is the *correct structural consequence* of taking relations seriously. Any framework that foregrounds relational structure and satisfies the minimal axioms of a category will, as a matter of mathematical necessity, yield the conclusion that identity is relational. Empirically, this is not idle: Renshon (2017) measures diplomatic status via PageRank centrality on the network of embassies—in effect, computing a representable functor—and shows that the gap between this relational identity and attribute-based expectations predicts conflict initiation.

Against the like-units assumption. Waltz's structural realism treats states as "like units"—functionally undifferentiated actors whose differences lie only in capabilities (Waltz, 1979). The Yoneda embedding says the opposite: states are distinguished *precisely* by their relational profiles. Note that the framework has heeded Waltz's own advice and ignored all within-unit variation—objects in a category have no internal structure, no properties, no hidden essence. And yet the Yoneda embedding differentiates them anyway. Even after stripping states of all intrinsic content—granting Waltz everything he asks for—the relational structure between them suffices to individuate them. The United States and Laos are not like units that happen to differ in GDP; they are fundamentally different objects in the trade category because the morphisms converging on them are different. Waltz was right that structure is primary, but wrong to conclude that structure leaves units interchangeable. The structure itself differentiates them.

States as canonical points in the space of models. The Yoneda embedding $\mathbf{y}\colon \mathcal{C} \hookrightarrow [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ maps each state to its relational profile, which is itself a model—a canonical, distinguished point in the vast space of all models. A state is not just an object that models describe; it is a *universal model of itself*. Every other model's assessment of X is obtained from $\mathrm{Hom}(-, X)$ via a natural transformation; so $\mathrm{Hom}(-, X)$ is the most complete model of X available within the theory, and every model extracts from it by projection.

This inverts a familiar picture. The USA—understood as the representable functor $\mathrm{Hom}(-, \mathrm{USA})$—is itself a model, and the richest model of itself that the theory can produce. The liberal model does not describe the USA from the outside; it *projects* the USA's relational self-description into a particular vocabulary. The state is not the described; it is the source from which all descriptions are derived.

6 Comparing Theories

The previous sections developed the framework within a single theory: a category $\mathcal{C}$, its models (functors to **Set**), translations between models (natural transformations), and the Yoneda lemma that pins identity to relational profile. But international relations does not consist of a single theory. The trade category $\mathcal{T}$, the alliance category $\mathcal{A}$, and a recognition category are three different theories of the same system of states. They share the same actors but differ in the relations they foreground. How do we compare them?

6.1 Functors between theories

The same tool that gives us models—the functor—also gives us theory translations.

Definition 6.1 (Theory translation)**.** A *theory translation* from $\mathcal{C}$ to $\mathcal{D}$ is a functor $P\colon \mathcal{C} \to \mathcal{D}$: an assignment of objects to objects and morphisms to morphisms that preserves composition and identity.

The definition is identical to Definition 3.1, but the conceptual shift is important. When the target is **Set**, a functor interprets a theory in the world of data. When the target is another theory, a functor translates one relational vocabulary into another.

Example 6.2 (From trade to alliance)**.** Suppose that major trade partnerships tend to generate security commitments—that states which trade heavily develop overlapping interests that eventually crystallize into alliance obligations. This hypothesis can be expressed as a functor $P\colon \mathcal{T} \to \mathcal{A}$ that maps each state to itself (the actors are the same) and each trade morphism to the alliance commitment it is hypothesized to induce. The functorial axioms impose constraints on this hypothesis: if the trade link from Laos to Vietnam and the link from Vietnam to China together generate a composite alliance commitment from Laos to China, then P must map the composite trade morphism $g \circ f$ to the composite alliance morphism $P(g) \circ P(f)$—the induced commitment must respect the chain of relationships, not just the bilateral links.

Not every hypothesized connection between theories will satisfy these constraints.

Example 6.3 (Trade-to-conflict in the running example)**.** Define a conflict category $\mathcal{D}$ on the same four states. Let $\mathcal{D}$ contain a rivalry morphism $r\colon \text{China} \to \text{USA}$ and, via composition, a mediated tension morphism from Vietnam to the USA and from Laos to the USA. A liberal functor $P\colon \mathcal{T} \to \mathcal{D}$ mapping each trade morphism to the conflict relationship it is hypothesized to dampen must respect composition: the composite trade chain Laos $\to$ Vietnam $\to$ China $\to$ USA must map to the composite conflict chain $P(h) \circ P(g) \circ P(f)$. If the liberal hypothesis holds that bilateral trade dampens bilateral conflict but the dampening effects do not compose—the chain of individually pacifying links does not yield a pacifying chain—then no such functor exists, and the hypothesis fails the compositional test.

