Structuring Relations Among Learning Paradigms via Protocol–Objective–Resource Reductions

Anonymous authors
Paper under double-blind review

Abstract

Modern machine learning increasingly spans supervised, transfer, continual, meta-learning, and other training regimes, often while reusing the same hypothesis families, architectures, and optimizers. Yet these regimes can differ substantially in information access, success criteria, memory assumptions, adaptation protocols, and sample accounting. As a result, it is often unclear whether a proposed paradigm is genuinely distinct, a special case of an existing paradigm, or part of a broader structural relationship among paradigms. This motivates a structural theory for comparing learning paradigms while separating representational capacity from protocol, objective, and resource choices. We develop a protocol–objective–resource (POR) framework for making such comparisons precise. In this framework, a learning paradigm is specified by its environment class, observation protocol, admissible learner class, performance functional, and resource accounting rule. We define POR reductions through environment embeddings, learner compilers, threshold maps, and calibrated sample-budget overheads. Our main theorem shows that a POR reduction implies worst-case sample-complexity domination on embedded comparison classes, transferring upper bounds in one direction and lower bounds in the other. We further provide reusable templates for proving reductions, show why explicit accuracy regimes are necessary to avoid vacuous comparisons, and prove that strengthening the objective can strictly increase minimax sample complexity even when the protocol and learner class remain unchanged. We instantiate the framework on supervised, transfer, continual, and meta-learning abstractions. Under explicit closure assumptions, the framework yields exact reductions showing that continual learning contains transfer learning as a special case, transfer learning contains supervised learning as a special case, and meta-learning also contains supervised learning as a special case. The same formalism captures monotone refinements within a paradigm, for example as replay memory or task identifiers vary in continual learning. Together, these results provide a principled structure for relating learning paradigms and enable the formal transfer of understanding and methods across them.

1 Introduction

Reductions are among the most powerful organizing principles in computer science. The theory of NP-completeness showed that apparently different computational problems can be related systematically: an instance of one problem can be encoded as an instance of another, so that a solver for the latter yields a solver for the former (Cook, 1971; KARP, 1972; Garey & Johnson, 1979; Papadimitriou, 1994; Arora & Barak, 2009). This idea did more than classify individual problems as hard. It revealed a network of relationships among difficult problems, allowing algorithms, proof techniques, lower bounds, and impossibility results to move across problem formulations rather than being rebuilt from scratch each time.

This structural perspective also protects research resources. When two formulations reduce to one another, progress on one formulation can be reused for the other; when they are separated, the separation identifies what is genuinely new. Reductions therefore help a field avoid spending repeated effort on the same underlying problem under different names. They turn informal similarity into explicit obligations, and they clarify which methods and theorems should transfer.

Modern machine learning increasingly needs an analogous language. The field is organized around supervised learning, transfer learning, continual learning, meta-learning, and many variants of these regimes (Caruana, 1997; Thrun; Pan & Yang, 2010; Ben-David et al., 2010; Weiss et al., 2016; Parisi et al., 2019; De Lange et al., 2022; Van de Ven et al., 2022; Baxter, 2000; Finn et al., 2017; Hospedales et al., 2021). These labels are useful, but they can hide distinct sources of difficulty. Different paradigms often reuse the same hypothesis families, representations, architectures, losses, and optimizers (Bengio et al., 2013; Caruana, 1997; Crammer et al., 2008; Maurer et al., 2016; Pentina & Lampert, 2015; Tripuraneni et al., 2020), while differing in what data are revealed, when the data are revealed, what counts as success, how memory and adaptation are constrained, and how samples are charged. A fixed predictor architecture, for example, may be trained from scratch, adapted from source tasks, updated sequentially across tasks, or used inside an episodic meta-learning procedure.

Without a structural theory, it is difficult to tell whether a proposed learning paradigm is genuinely new, a special case of an existing paradigm, or a variant whose apparent novelty comes from a change in protocol, objective, or resource accounting. This ambiguity can waste research effort: a method may be developed for a new label even though its guarantee follows from an existing setting, or a benchmark difference may be attributed to a new paradigm when it is actually caused by extra side information, a weaker success criterion, or uncounted resources. It also makes method-transfer claims hard to interpret. Saying that a method transfers from one paradigm to another is not only a statement about using the same model class; it is also a claim that the interaction protocol, success criterion, and resource budget can be matched.

Existing learning-theoretic complexity measures do not resolve this issue. Tools such as VC dimension, Rademacher complexity, covering numbers, stability, and norm-based capacity bounds compare hypothesis classes under a fixed learning setup (Vapnik & Chervonenkis, 1971; Valiant, 1984; Blumer et al., 1989; Vapnik, 1998; Anthony & Bartlett, 1999; Shalev-Shwartz & Ben-David, 2014; Mohri et al., 2018; Bartlett & Mendelson, 2002; Bousquet & Elisseeff, 2002). They are not designed to separate the effect of representational capacity from the effect of changing the learner's information interface, objective, memory constraint, adaptation protocol, or sample-accounting rule. Consequently, a statement such as "continual learning is harder than transfer learning" is formally incomplete. The claimed gap might come from the need to control forgetting, the absence of task identifiers, a stronger all-task objective, a smaller replay budget, or a different convention for counting source and target samples (Robins, 1995; French, 1999; Kirkpatrick et al., 2017; Lopez-Paz & Ranzato, 2017; Rebuffi et al., 2017; De Lange et al., 2022; Van de Ven et al., 2022; Kalan et al., 2020). This paper asks the following question:

> *What exactly makes one learning paradigm harder than another, and under what conditions can understanding and methods be systematically transferred across paradigms?*

We answer this question by developing a reduction-based framework for comparing learning paradigms. The analogy is with reductions in complexity theory, but the objects are statistical learning interfaces rather than decision problems. A learning paradigm is represented as a protocol–objective–resource (POR) specification: it consists of an environment class, an observation protocol, an admissible learner class, a performance functional, and a resource accounting rule. A POR reduction certifies that any successful learner for one paradigm can be systematically compiled into a successful learner for another, after translating environments, success requirements, and sample budgets.

The practical role of the framework is to separate representation from the rest of the learning interface. Shared architectures or optimizers do not by themselves imply that understanding or methods transfer. For understanding transfer, a reduction states when upper bounds, lower bounds, and impossibility results in one paradigm imply corresponding results in another. For method transfer, the same reduction gives a checklist: the target setting must be able to reproduce the source method's data transcript, source success must imply target success after the threshold translation, and the simulation must not hide extra sample, memory, replay, or adaptation cost. This makes informal claims of transfer auditable rather than merely intuitive.

Contributions. The main contributions of this paper are as follows.

- **A reduction framework for learning paradigms.** Section 2 formalizes learning paradigms as POR objects and defines POR reductions. The definition records four obligations: environments must be embedded, learners must be compiled, success requirements must be translated, and sample budgets must be calibrated. Accuracy regimes are included so that comparisons cannot be made meaningful only by choosing vacuously loose thresholds. Collectively, these elements establish a reduction theory for machine learning paradigms, analogous in spirit to the role of reductions in complexity theory.

- **Consequence and structural properties of POR reduction.** Section 3 develops the main theoretical consequences of the POR reduction framework. Proposition 3.1 characterizes how sample budgets are calibrated when a simulation consumes a bounded number of observations. Theorem 3.3 establishes the central result of the framework: a POR reduction implies minimax sample-complexity domination on embedded comparison classes. Building on this theorem, Corollaries 3.4, 3.5, and 3.6 provide lower-bound transfer, exact-domination results, and a sample-complexity obstruction for ruling out proposed reductions. Beyond complexity transfer, we establish several structural properties of the framework. We prove monotonicity with respect to thresholds and comparison classes (Propositions 3.7 and 3.8), show that POR hardness forms a preorder whose induced equivalence classes capture mutual reducibility (Proposition 3.9 and Corollary 3.10), and derive reusable proof templates based on special-case embeddings, objective strengthening, additional information, and enlarged learner resources (Proposition 3.11). Finally, Propositions 3.12 and 3.13 demonstrate that modifying the learning objective alone can strictly increase minimax sample complexity, even when the environment class, observation protocol, learner class, hypothesis family, and resource accounting rule remain unchanged. Together, these results establish POR as a mathematically structured reduction theory and clarify how protocol, objective, and resource choices shape the relative difficulty of learning paradigms.