The failure is informative: it says that the two theories' relational structures are incompatible in a specific way. The trade category composes transitively by assumption; if the hypothesized conflict effects do not, the translation breaks.

6.2 Pulling back models

Theory translations do more than relate relational structures. They transport models.

If $P\colon \mathcal{C} \to \mathcal{D}$ is a functor between theories and $F\colon \mathcal{D} \to \mathbf{Set}$ is a model of $\mathcal{D}$, then the composite $F \circ P\colon \mathcal{C} \to \mathbf{Set}$ is a model of $\mathcal{C}$. It assigns to each object of $\mathcal{C}$ whatever F assigns to its image in $\mathcal{D}$, and to each morphism the corresponding function. The composite is called the *pullback* of F along P.

This is how cross-theory interpretation works. A conflict model $F\colon \mathcal{D} \to \mathbf{Set}$ assigns conflict probabilities to states and conflict-transmission functions to the relations in a conflict category $\mathcal{D}$. If a functor $P\colon \mathcal{T} \to \mathcal{D}$ translates trade relations into conflict relations, then $F \circ P$ is a model of the trade category that reads trade as conflict: it assigns to each state the conflict probability induced by its trade position, and to each trade morphism the corresponding conflict-transmission function.

6.3 Does trade cause peace?

Section 2.4 observed that the question "does trade cause peace?" does not live inside a single category. We can now say precisely where it lives.

The question requires three ingredients:

(i) A trade category $\mathcal{T}$ and a conflict category $\mathcal{D}$ (two theories of the same system of states).

(ii) A functor $P \colon \mathcal{T} \to \mathcal{D}$ that translates trade relations into conflict relations (a specific causal hypothesis).

(iii) A model $F \colon \mathcal{D} \to \mathbf{Set}$ that interprets conflict relations as data—probability of armed dispute, crisis escalation, or whatever the analyst measures.

The composite $F \circ P \colon \mathcal{T} \to \mathbf{Set}$ is then a model of trade that reads trade as conflict. Different functors P encode different causal stories. The liberal hypothesis that trade promotes peace corresponds to one P—one that maps trade morphisms to conflict-dampening relationships. The realist hypothesis that trade dependence generates vulnerability corresponds to a different P—one that maps trade morphisms to conflict-exacerbating relationships.

The two hypotheses are not merely different verbal accounts. They are different functors, and they generate different pullback models $F \circ P_{\text{lib}}$ and $F \circ P_{\text{real}}$ of the trade category. Whether a natural transformation exists between these pullback models—whether the liberal and realist readings of trade-as-conflict are coherently translatable—is a determinate structural question, answerable by the methods of Section 4.

The framework does not answer the empirical question of which functor P is correct. It clarifies what kind of object the question is about: a functor between theories, subject to the compositional constraints of the source and target categories.

6.4 Reducibility and embedding

The properties of a theory translation $P \colon \mathcal{C} \to \mathcal{D}$ characterize the structural relationship between the two theories.

- P is *faithful* if it preserves distinctions: whenever $\mathcal{C}$ has two different morphisms $f \neq g$ between the same pair of objects, $P(f) \neq P(g)$ in

$\mathcal{D}$. A faithful translation says that the target theory can express every relational distinction that the source makes.

- P is *full* if the source theory captures every relation in the target: for any morphism $h\colon P(A) \to P(B)$ in $\mathcal{D}$, there exists a morphism $f\colon A \to B$ in $\mathcal{C}$ with $P(f) = h$. A full translation says that the source theory is relationally complete with respect to the target.

- P is *fully faithful* if both conditions hold. This is an *embedding*: the source theory is a sub-theory of the target, with no relational information lost or gained.

These distinctions bear on long-standing debates about theory reduction in IR. Can the alliance category be reduced to the trade category—do all alliance commitments arise from trade relationships? This is the claim that a full functor $P\colon \mathcal{T} \to \mathcal{A}$ exists. Can trade relationships be distinguished within the alliance framework—do alliance commitments preserve the fine-grained structure of trade? This is the claim that P is faithful.