- **POR Structure among representative learning paradigms.** Section 4 instantiates the POR framework on abstractions of supervised, transfer, continual, and meta-learning. Under explicit closure assumptions, we establish the exact reduction relations

  $$\text{Cont} \succeq_{\text{ex}} \text{Trans} \succeq_{\text{ex}} \text{Sup}, \qquad \text{Meta} \succeq_{\text{ex}} \text{Sup},$$

  where ($A \succeq_{\text{ex}} B$) denotes that (B) exact-reduces to (A). These results formally identify transfer learning as a special case of continual learning, supervised learning as a special case of transfer learning, and supervised learning as a special case of meta-learning. The reductions are established in Propositions 4.4–4.7, and their sample-complexity consequences are summarized in Corollary 4.9. To demonstrate the practical utility of this structure, we present a case study relating transfer learning and continual learning. Using the reduction framework, we show how impossibility and lower-bound results established for transfer learning can be transferred directly to continual learning without requiring a separate analysis. This illustrates how POR reductions enable the systematic transfer of theoretical understanding across learning paradigms.

Scope. The reductions studied here are structural worst-case statements. They do not claim that optimization difficulty, constants, empirical benchmark behavior, or typical-case performance are identical across paradigms. They also do not impose a universal total order over supervised, transfer, continual, and meta-learning. Instead, the framework identifies the precise assumptions under which one learning interface contains another for minimax sample-complexity purposes, and it explains which obligation fails when such a comparison cannot be justified.

2 Preliminaries: POR Objects and Reductions

We use $\mathbb{N} = \{0, 1, 2, \dots\}$, and log denotes the natural logarithm. All paradigms compared in a reduction share an input space $\mathcal{X}$, output space $\mathcal{Y}$, base loss $\ell : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}_+$, and hypothesis family

$$\mathcal{H} = \{h_\theta : \theta \in \Theta\}.$$

For a distribution D over $\mathcal{X} \times \mathcal{Y}$, write

$$R_D(h) := \mathbb{E}_{(x,y)\sim D}[\ell(h(x), y)].$$

The shared $\mathcal{H}$ assumption isolates protocol, objective, and resource effects from representational capacity. The default resource unit is one revealed labeled example, unless an explicit conversion rule is stated.

2.1 POR objects

A learning paradigm over $\mathcal{H}$ is a tuple

$$P = (\mathcal{E}_P, \mathcal{O}_P, \mathfrak{T}_P, \mathbf{J}_P, \kappa_P).$$

Here $\mathcal{E}_P$ is the environment class; $\mathcal{O}_P$ is the observation protocol; $\mathfrak{T}_P$ is the admissible learner class; $\mathbf{J}_P$ is the performance functional; and κ_P is the resource accounting rule. Environments may be single distributions, source and target task tuples, sequential task streams, or task distributions for episodic meta-learning. The protocol specifies what information is revealed, and $\mathfrak{T}_P$ encodes constraints such as memory, replay, or adaptation depth.

Example: supervised learning. For ordinary supervised learning, $\mathcal{E}_P$ may be a class of distributions D on $\mathcal{X} \times \mathcal{Y}$; $\mathcal{O}_P$ reveals i.i.d. labeled samples from D; $\mathfrak{T}_P$ contains the training procedures allowed to output a predictor $h \in \mathcal{H}$; the performance functional can be the scalar risk

$$\mathbf{J}_P(T, D; M) = R_D(h_T(S_M)), \qquad S_M \sim D^M,$$

where D is the environment, S_M is the random training sample generated by the protocol, and $h_T(S_M)$ is the predictor output by T. Thus $\mathbf{J}_P(T, D; M)$ is random through the sampled dataset S_M, but its explicit environment argument is the distribution D. The resource rule κ_P counts the number of labeled examples used.

Description and purpose. This definition describes a learning paradigm as the full interface between an environment, a learner, an evaluation rule, and a resource model. We need this object because the paper compares paradigms while holding the hypothesis family fixed; the tuple makes the protocol, objective, learner constraints, and accounting convention explicit rather than hiding them behind a paradigm label.

Assumption 2.1 (Budget capping)**.** Learner classes are closed under budget capping. For any admissible learner $T \in \mathfrak{T}_P$ and any budget-capping rule $q : \mathbb{N} \to \mathbb{N}$ satisfying $q(M) \le M$, define T^q to be the learner that, under external budget M, runs T with internal budget $q(M)$ and ignores the remaining $M - q(M)$ observations. Then $T^q \in \mathfrak{T}_P$, and for every environment $e \in \mathcal{E}_P$ and budget $M \in \mathbb{N}$,

$$\mathbf{J}_P(T^q, e; M) = \mathbf{J}_P(T, e; q(M))$$

as random variables.

Description and reasonableness. This assumption states that a sample budget is an upper bound: a valid learner may use fewer than the available observations and ignore the rest. This is reasonable because most sample-complexity guarantees are stated in terms of "at most M" samples rather than an obligation to consume exactly M samples. The assumption is also needed for resource calibration across paradigms. If a simulation uses only $\psi(M) \le aM + b$ samples while the reduction grants budget $aM + b$, the compiled learner must be allowed to run the simulation on $\psi(M)$ samples and discard the rest. Without this closure, reductions would depend on the unnatural distinction between "exactly M samples" and "at most M samples."

The performance functional is an $\mathbb{R}^{d_P}$-valued random vector $\mathbf{J}_P(T, e; M)$, whose law includes environment, sampling, protocol, and learner randomness. Its coordinates may include risk, regret, forgetting, or post-adaptation query error. For $u, v \in \mathbb{R}^d$, write $u \preceq v$ when $u_i \le v_i$ for every coordinate. We assume all events $\{\mathbf{J}_P(T, e; M) \preceq \tau\}$ are measurable.

2.2 Success and minimax sample complexity

An accuracy regime for P is a set $\mathscr{T}_P \subseteq \mathbb{R}^{d_P}$ of meaningful threshold vectors. Regimes are part of every reduction statement. For a reduction $B \preceq_{\mathrm{POR}} A$, the threshold map sends thresholds for the solver paradigm A to thresholds for the reduced paradigm B. Regimes also prevent vacuous reductions that map every requirement to an automatically satisfied threshold.

Description and purpose. This definition describes the set of accuracy thresholds that are considered meaningful for a paradigm. We need the regime because reductions compare success requirements, and without restricting the thresholds one could obtain formally correct but vacuous comparisons by mapping every requirement to an automatically satisfied target.

For $T \in \mathfrak{T}_P$, environment e, threshold τ_P, and confidence $\delta \in (0,1)$, define

$$N_P(T, e; \tau_P, \delta) := \inf\Big\{ M \in \mathbb{N} : \Pr[\mathbf{J}_P(T, e; M) \preceq \tau_P] \geq 1 - \delta \Big\},$$

with $\inf \emptyset = \infty$. A comparison class is a nonempty subset $\mathcal{C}_P \subseteq \mathcal{E}_P$, and the minimax sample complexity is

$$N_P^\star(\mathcal{C}_P; \tau_P, \delta) := \inf_{T \in \mathfrak{T}_P} \sup_{e \in \mathcal{C}_P} N_P(T, e; \tau_P, \delta).$$

All affine expressions involving $N_P^\star$ use the usual extended-real conventions. Comparison classes matter: unrelated transfer tasks, shared representation tasks, and slowly drifting continual streams are different statistical problems.

Description and purpose. These definitions describe the sample complexity of a fixed learner on a fixed environment, and then the minimax sample complexity of the best admissible learner on a comparison class. We need them because POR reductions are ultimately evaluated by their consequences for worst-case sample complexity on explicitly chosen environment classes.

2.3 Reduction between learning paradigms

The central idea is that paradigm B reduces to paradigm A if any algorithm that succeeds under A can be systematically transformed into an algorithm that succeeds under B, with controlled change in the success threshold and sample budget.