The answers are empirical, but the framework makes the structural content of the claims precise. A claim of reducibility is a claim about a specific functor with specific properties, not a vague assertion that "everything is really about trade" or "alliances subsume trade."

These tools handle theories that share the same actors and differ in their relations. More complex comparisons—theories with different actor sets, or the question of complementarity rather than translation—require richer structure (adjunctions, fibered categories) that lies beyond the present paper but within the categorical framework. Section 7.3 sketches these directions.

7 Discussion

The preceding sections have built a framework in which theories are categories, models are functors, translations between models are natural transformations, and the Yoneda lemma proves that identity is relational. This section takes stock: what does the framework buy, what does it not buy, and where might it go?

7.1 Relation to existing IR metatheory

Wendt and constructivism. The paper's central result—that identity is constituted by relational profile—vindicates the core claim of the constructivist research program (Adler, 1997; Wendt, 1992; Finnemore and Sikkink, 2001) in a precise sense. "Anarchy is what states make of it" (Wendt, 1992) can now be read as a statement about models: the structure of the international system (the category) underdetermines its meaning, which is supplied by the choice of functor. "Identities are constituted by interaction" (Wendt, 1999) is the Yoneda embedding applied to social categories. Finnemore and Sikkink (2001) characterized constructivism as a "framework, not a substantive theory"; the distinction maps precisely onto ours: a category is a framework (a structural signature), while a functor is a substantive theory (an interpretation). Checkel (1998) argued that constructivism's central challenge is to develop theory, not merely demonstrate that social construction matters. The Yoneda lemma is one response: it shows that the constructivist insight follows as a theorem from minimal structural axioms, not from philosophical argument alone.

But the framework also clarifies where Wendt's argument needed sharpening. Wendt's constructivism is often criticized for vagueness about what "constitution" means and how it differs from causal influence (Jackson, 2011). The Yoneda lemma gives "constitution" a precise content: X's identity is not *caused* by its relational profile; it *is* its relational profile, in the sense that the two are in natural bijection. The distinction between constitutive and causal claims, which has generated extensive philosophical debate in IR (Wendt, 1998), is here a distinction between the Yoneda embedding (constitutive: identity is profile) and a functor between categories (causal: one relational structure produces another, as in Section 6.3).

Waltz, Lake, and the structure of the international system. Section 5.7 showed that the Yoneda embedding contradicts Waltz's like-units assumption while vindicating his insistence that structure is primary (Waltz, 1979). Lake (2009) pushes back on the anarchy assumption from the opposite direction, arguing that international relations exhibit varying degrees of hierarchy, with authority constituted relationally through a social contract between dominant and subordinate states. In the present framework, Lake's claim amounts to the observation that the international category possesses non-trivial authority morphisms—relations of legitimate command and compliance—that the anarchy assumption erases. The richer this hierarchical structure, the more information the Yoneda embedding carries about each state's identity.

Jackson and metatheory. Jackson (2011) argues that IR scholars operate within distinct "philosophical ontologies" that shape what counts as knowledge. The framework developed here does not resolve Jackson's pluralism, but it gives it a formal counterpart. Different theories (categories) reflect different ontological commitments about which relations are primitive. Different models (functors) on the same theory reflect different interpretive commitments about what those relations mean. The question of commensurability between paradigms—which Jackson treats as a philosophical question about the compatibility of ontologies—becomes a mathematical question about the existence of natural transformations (Section 4.3) or functors between categories (Section 6).

This does not make the philosophical questions disappear. The choice of which category to work in—whether to foreground trade, alliance, recognition, or some other relational type—remains a substantive theoretical commitment that the mathematics cannot adjudicate. But once that choice is made, the framework constrains what can coherently be said within it, and the Yoneda lemma is the tightest of those constraints.

The analytic eclecticism advocated by Sil and Katzenstein (2010)—drawing on constructs from multiple research traditions to address complex substantive problems—finds formal support here. Different research traditions are different categories; the existence or non-existence of natural transformations between their models makes the commensurability question precise rather than interminable.

Kurki and causal pluralism. Kurki (2008) argues that IR should embrace a pluralism of causal concepts beyond Humean regularity. The framework accommodates this: Humean causation corresponds to models on discrete categories (attribute-table data), while structural and constitutive causation correspond to models on richer categories where the Yoneda lemma propagates consequences through relational structure.