Definition 2.2 (Protocol–objective–resource reduction)**.** Let A and B be learning paradigms over the common hypothesis family $\mathcal{H}$, with accuracy regimes $\mathscr{T}_A \subseteq \mathbb{R}^{d_A}$ and $\mathscr{T}_B \subseteq \mathbb{R}^{d_B}$. We say that B *POR-reduces* to A relative to $(\mathscr{T}_A, \mathscr{T}_B)$, written

$$B \preceq_{\mathrm{POR}} A,$$

if there exist:

1. an environment embedding
$$\Phi : \mathcal{E}_B \to \mathcal{E}_A,$$
2. a compiler
$$\mathfrak{C} : \mathfrak{T}_A \to \mathfrak{T}_B,$$
3. a threshold transfer map
$$\Gamma : \mathscr{T}_A \to \mathscr{T}_B,$$
4. integer overhead parameters $a \in \mathbb{N}$ and $b \in \mathbb{N}$, with $a \geq 1$,

such that for every environment $e \in \mathcal{E}_B$, every learner $T_A \in \mathfrak{T}_A$, every threshold $\tau_A \in \mathscr{T}_A$, every confidence parameter $\delta \in (0,1)$, and every sample budget $M \in \mathbb{N}$, the implication

$$\Pr[\mathbf{J}_A(T_A, \Phi(e); M) \preceq \tau_A] \geq 1 - \delta$$

implies

$$\Pr[\mathbf{J}_B(\mathfrak{C}(T_A), e; aM + b) \preceq \Gamma(\tau_A)] \geq 1 - \delta.$$

Description and purpose. This definition describes when one learning paradigm can be simulated through another: environments are embedded, learners are compiled, thresholds are translated, and sample budgets are calibrated. We need this definition because informal claims that one paradigm contains another, or that a method transfers between paradigms, are ambiguous unless these four obligations are made explicit.

When the regimes are clear from context, we omit them from the notation. Exact reductions, hardness notation, and the transfer theorems below are always understood relative to the same fixed regimes. The word "embedding" denotes the environment map used by the reduction; injectivity is not required unless explicitly stated.

Remark 2.3. The compiler $\mathfrak{C}$ is required to transform *any* successful learner for A into a learner for B. This uniformity is what makes the reduction useful for proving hardness. It is not enough to show that some algorithm solves A and some possibly unrelated algorithm solves B.

Remark 2.4 (Why the accuracy regime matters). The definition packages four nontrivial obligations: an embedding of environments, a simulation of admissible interaction, a threshold transfer map, and a calibrated sample overhead. In applications, most proposed hardness comparisons fail because at least one of these ingredients is unavailable. The accuracy regimes are part of these obligations. Without them, one could map every target threshold to an uninformatively loose threshold and obtain a formally correct but meaningless comparison.

Definition 2.5 (Exact POR-reduction)**.** A POR-reduction is called *exact* if $a = 1$ and $b = 0$. In this case we write

$$B \preceq_{\mathrm{ex}} A.$$

Thus, exact reductions preserve the sample budget exactly.

Description and purpose. This definition describes the zero-overhead case of a POR reduction, where the simulation preserves the sample budget exactly. We need it to distinguish true special-case containment from reductions that require additional samples through an affine overhead.

Definition 2.6 (Relative hardness)**.** We say that A is at least as hard as B, written

$$A \succeq_{\mathrm{POR}} B,$$

if

$$B \preceq_{\mathrm{POR}} A.$$

Similarly, $A \succeq_{\mathrm{ex}} B$ means $B \preceq_{\mathrm{ex}} A$.

Description and purpose. This definition describes the hardness ordering induced by POR reductions: A is at least as hard as B when a solver for A can be used to solve B. We need this reverse-direction notation because the reduction relation points from the simulated problem to the solver problem, while hardness comparisons are usually read in the opposite direction.

3 Structural Results of the Framework

The preliminary setup of Section 2 identifies the basic objects that define a learning paradigm and the POR-reduction notion used to compare paradigms. We now prove the main structural consequences of that definition: reductions transfer sample-complexity upper bounds, transfer lower bounds in the reverse direction, and generate obstructions to proposed comparisons.

3.1 Sample-complexity transfer

The first result formalizes a useful budget-calibration step for POR-reductions.

Proposition 3.1 (Budget calibration)**.** *Fix accuracy regimes $\mathscr{T}_A$ and $\mathscr{T}_B$, and suppose the reduced paradigm B satisfies the budget-capping closure in Assumption 2.1. Suppose there exist an environment embedding Φ, a pre-compiler $\mathfrak{C}_0 : \mathfrak{T}_A \to \mathfrak{T}_B$, a threshold transfer map $\Gamma : \mathscr{T}_A \to \mathscr{T}_B$, and a nondecreasing map $\psi : \mathbb{N} \to \mathbb{N}$ such that for every $e \in \mathcal{E}_B$, $T_A \in \mathfrak{T}_A$, $\tau_A \in \mathscr{T}_A$, $\delta \in (0, 1)$, and $M \in \mathbb{N}$,*

$$\Pr[\mathbf{J}_A(T_A, \Phi(e); M) \preceq \tau_A] \geq 1 - \delta$$

implies

$$\Pr[\mathbf{J}_B(\mathfrak{C}_0(T_A), e; \psi(M)) \preceq \Gamma(\tau_A)] \geq 1 - \delta.$$

If there exist integer overhead parameters $a \in \mathbb{N}$ and $b \in \mathbb{N}$, with $a \geq 1$, such that

$$\psi(M) \leq aM + b \qquad \text{for all } M \in \mathbb{N},$$

then $B \preceq_{\mathrm{POR}} A$ with overhead parameters a, b.

Description and significance. This result shows that a simulation whose internal sample use is bounded by $aM + b$ can be converted into a valid POR-reduction by capping the reduced learner's budget. It is useful because reduction proofs can first build a natural simulation and then certify its resource overhead separately.

Remark 3.2 (Episode-to-example conversion). Proposition 3.1 is especially useful when two paradigms count different atomic units. If one unit of paradigm A can always be implemented using at most q units of paradigm B on the embedded class, then $\psi(M) = qM$ and the resulting POR-reduction has multiplicative overhead $a = q$, $b = 0$.

Theorem 3.3 (Reduction implies sample-complexity domination)**.** *Fix accuracy regimes $\mathscr{T}_A$ and $\mathscr{T}_B$. Suppose $B \preceq_{\mathrm{POR}} A$ relative to these regimes, with embedding Φ, threshold map Γ, and overhead parameters a, b. Then for every comparison class $\mathcal{C}_B \subseteq \mathcal{E}_B$, every threshold $\tau_A \in \mathscr{T}_A$, and every confidence level $\delta \in (0, 1)$,*

$$N_B^\star(\mathcal{C}_B; \Gamma(\tau_A), \delta) \;\leq\; a\, N_A^\star(\Phi(\mathcal{C}_B); \tau_A, \delta) + b,$$

where

$$\Phi(\mathcal{C}_B) := \{\Phi(e) : e \in \mathcal{C}_B\}.$$

Description and significance. The theorem states that a POR-reduction from B to A forces the minimax sample complexity of the embedded B-class to be bounded by that of A, up to the reduction overhead. This is the central guarantee of the framework: a valid reduction is not merely a formal simulation, but a certificate of worst-case sample-complexity domination.

Corollary 3.4 (Lower-bound transfer)**.** *Under the assumptions of Theorem 3.3, for every comparison class $\mathcal{C}_B \subseteq \mathcal{E}_B$, every threshold $\tau_A \in \mathscr{T}_A$, and every $\delta \in (0, 1)$,*

$$N_A^\star(\Phi(\mathcal{C}_B); \tau_A, \delta) \;\geq\; \frac{N_B^\star(\mathcal{C}_B; \Gamma(\tau_A), \delta) - b}{a}.$$

Hence, any lower bound for B induces a lower bound for A.

Description and significance. This corollary rewrites the main domination inequality as a lower-bound transfer principle. Its significance is that a hard embedded subproblem of B immediately yields a lower bound for the richer or more demanding paradigm A.

Corollary 3.5 (Exact sample-complexity domination)**.** *Under the assumptions of Theorem 3.3, if $B \preceq_{\mathrm{ex}} A$, then for every comparison class $\mathcal{C}_B \subseteq \mathcal{E}_B$, every threshold $\tau_A \in \mathscr{T}_A$, and every $\delta \in (0, 1)$,*

$$N_B^\star(\mathcal{C}_B; \Gamma(\tau_A), \delta) \leq N_A^\star(\Phi(\mathcal{C}_B); \tau_A, \delta).$$

Description and significance. Exact reductions remove the additive and multiplicative calibration terms from the comparison. Thus, when one paradigm contains another as a genuine special case, the resulting minimax sample-complexity inequality is direct and has no budget loss.

Corollary 3.6 (Sample-complexity obstruction to reductions)**.** *Fix candidate reduction data consisting of an environment embedding Φ, a threshold map $\Gamma : \mathscr{T}_A \to \mathscr{T}_B$, and overhead parameters a, b. If there exist a comparison class $\mathcal{C}_B$, a threshold $\tau_A \in \mathscr{T}_A$, and a confidence level $\delta \in (0, 1)$ such that*

$$N_B^\star(\mathcal{C}_B; \Gamma(\tau_A), \delta) > a\, N_A^\star(\Phi(\mathcal{C}_B); \tau_A, \delta) + b,$$

then no POR-reduction with this candidate data exists. In particular, if the displayed inequality can be witnessed for every possible embedding and threshold map, then $B \not\preceq_{\mathrm{POR}} A$ for the specified overhead parameters.

Description and significance. This obstruction gives a way to falsify proposed POR-reductions: if the sample complexities violate the inequality that a reduction would imply, that reduction cannot exist. It turns sample-complexity separations into diagnostic tools for checking informal claims of containment.

3.2 General structural properties

The next propositions record the main structural consequences of the framework.

Proposition 3.7 (Monotonicity in thresholds)**.** *Fix a learning paradigm P, a comparison class $\mathcal{C}_P \subseteq \mathcal{E}_P$, and $\delta \in (0,1)$. If $\tau_P, \tau'_P \in \mathbb{R}^{d_P}$ satisfy*

$$\tau_P \preceq \tau'_P,$$

then

$$N^\star_P(\mathcal{C}_P; \tau_P, \delta) \geq N^\star_P(\mathcal{C}_P; \tau'_P, \delta).$$

Description and significance. The proposition records that weaker success requirements cannot increase minimax sample complexity. This sanity check matters because POR comparisons depend on the chosen accuracy regime, and monotonicity ensures that threshold relaxations behave as expected.

Proposition 3.8 (Monotonicity in comparison classes)**.** *Fix a learning paradigm P, a threshold τ_P, and $\delta \in (0,1)$. If*

$$\mathcal{C}_P \subseteq \mathcal{C}'_P \subseteq \mathcal{E}_P,$$

then

$$N^\star_P(\mathcal{C}_P; \tau_P, \delta) \leq N^\star_P(\mathcal{C}'_P; \tau_P, \delta).$$

Description and significance. The result states that enlarging the comparison class cannot make a minimax learning problem easier. It is important because claims about paradigm hardness are only meaningful relative to the environment class over which uniform performance is required.

Proposition 3.9 (Preorder structure for regime-equipped paradigms)**.** *Fix an accuracy regime for each paradigm under consideration. The relation $\succeq_{\mathrm{POR}}$ is a preorder on these regime-equipped paradigms. In particular:*

1. *it is reflexive;*
2. *it is transitive.*

The same is true for $\succeq_{\mathrm{ex}}$.

Description and significance. This proposition shows that the POR hardness relation is reflexive and transitive once accuracy regimes are fixed. The result justifies chaining reductions across paradigms and treating the framework as an organizing structure rather than a collection of isolated comparisons.

Corollary 3.10 (Equivalence classes)**.** *For regime-equipped paradigms, define $A \equiv_{\mathrm{POR}} B$ if $A \succeq_{\mathrm{POR}} B$ and $B \succeq_{\mathrm{POR}} A$. Then $\equiv_{\mathrm{POR}}$ is an equivalence relation, and $\succeq_{\mathrm{POR}}$ induces a partial order on the equivalence classes.*

Description and significance. The corollary says that mutually reducible paradigms form equivalence classes, and that these classes are partially ordered by POR hardness. This makes precise when two superficially different learning setups should be treated as structurally equivalent for the purposes of worst-case sample complexity.

3.3 Generic sufficient conditions for hardness

The next proposition gives several reusable proof templates. In practice, most cross-paradigm hardness results are instances of one of these patterns.

Proposition 3.11 (Generic sufficient conditions)**.** *Let A and B be learning paradigms over the same hypothesis family, with fixed accuracy regimes $\mathscr{T}_A$ and $\mathscr{T}_B$.*

1. ***Special-case embedding.*** *Suppose there exist an embedding $\Phi : \mathcal{E}_B \to \mathcal{E}_A$ and a compiler $\mathfrak{C} : \mathfrak{T}_A \to \mathfrak{T}_B$ such that $\mathfrak{C}(T)$ simulates every execution of learner T for A on $\Phi(e)$ using the information available from e. Suppose also that there exists $\Gamma : \mathscr{T}_A \to \mathscr{T}_B$ such that*

$$\{\mathbf{J}_A(T, \Phi(e); M) \preceq \tau_A\} \subseteq \{\mathbf{J}_B(\mathfrak{C}(T), e; M) \preceq \Gamma(\tau_A)\}$$

up to null events, for all e, T, M, and all $\tau_A \in \mathscr{T}_A$. Then $A \succeq_{\text{ex}} B$.

2. ***Objective strengthening.*** *Suppose A and B have the same environments, protocol, learner class, and resource accounting, and suppose there exists $\Gamma : \mathscr{T}_A \to \mathscr{T}_B$ such that*

$$\{\mathbf{J}_A(T, e; M) \preceq \tau_A\} \subseteq \{\mathbf{J}_B(T, e; M) \preceq \Gamma(\tau_A)\}$$

up to null events, for all e, T, M, and all $\tau_A \in \mathscr{T}_A$. Then $A \succeq_{\text{ex}} B$.

3. ***Information monotonicity.*** *Suppose A and B have the same environments, learner class, objective, resource accounting, and accuracy regime, but B's protocol reveals at least as much side information as A's protocol, and learners under B are free to ignore the extra information. Then $A \succeq_{\text{ex}} B$.*

4. ***Resource monotonicity.*** *Suppose A and B have the same environments, protocol, objective, resource accounting, and accuracy regime, but*

$$\mathfrak{T}_A \subseteq \mathfrak{T}_B.$$

Then $A \succeq_{\text{ex}} B$.

Description and significance. The proposition gives reusable routes for proving POR-reductions through special-case embeddings, objective strengthening, information monotonicity, and resource monotonicity. Its value is practical: it converts common informal arguments about one paradigm containing another into a short list of formal obligations.

3.4 Strictness of objective-based hardness

Objective strengthening is not merely a formal inclusion: changing an objective can strictly change minimax sample complexity even when the environment class, protocol, learner class, and hypothesis family are fixed. The next proposition gives a minimal example. In the two constructions below, h_T^M denotes the predictor output by learner T after the M-sample transcript generated by the stated i.i.d. protocol.

Proposition 3.12 (Objective strengthening can be strict)**.** *Let $\mathcal{X} = \{1, 2\}$, $\mathcal{Y} = \{0, 1\}$, let ℓ be the zero-one loss, and let $\mathcal{H}$ contain all functions $\mathcal{X} \to \mathcal{Y}$. For each $\sigma = (\sigma_1, \sigma_2) \in \{0, 1\}^2$, let D_σ be the distribution that draws x uniformly from $\{1, 2\}$ and returns the deterministic label $y = \sigma_x$. Consider two paradigms with the same environment class $\{D_\sigma : \sigma \in \{0, 1\}^2\}$, the same i.i.d. sampling protocol, and the same unrestricted learner class:*

$$\mathbf{J}_{\text{one}}(T, D_\sigma; M) = \mathbf{1}\{h_T^M(1) \neq \sigma_1\},$$

and

$$\mathbf{J}_{\text{all}}(T, D_\sigma; M) = \big(\mathbf{1}\{h_T^M(1) \neq \sigma_1\}, \mathbf{1}\{h_T^M(2) \neq \sigma_2\}\big).$$

Use the zero-threshold accuracy regimes $\mathscr{T}_{\text{one}} = \{0\}$ and $\mathscr{T}_{\text{all}} = \{(0, 0)\}$, and let $\mathcal{C} = \{D_\sigma : \sigma \in \{0, 1\}^2\}$. Then

$$P_{\text{all}} \succeq_{\text{ex}} P_{\text{one}},$$

but the domination is strict: for every $\delta \in (0, 1/2)$,

$$N^\star_{P_{\text{one}}}(\mathcal{C}; 0, \delta) = \left\lceil \log_2 \frac{1}{2\delta} \right\rceil, \qquad N^\star_{P_{\text{all}}}(\mathcal{C}; (0, 0), \delta) = \left\lceil \log_2 \frac{1}{\delta} \right\rceil.$$

Consequently P_{all} and P_{one} are not equivalent under exact POR-reductions relative to these zero-threshold regimes.