7.2 Limitations

The framework has clear boundaries, and honesty about them is important.

Structure, not dynamics. Categories, as presented here, are static: they encode a fixed set of actors and relations. The international system is not static. States enter and exit, relations form and dissolve, and the relational structure itself changes over time. The framework captures the structure of the system at a given moment (or under a given theoretical idealization) but does not model the processes by which that structure evolves.

No agency. The framework says nothing about intentionality, strategic choice, or the mechanisms by which actors create and sustain relationships. The bargaining models of Section 3.5 are included as examples of multi-sorted theories, but the equilibrium analysis sits inside the functor—inside the model—not in the category itself. The category records that a contest morphism exists; it does not model the decision-making process that determines its outcome.

Theory-dependence. The Yoneda lemma proves that identity is relational *within a given theory*. A state's identity in the trade category may differ from its identity in the alliance category, because the two categories carry different relational structures. This is the formal counterpart of the sensitivity documented by Zhukov and Stewart (2013): different relational specifications yield different identities precisely because they are different categories. The framework does not eliminate this sensitivity; it forces the analyst to announce the theory-dependence of the identities assigned, rather than leaving it implicit in a choice of dataset or network specification. For a constructivist, the theory-dependence of identity is the point. For a positivist seeking a view from nowhere, it may be unsatisfying—but the unsatisfying conclusion is a theorem, not an evasion.

7.3 Extensions

Several directions suggest themselves for future work.

Enriched categories. The categories here are "ordinary"—morphisms either exist or they don't. *Enriched categories*, in which hom-sets are replaced by objects in a monoidal category (e.g., non-negative reals, probability distributions), can accommodate weighted relations. An enriched Yoneda lemma would yield a weighted relational-identity claim: identity is not just the pattern of relations but their intensities.

Higher categories. *2-categories* formalize relationships between relationships: morphisms between morphisms can encode renegotiations of treaties or revisions of trade agreements. This is one route to addressing the dynamics limitation, modeling the evolution of relational structure as higher-dimensional morphisms.

Sheaves. The Yoneda embedding maps a category into its category of presheaves $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$. When the category carries a notion of "covering"—a specification of which collections of relationships jointly determine a global property—the relevant models are not arbitrary presheaves but *sheaves*: presheaves satisfying a local-to-global consistency condition. Sheaf theory could formalize the idea that a state's identity is determined not by its bilateral relationships individually but by coherent combinations of relationships that "cover" the state's position in the system.

Dependent type theory. Dependent type theory, which allows types to depend on terms, corresponds to categories with richer internal structure (locally cartesian closed categories). An IR theory formalized in dependent type theory could express claims like "the nature of state A's alliance with B depends on the specific trade relationship between them"—dependencies between relations that ordinary categories cannot capture.

8 Conclusion

This paper asked what follows, formally, from the commitment to a relational ontology in international relations. The answer: a state's identity, within any relational theory, is nothing more and nothing less than its complete relational profile—the totality of relationships that every other actor in the system bears to it.

The argument proceeded in layers. A theory of international relations is a category: a collection of actors and typed relations, with composition and identity (Section 2). A model is a functor to **Set**: a structure-preserving interpretation that assigns data to actors and functions to relations (Section 3). A translation between models is a natural transformation: a systematic conversion of one interpretation into another that respects the relational structure at every point (Section 4). The Yoneda lemma (Section 5) then proved that every datum a model assigns to a state is a global structural commitment—a complete reading of the state's relational position, compressed into a single element—and the Yoneda embedding proved that two states are structurally identical if and only if their relational profiles agree.

These are not philosophical claims. They are theorems, valid in any category. The constructivist insight that identity is constituted by interaction, the structuralist insight that position matters, and the metatheoretic demand for precise conditions of commensurability between paradigms all find formal expression in the same framework.

Along the way, the framework delivered concrete returns: a sharp constraint on when the Fearon and Beví́á–Corchón bargaining models are commensurable (Section 4.5), a precise account of what is lost when relational structure is flattened into attribute tables (Section 5.5), and a determinate formalization of "does trade cause peace?" as a functor between categories (Section 6.3).

The framework has limitations. It is static, not dynamic. It says nothing about agency or intentionality. It makes identity theory-dependent—a feature for constructivists, a cost for those seeking theory-independent foundations. Extensions to enriched categories, higher categories, and dependent type theory may address some of these limitations; others may require fundamentally different tools.