Description and significance. This example proves that a stronger objective can strictly increase minimax sample complexity even when the environments, data protocol, hypothesis class, and learner class are unchanged. It demonstrates that POR hardness is not only about information access; the definition of success itself can create genuine statistical difficulty.

Proposition 3.13 (A scalable objective-strengthening gap)**.** *For each $m \geq 3$, let $\mathcal{X}_m = \{1, \ldots, m\}$, $\mathcal{Y} = \{0, 1\}$, and let $\mathcal{H}_m$ contain all functions $\mathcal{X}_m \to \mathcal{Y}$. For each $\sigma \in \{0,1\}^m$, let D_σ draw x uniformly from $\mathcal{X}_m$ and return $y = \sigma_x$. Define two paradigms with the same environment class, i.i.d. sampling protocol, and unrestricted learner class:*

$$\mathbf{J}^{(m)}_{\text{one}}(T, D_\sigma; M) = \mathbf{1}\{h^M_T(1) \neq \sigma_1\},$$

and

$$\mathbf{J}^{(m)}_{\text{all}}(T, D_\sigma; M) = \big(\mathbf{1}\{h^M_T(j) \neq \sigma_j\}\big)^m_{j=1}.$$

Let $\mathcal{C}_m = \{D_\sigma : \sigma \in \{0,1\}^m\}$, use the singleton accuracy regimes

$$\mathscr{T}^{(m)}_{\text{one}} = \{0\}, \qquad \mathscr{T}^{(m)}_{\text{all}} = \{\mathbf{0}_m\},$$

where $\mathbf{0}_m$ is the zero vector in $\mathbb{R}^m$, and set $\delta = 1/8$. Then

$$N^\star_{P^{(m)}_{\text{one}}}(\mathcal{C}_m; 0, \delta) \leq 2m, \qquad N^\star_{P^{(m)}_{\text{all}}}(\mathcal{C}_m; \mathbf{0}_m, \delta) \geq \frac{m}{4} \log m.$$

Consequently, no family of reverse reductions $P^{(m)}_{\text{all}} \preceq_{\text{POR}} P^{(m)}_{\text{one}}$, relative to these singleton regimes, can have affine overhead parameters a, b that are uniform in m.

Description and significance. This scalable example amplifies the previous strictness result by producing a growing gap between one-coordinate and all-coordinate objectives. The growth rules out a uniform affine-overhead reverse reduction, showing that objective strengthening can change the structural hardness class rather than only the constants.

Remark 3.14 (Exact versus affine overhead)*.* Proposition 3.12 rules out exact reverse reductions in a minimal two-point example, while Proposition 3.13 rules out reverse reductions with any fixed affine overhead across the scalable family. This distinction matters: constant sample gaps are meaningful for exact containment, but affine-overhead POR reductions intentionally ignore fixed additive and multiplicative calibration costs.

4 Hardness Relations Among Representative Paradigms

We now instantiate the general framework on four representative paradigms: supervised learning, transfer learning, continual learning, and meta-learning. The purpose of this section is not to impose a universal total ordering on all paradigms, but rather to identify the hardness relations that follow directly from the framework under natural closure assumptions.

4.1 Canonical forms of the representative paradigms

We use the following simplified canonical models. Whenever K source or sequential tasks appear, K is fixed for that comparison; varying K can be treated as a family of POR objects, keeping the performance dimension fixed inside each reduction statement. Throughout this subsection, $h^M_{T,e}$ denotes the predictor output by T after the M-sample interaction with environment e; when the environment is a single distribution D, we also write $h^M_{T,D}$.

Supervised learning. An environment is a single distribution D over $\mathcal{X} \times \mathcal{Y}$. The protocol reveals i.i.d. labeled samples from D. The performance functional is

$$\mathbf{J}_{\text{Sup}}(T, D; M) = R_D(h^M_{T,D}),$$

where $h^M_{T,D} \in \mathcal{H}$ is the predictor produced by learner T after budget M. This canonical form follows the standard distributional setup in statistical learning theory (Vapnik, 1998; Shalev-Shwartz & Ben-David, 2014; Mohri et al., 2018).

Transfer learning. An environment is a tuple

$$e = (D_1, \ldots, D_K; D_\star),$$

where $D_1, \ldots, D_K$ are source tasks and $D_\star$ is the target task. The protocol reveals source samples first and target samples afterward. The performance functional is

$$\mathbf{J}_{\mathrm{Trans}}(T, e; M) = R_{D_\star}(h^M_{T,e}).$$

This abstraction captures the source-target viewpoint common in transfer and domain-adaptation theory (Pan & Yang, 2010; Ben-David et al., 2010; Blitzer et al., 2007; Mansour et al., 2008; Kalan et al., 2020).

Continual learning. An environment is a sequential task stream

$$e = (D_1, \ldots, D_K).$$

The protocol reveals tasks sequentially. We use the performance functional

$$\mathbf{J}_{\mathrm{Cont}}(T, e; M) = \big(R_{D_1}(h^M_{T,e}), \ldots, R_{D_K}(h^M_{T,e}), F(T, e; M)\big),$$

where $F(T, e; M)$ is a forgetting functional. This form reflects the standard emphasis on sequential exposure, replay, task information, and forgetting in continual learning (Robins, 1995; French, 1999; Parisi et al., 2019; Lopez-Paz & Ranzato, 2017; De Lange et al., 2022; Van de Ven et al., 2022).

Meta-learning. An environment is a distribution Π over tasks D. The protocol reveals support/query episodes sampled from Π. If $A^M_{T,\Pi}(S) \in \mathcal{H}$ denotes the predictor obtained after the learner adapts to support set S, we write the performance functional abstractly as

$$\mathbf{J}_{\mathrm{Meta}}(T, \Pi; M) = \mathbb{E}_{D\sim\Pi}\Big[R_D\big(A^M_{T,\Pi}(S_D)\big)\Big],$$

where S_D is the support sample drawn for task D. The support/query abstraction is standard in learning-to-learn and few-shot meta-learning (Baxter, 2000; Vinyals et al., 2016; Santoro et al., 2016; Ravi & Larochelle, 2017; Snell et al., 2017; Finn et al., 2017; Hospedales et al., 2021).

Table 1: Canonical instantiations of the four representative paradigms. The table suppresses algorithmic details and keeps only the ingredients needed for POR comparisons.

Paradigm	Environment class	Controlled performance
Supervised	single task distribution D	population risk on D
Transfer	$(D_1, \ldots, D_K; D_\star)$	target risk after source-then-target interaction
Continual	ordered task stream $(D_1, \ldots, D_K)$	risks on all tasks together with forgetting
Meta	task distribution Π	expected post-adaptation query risk

4.2 Closure assumptions

To derive concrete hardness statements, we use the following natural closure assumptions. Fix a nonempty scalar risk regime $\mathscr{I} \subseteq \mathbb{R}_+$. Supervised, transfer, and meta-learning use $\mathscr{I}$; continual learning uses threshold vectors whose final embedded-task coordinate lies in $\mathscr{I}$. The threshold maps below are identity maps on $\mathscr{I}$ for scalar-risk reductions and the final-task coordinate projection for transfer-to-continual reduction.

Assumption 4.1 (Degenerate-source closure for transfer)**.** The transfer-learning environment class contains empty-source instances of the form

$$(\varnothing; D)$$

for every supervised environment D.

Description and reasonableness. This assumption states that the transfer-learning environment class includes the degenerate case in which there are no source tasks and only the target task remains. It is reasonable because transfer learning is normally an extension of supervised learning: when the source phase is empty, the learner simply receives target samples and is evaluated on target risk. The assumption therefore formalizes the intended special-case containment rather than adding a new statistical advantage.

Assumption 4.2 (Target-risk projection for continual learning)**.** If

$$e = (D_1, \ldots, D_K; D_\star)$$

is viewed as a transfer-learning environment, then its sequential embedding

$$\Phi_{\text{Trans}\to\text{Cont}}(e) := (D_1, \ldots, D_K, D_\star)$$

belongs to the continual-learning environment class, and the continual-learning objective contains the final-task risk $R_{D_\star}$ as one of its controlled coordinates.

Description and reasonableness. This assumption states that a source-target transfer environment can be embedded into a continual-learning stream by presenting the source tasks first and the target task last, and that the continual objective records the final target risk. It is reasonable for continual-learning formulations that evaluate performance across the task stream, because the final task is one of the tasks on which performance is controlled. The assumption is also explicit about the needed objective alignment: without a final-task-risk coordinate, a continual learner could succeed while failing the transfer objective.

Assumption 4.3 (Singleton-task closure for meta-learning)**.** For every supervised environment D, the meta-learning environment class contains the degenerate task distribution

$$\Pi_D := \delta_D,$$

the Dirac measure concentrated on the single task D. Moreover, the meta learner is allowed to use its supervised samples to synthesize the support/query episodes required by the meta protocol.

Description and reasonableness. This assumption states that the meta-learning environment class includes the degenerate task distribution concentrated on a single supervised task, and that samples from this task can be organized into the support/query interface used by the meta protocol. It is reasonable because episodic meta-learning is usually defined for task distributions, and a point mass on one task is the simplest valid task distribution. The sample-partition condition makes the resource convention explicit: the reduction is exact only when the supervised samples can be reused to instantiate the required episodes without extra sample cost.

Table 2 records the reduction data used below. The following propositions then verify that these data satisfy the POR definition under the stated closure assumptions.

Table 2: Reduction data for the canonical exact reductions. Each row lists the four POR obligations: environment embedding Φ, learner compiler $\mathfrak{C}$, threshold map Γ, and sample overhead.

Red.	Embedding Φ	Compiler $\mathfrak{C}$	Γ	Cost
Sup $\preceq$ Trans	$D \mapsto (\varnothing; D)$	Run the transfer learner with no source data and target samples from D.	Identity on risk	$(1, 0)$
Trans $\preceq$ Cont	$(D_1, \ldots, D_K; D_\star) \mapsto (D_1, \ldots, D_K, D_\star)$	Feed source and target tasks as a continual stream; output the final predictor.	Project final-task risk	$(1, 0)$
Sup $\preceq$ Meta	$D \mapsto \delta_D$	Partition samples from D into the allowed support/query interface.	Identity on post-adaptation risk	$(1, 0)$

4.3 Concrete hardness relations

The first result says that transfer learning is at least as hard as supervised learning.

Proposition 4.4 (Transfer learning dominates supervised learning)**.** *Under Assumption 4.1,*

$$\text{Trans} \succeq_{\text{ex}} \text{Sup}.$$

Description and significance. The proposition embeds supervised learning into transfer learning by treating the source phase as empty and the target phase as the supervised problem. It anchors the hierarchy with an exact special-case containment and shows that worst-case transfer learning is at least as hard as supervised learning under the stated closure assumption.

The next proposition formalizes the intuition that continual learning is at least as hard as transfer learning when the continual objective requires control of all tasks, including the final one.

Proposition 4.5 (Continual learning dominates transfer learning)**.** *Under Assumption 4.2,*

$$\text{Cont} \succeq_{\text{ex}} \text{Trans}.$$

Description and significance. This result embeds a transfer problem into a continual-learning stream by placing the source tasks before the target task. Its significance is that continual learning formally contains the source-target transfer setting when the continual objective includes the target-risk requirement.

Case study: impossibility transfer. The reduction immediately turns transfer-learning impossibility results into continual-learning impossibility results on the corresponding embedded streams. Existing transfer-learning lower bounds show that, without appropriate task relatedness or diversity assumptions, source tasks can be uninformative for the target and transfer may fail to achieve a desired target-risk guarantee within a given sample budget (Ben-David et al., 2010; Kalan et al., 2020; Tripuraneni et al., 2020). Suppose such a result gives, for a transfer comparison class $\mathcal{C}_{\text{Trans}}$, threshold τ_{Trans}, confidence level δ, and budget M,

$$N^{\star}_{\text{Trans}}(\mathcal{C}_{\text{Trans}}; \tau_{\text{Trans}}, \delta) > M,$$

or even $N^{\star}_{\text{Trans}}(\mathcal{C}_{\text{Trans}}; \tau_{\text{Trans}}, \delta) = \infty$. Then for every continual-learning threshold τ_{Cont} whose final-task projection is τ_{Trans}, Proposition 4.5 and Corollary 3.4 imply

$$N^{\star}_{\text{Cont}}(\Phi_{\text{Trans}\to\text{Cont}}(\mathcal{C}_{\text{Trans}}); \tau_{\text{Cont}}, \delta) \geq N^{\star}_{\text{Trans}}(\mathcal{C}_{\text{Trans}}; \tau_{\text{Trans}}, \delta).$$

Thus the same impossibility or lower bound holds for continual learning on the embedded task streams. No separate continual-learning lower-bound proof is needed: any continual learner that solved these streams would also solve the original transfer problem through the reduction.

Corollary 4.6 (Continual learning dominates supervised learning)**.** *Under Assumptions 4.1 and 4.2,*

$$\text{Cont} \succeq_{\text{ex}} \text{Sup}.$$

Description and significance. The corollary uses transitivity to show that supervised learning is exactly contained in continual learning under the two closure assumptions. It illustrates how the preorder structure lets local reductions compose into broader relationships among paradigms.

The next result shows that meta-learning also contains supervised learning as a degenerate special case.

Proposition 4.7 (Meta-learning dominates supervised learning)**.** *Under Assumption 4.3,*

$$\text{Meta} \succeq_{\text{ex}} \text{Sup}.$$

Description and significance. The proposition embeds supervised learning into meta-learning by using a task distribution concentrated on one task and reusing samples through the support/query interface. It shows that meta-learning's distinctiveness comes from non-degenerate task variation and adaptation structure, not from the mere presence of an episodic protocol.

Remark 4.8 (Sensitivity of the meta-learning reduction). Proposition 4.7 relies on the ability to recycle supervised samples into the support/query structure required by the meta protocol. If instead the meta setup required externally supplied episodic partitions, or charged separate resource costs for support and adaptation, the same argument would generally yield only a non-exact POR-reduction with explicit overhead. This illustrates why the closure assumptions are substantive rather than merely notational.

4.4 Resulting picture

Collecting the previous propositions yields the following canonical hardness diagram:

$$\mathrm{Cont} \succeq_{\mathrm{ex}} \mathrm{Trans} \succeq_{\mathrm{ex}} \mathrm{Sup}, \qquad \mathrm{Meta} \succeq_{\mathrm{ex}} \mathrm{Sup}.$$

Corollary 4.9 (Sample-complexity consequences for the canonical hierarchy)**.** *Under Assumptions 4.1, 4.2, and 4.3, the exact reductions above imply the following minimax inequalities on embedded comparison classes.*

For every supervised comparison class $\mathcal{C}_{\mathrm{Sup}}$, every supervised threshold τ_{Sup}, and every $\delta \in (0,1)$,

$$N^{\star}_{\mathrm{Sup}}(\mathcal{C}_{\mathrm{Sup}}; \tau_{\mathrm{Sup}}, \delta) \le N^{\star}_{\mathrm{Trans}}(\Phi_{\mathrm{Sup}\to\mathrm{Trans}}(\mathcal{C}_{\mathrm{Sup}}); \tau_{\mathrm{Sup}}, \delta),$$

$$N^{\star}_{\mathrm{Sup}}(\mathcal{C}_{\mathrm{Sup}}; \tau_{\mathrm{Sup}}, \delta) \le N^{\star}_{\mathrm{Cont}}((\Phi_{\mathrm{Trans}\to\mathrm{Cont}} \circ \Phi_{\mathrm{Sup}\to\mathrm{Trans}})(\mathcal{C}_{\mathrm{Sup}}); \tau_{\mathrm{Cont}}, \delta),$$

for every continual-learning threshold τ_{Cont} whose final-task projection equals τ_{Sup}, and

$$N^{\star}_{\mathrm{Sup}}(\mathcal{C}_{\mathrm{Sup}}; \tau_{\mathrm{Sup}}, \delta) \le N^{\star}_{\mathrm{Meta}}(\Phi_{\mathrm{Sup}\to\mathrm{Meta}}(\mathcal{C}_{\mathrm{Sup}}); \tau_{\mathrm{Sup}}, \delta).$$

Description and significance. This corollary translates the canonical exact reductions into explicit minimax sample-complexity inequalities for the embedded supervised, transfer, continual, and meta-learning classes. It is significant because it turns the structural hierarchy into quantitative consequences that can be used to transfer upper and lower bounds.