But the core result stands without qualification. In any theory that specifies actors and relations and satisfies the minimal axioms of composition, identity, and associativity, the Yoneda lemma guarantees that identity is relational. No hidden essence, no intrinsic property, no attribute not already encoded in the relational profile can distinguish two actors whose profiles agree. The state is its relations.

References

Adler, E. (1997). Seizing the middle ground: Constructivism in world politics. *European Journal of International Relations*, 3(3):319–362.

Beckfield, J. (2010). The social structure of the world polity. *American Journal of Sociology*, 115(4):1018–1068.

Beviá, C. and Corchón, L. C. (2010). Peace agreements without commitment. *Games and Economic Behavior*, 68(2):469–487.

Checkel, J. T. (1998). The constructivist turn in international relations theory. *World Politics*, 50(2):324–348.

Cranmer, S. J., Desmarais, B. A., and Menninga, E. J. (2012). Complex dependencies in the alliance network. *Conflict Management and Peace Science*, 29(3):279–313.

Duque, M. G. (2018). Recognizing international status: A relational approach. *International Studies Quarterly*, 62(3):577–592.

Emirbayer, M. (1997). Manifesto for a relational sociology. *American Journal of Sociology*, 103(2):281–317.

Fearon, J. D. (1995). Rationalist explanations for war. *International Organization*, 49(3):379–414.

Finnemore, M. and Sikkink, K. (2001). Taking stock: The constructivist research program in international relations and comparative politics. *Annual Review of Political Science*, 4:391–416.

Fong, B. and Spivak, D. I. (2019). *An Invitation to Applied Category Theory: Seven Sketches in Compositionality*. Cambridge University Press, Cambridge.

Hafner-Burton, E. M., Kahler, M., and Montgomery, A. H. (2009). Network analysis for international relations. *International Organization*, 63(3):559–592.

Hoff, P. D. and Ward, M. D. (2004). Modeling dependencies in international relations networks. *Political Analysis*, 12(2):160–175.

Hopf, T. (1998). The promise of constructivism in international relations theory. *International Security*, 23(1):171–200.

Jackson, P. T. (2011). *The Conduct of Inquiry in International Relations: Philosophy of Science and Its Implications for the Study of World Politics*. Routledge, London.

Kurki, M. (2008). *Causation in International Relations: Reclaiming Causal Analysis*. Cambridge University Press, Cambridge.

Lake, D. A. (2009). *Hierarchy in International Relations*. Cornell University Press, Ithaca, NY.

Lawvere, F. W. (1963). Functorial semantics of algebraic theories. *Proceedings of the National Academy of Sciences*, 50(5):869–872.

Mac Lane, S. (1998). *Categories for the Working Mathematician*. Springer, New York, 2nd edition.

Maoz, Z. (2012). How network analysis can inform the study of international relations. *Conflict Management and Peace Science*, 29(3):247–256.

Monteiro, N. P. and Ruby, K. G. (2009). IR and the false promise of philosophical foundations. *International Theory*, 1(1):15–48.

Nexon, D. H. (2009). *The Struggle for Power in Early Modern Europe: Religious Conflict, Dynastic Empires, and International Change*. Princeton University Press, Princeton, NJ.

Renshon, J. (2017). *Fighting for Status: Hierarchy and Conflict in World Politics*. Princeton University Press, Princeton, NJ.

Ruggie, J. G. (1998). *Constructing the World Polity: Essays on International Institutionalization*. Routledge, London.

Sil, R. and Katzenstein, P. J. (2010). Analytical eclecticism in the study of world politics. *Perspectives on Politics*, 8(2):411–431.

Waltz, K. N. (1979). *Theory of International Politics*. Addison-Wesley, Reading, MA.

Ward, M. D., Stovel, K., and Sacks, A. (2011). Network analysis and political science. *Annual Review of Political Science*, 14:245–264.

Wendt, A. (1992). Anarchy is what states make of it: The social construction of power politics. *International Organization*, 46(2):391–425.

Wendt, A. (1998). On constitution and causation in international relations. *Review of International Studies*, 24(special issue):101–117.

Wendt, A. (1999). *Social Theory of International Politics*. Cambridge University Press, Cambridge.

Zhukov, Y. M. and Stewart, B. M. (2013). Choosing your neighbors: Networks of diffusion in international relations. *International Studies Quarterly*, 57(2):271–287.