Remark 4.10 (No universal total order). The framework does not by itself imply a universal ordering between Meta and Trans, or between Meta and Cont. Such comparisons depend on additional structural assumptions about the environment classes and the objectives. For example, one would need a formal embedding of source-target problems into episodic task distributions, or conversely an embedding of episodic adaptation problems into sequential task streams. In this sense, the framework naturally produces a preorder rather than a total order.

Remark 4.11 (Interpretation of exact hardness). An exact POR-reduction says that one paradigm contains another as a special case without inflating the sample budget. This should be interpreted as a structural statement about worst-case sample difficulty under aligned sample units. It does not claim that optimization, constants, or empirical benchmark performance are identical across paradigms.

Remark 4.12 (How existing lower bounds plug into the framework). The point of the framework is not to replace within-paradigm lower-bound analysis, but to make such results reusable. For instance, any lower bound proved for a task-aware continual-learning model or for a large-replay variant immediately transfers to the harder task-agnostic or smaller-memory variants by Corollary 3.4 together with Propositions A.2 and A.3. Likewise, lower bounds proved within source-target transfer models (Kalan et al., 2020) can be imported into stronger structural settings once an appropriate POR embedding is specified. The same principle applies to continual-learning variants whenever the relevant replay, memory, or side-information choices are represented as POR resources.

5 Related Work

Classical learning theory relates sample complexity to a fixed hypothesis class and loss through tools such as uniform convergence, stability, and capacity measures (Vapnik & Chervonenkis, 1971; Valiant, 1984; Blumer et al., 1989; Vapnik, 1998; Anthony & Bartlett, 1999; Shalev-Shwartz & Ben-David, 2014; Mohri et al., 2018; Bartlett & Mendelson, 2002; Bousquet & Elisseeff, 2002). Our focus is orthogonal: the hypothesis family is fixed, while the protocol, objective, or resources change. Transfer-learning theory studies when source tasks

improve target performance, including representation sharing, task relatedness, and minimax lower bounds (Caruana, 1997; Pan & Yang, 2010; Ben-David et al., 2010; Blitzer et al., 2007; Crammer et al., 2008; Mansour et al., 2008; Weiss et al., 2016; Maurer et al., 2016; Tripuraneni et al., 2020; Kalan et al., 2020). Meta-learning and learning-to-learn analyze experience across related tasks and episodic adaptation (Thrun; Baxter, 2000; Maurer & Jaakkola, 2005; Lake et al., 2015; Andrychowicz et al., 2016; Santoro et al., 2016; Vinyals et al., 2016; Ravi & Larochelle, 2017; Snell et al., 2017; Finn et al., 2017; Hospedales et al., 2021). Continual learning studies sequential exposure, stability-plasticity tradeoffs, replay, task identifiers, and forgetting (Robins, 1995; French, 1999; Parisi et al., 2019; Kirkpatrick et al., 2017; Lopez-Paz & Ranzato, 2017; Rebuffi et al., 2017; Pentina & Lampert, 2015; De Lange et al., 2022; Van de Ven et al., 2022). POR-reductions do not replace these within-paradigm analyses; they specify when their guarantees or lower bounds transfer across paradigm interfaces.

The closest classical statistical analogue is the comparison of statistical experiments. Blackwell's comparison theorem orders experiments by whether one experiment can simulate another through randomization without increasing decision risk (Blackwell, 1953). Le Cam's theory of deficiency and asymptotic comparison extends this viewpoint to approximate comparison of statistical models (Le, 1964; Le Cam, 2012), with later systematic treatments such as (Torgersen, 1991). These works compare information structures through decision performance, and they are therefore conceptually close to the present paper. POR-reductions differ in the objects being compared: rather than comparing statistical experiments for a fixed decision problem, they compare learning paradigms that may differ in environment class, observation protocol, admissible learner class, performance functional, and resource accounting. The threshold map and budget overhead in a POR-reduction make explicit two ingredients that are usually outside classical experiment comparison: how learning objectives are translated and how sample resources are calibrated across paradigms. Moreover, POR comparisons are algorithm-interface comparisons: the compiler must transform every admissible learner for the solver paradigm into a learner for the reduced paradigm under the target protocol. This learner-class and resource-sensitive perspective is what allows the framework to discuss method transfer, replay memory, task identifiers, and adaptation protocols, which are not naturally represented by an information ordering between experiments alone.

Reduction arguments are central in computational complexity (Cook, 1971; KARP, 1972; Garey & Johnson, 1979; Papadimitriou, 1994; Arora & Barak, 2009) and also appear within learning theory, for example in statistical-query learning (Kearns, 1998) and classification reductions (Beygelzimer et al., 2005). The present work uses reductions at the level of learning paradigms. The reduction data therefore include an environment embedding, learner compiler, threshold map, and budget calibration, which are precisely the ingredients needed for minimax sample-complexity transfer across protocol–objective–resource specifications.

6 Conclusion

We introduced a protocol–objective–resource framework for comparing learning paradigms while holding the hypothesis family fixed. The central message is that relationships among learning paradigms cannot be read from model classes alone. They depend on how environments are embedded, how learners are compiled, how success criteria are translated, and how samples and other resources are accounted for. POR-reductions make these obligations explicit and convert them into minimax sample-complexity comparisons.

The framework gives a formal language for transferring understanding and methods across learning regimes. On the understanding side, a valid reduction transports upper bounds, lower bounds, and impossibility evidence between embedded comparison classes. On the method side, the same reduction specifies what must be preserved for an algorithmic idea to move from one paradigm to another: the interaction transcript, the performance target, and the resource budget. This helps distinguish genuine paradigm-level novelty from differences caused by side information, objective strength, memory assumptions, or sample accounting. In this sense, POR-reductions provide a way to organize the growing landscape of learning regimes and to avoid rederiving the same insight under different names.

Our structural results support this view. The reduction theorem gives sample-complexity domination on embedded classes, its converse use yields obstructions to proposed reductions, and the preorder structure allows local comparisons to be composed into larger diagrams. The objective-strengthening examples show

that the objective component is not cosmetic: even with the same protocol and learner class, changing what must be controlled can strictly increase minimax sample complexity. The instantiations on supervised, transfer, continual, and meta-learning abstractions demonstrate how familiar learning regimes can be related through exact special-case containments under explicit closure assumptions, while the replay-memory and task-identifier examples show that the same language also compares variants inside a single paradigm.

6.1 Limitations and Future Work

The framework is intentionally structural and worst-case. It does not claim that a POR reduction preserves optimization difficulty, constants, empirical benchmark performance, or typical-case behavior. It also does not impose a universal total order over learning paradigms. Comparisons depend on the chosen environment classes, accuracy regimes, closure assumptions, and resource conventions; changing any of these ingredients can change the resulting hardness relation.

Several extensions are natural. First, the resource model should be enriched beyond a single sample budget. Many practical comparisons also depend on memory, computation, communication, adaptation rounds, replay buffers, and access to task identifiers. A multi-resource POR theory could make these tradeoffs explicit. Second, approximate reductions could allow controlled slack in the performance threshold as well as in the resource budget, better matching empirical method transfer where small losses are often accepted. Third, distribution-dependent or randomized reductions could capture settings in which structural transfer holds for most tasks under a meta-distribution rather than uniformly over a comparison class. Finally, the framework should be paired with sharper within-paradigm sample-complexity results. POR-reductions are most useful when concrete upper and lower bounds are available to transport across learning interfaces.

Broader Impact Statement

This paper is theoretical and does not introduce a deployable learning system, dataset, or benchmark. Its main purpose is to clarify when sample-complexity claims can be transferred across learning paradigms. The most direct positive impact is methodological: making protocol, objective, and resource assumptions explicit can reduce misleading comparisons between learning settings. Because the work is conceptual rather than application-oriented, we do not anticipate significant immediate societal risks or direct downstream impacts beyond its potential influence on future machine learning research.

References

Marcin Andrychowicz, Misha Denil, Sergio Gomez, Matthew W Hoffman, David Pfau, Tom Schaul, Brendan Shillingford, and Nando De Freitas. Learning to learn by gradient descent by gradient descent. *Advances in neural information processing systems*, 29, 2016.

Martin Anthony and P Bartlett. *Neural network learning: theoretical foundations*. Cambridge University Press, 1999.

Sanjeev Arora and Boaz Barak. *Computational complexity: a modern approach*. Cambridge University Press, 2009.

Peter L Bartlett and Shahar Mendelson. Rademacher and gaussian complexities: Risk bounds and structural results. *Journal of machine learning research*, 3(Nov):463–482, 2002.

J Baxter. A model of inductive bias learning. *The Journal of Artificial Intelligence Research*, 12:149, 2000.

Shai Ben-David, John Blitzer, Koby Crammer, Alex Kulesza, Fernando Pereira, and Jennifer Wortman Vaughan. A theory of learning from different domains. *Machine learning*, 79(1):151–175, 2010.

Yoshua Bengio, Aaron Courville, and Pascal Vincent. Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence*, 35(8):1798–1828, 2013.

Alina Beygelzimer, Varsha Dani, Tom Hayes, John Langford, and Bianca Zadrozny. Error limiting reductions between classification tasks. In *Proceedings of the 22nd international conference on Machine learning*, pp. 49–56, 2005.

David Blackwell. Equivalent comparisons of experiments. *The annals of mathematical statistics*, pp. 265–272, 1953.

John Blitzer, Koby Crammer, Alex Kulesza, Fernando Pereira, and Jennifer Wortman. Learning bounds for domain adaptation. *Advances in neural information processing systems*, 20, 2007.

Anselm Blumer, David Haussler, and Manfred K Warmuth. Learnability and the vapnik-chervonenkis dimension. *Journal of the ACM (JACM)*, 36(4):929–965, 1989.

Olivier Bousquet and André Elisseeff. Stability and generalization. *Journal of machine learning research*, 2 (Mar):499–526, 2002.

Rich Caruana. Multitask learning. *Machine learning*, 28(1):41–75, 1997.

Stephen A Cook. The complexity of theorem-proving procedures. In *Proceedings of the third annual ACM symposium on Theory of computing*, pp. 151–158, 1971.

Koby Crammer, Michael Kearns, and Jennifer Wortman. Learning from multiple sources. *Journal of machine learning research*, 9(8), 2008.

Matthias De Lange, Rahaf Aljundi, Marc Masana, Sarah Parisot, Xu Jia, Ales Leonardis, Gregory Slabaugh, and Tinne Tuytelaars. A continual learning survey: Defying forgetting in classification tasks. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 44(07):3366–3385, 2022.

Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *International conference on machine learning*, pp. 1126–1135. PMLR, 2017.

Robert M French. Catastrophic forgetting in connectionist networks. *Trends in cognitive sciences*, 3(4): 128–135, 1999.

Michael R Garey and David S Johnson. Computers and intractability: A guide to the theory of np-completeness, 1979.

Timothy Hospedales, Antreas Antoniou, Paul Micaelli, and Amos Storkey. Meta-learning in neural networks: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 44(9):5149–5169, 2021.

Seyed Mohammadreza Mousavi Kalan, Zalan Fabian, Salman Avestimehr, and Mahdi Soltanolkotabi. Minimax lower bounds for transfer learning with linear and one-hidden layer neural networks. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, pp. 1959–1969, 2020.

RM KARP. Reducibility among combinatorial problems. *Complexity of Computer Computations*, pp. 85–103, 1972.

Michael Kearns. Efficient noise-tolerant learning from statistical queries. *Journal of the ACM (JACM)*, 45(6): 983–1006, 1998.

James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, et al. Overcoming catastrophic forgetting in neural networks. *Proceedings of the national academy of sciences*, 114(13):3521–3526, 2017.

Brenden M Lake, Ruslan Salakhutdinov, and Joshua B Tenenbaum. Human-level concept learning through probabilistic program induction. *Science*, 350(6266):1332–1338, 2015.

L Le. Sufficiency and approximate sufficiency. *The Annals of Mathematical Statistics*, pp. 1419–1455, 1964.

Lucien Le Cam. *Asymptotic methods in statistical decision theory*. Springer Science & Business Media, 2012.

David Lopez-Paz and Marc'Aurelio Ranzato. Gradient episodic memory for continual learning. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pp. 6470–6479, 2017.

Yishay Mansour, Mehryar Mohri, and Afshin Rostamizadeh. Domain adaptation with multiple sources. *Advances in neural information processing systems*, 21, 2008.

Andreas Maurer and Tommi Jaakkola. Algorithmic stability and meta-learning. *Journal of Machine Learning Research*, 6(6), 2005.

Andreas Maurer, Massimiliano Pontil, and Bernardino Romera-Paredes. The benefit of multitask representation learning. *Journal of Machine Learning Research*, 17(81):1–32, 2016.

Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of Machine Learning*. MIT Press, 2018.

Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10):1345–1359, 2010.

Christos H Papadimitriou. Computational complexity. 1994.

German I Parisi, Ronald Kemker, Jose L Part, Christopher Kanan, and Stefan Wermter. Continual lifelong learning with neural networks: A review. *Neural networks*, 113:54–71, 2019.

Anastasia Pentina and Christoph H Lampert. Lifelong learning with non-iid tasks. *Advances in Neural Information Processing Systems*, 28, 2015.

Sachin Ravi and Hugo Larochelle. Optimization as a model for few-shot learning. In *International conference on learning representations*, 2017.

Sylvestre-Alvise Rebuffi, Alexander Kolesnikov, Georg Sperl, and Christoph H Lampert. icarl: Incremental classifier and representation learning. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pp. 2001–2010, 2017.

Anthony Robins. Catastrophic forgetting, rehearsal and pseudorehearsal. *Connection Science*, 7(2):123–146, 1995.

Adam Santoro, Sergey Bartunov, Matthew Botvinick, Daan Wierstra, and Timothy Lillicrap. Meta-learning with memory-augmented neural networks. In *International conference on machine learning*, pp. 1842–1850. PMLR, 2016.

Shai Shalev-Shwartz and Shai Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014.

Jake Snell, Kevin Swersky, and Richard Zemel. Prototypical networks for few-shot learning. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pp. 4080–4090, 2017.

Sebastian Thrun. Lifelong learning algorithms. In *Learning to learn*, pp. 181–209. Springer.

Erik Torgersen. *Comparison of statistical experiments*, volume 36. Cambridge University Press, 1991.

Nilesh Tripuraneni, Michael I Jordan, and Chi Jin. On the theory of transfer learning: the importance of task diversity. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, pp. 7852–7862, 2020.

LG Valiant. A theory of the learnable. *Communications of the ACM*, 27(11):1134–1142, 1984.

Gido M Van de Ven, Tinne Tuytelaars, and Andreas S Tolias. Three types of incremental learning. *Nature Machine Intelligence*, 4(12):1185–1197, 2022.

Vladimir Vapnik. Statistical learning theory. *John Wiley & Sons google schola*, 2:831–842, 1998.

VN Vapnik and A Ya Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability & Its Applications*, 16(2):264–280, 1971.

Oriol Vinyals, Charles Blundell, Timothy Lillicrap, Daan Wierstra, et al. Matching networks for one shot learning. *Advances in neural information processing systems*, 29, 2016.

Karl Weiss, Taghi M Khoshgoftaar, and DingDing Wang. A survey of transfer learning. *Journal of Big data*, 3(1):9, 2016.

A Further Discussion of the Framework

This appendix expands on how the protocol–objective–resource framework should be interpreted and used. The main text states the formal definitions and main sample-complexity consequences; Appendix B gives the detailed proofs. The purpose here is more conceptual: to clarify what a reduction certifies, what kinds of transfer it supports, and where additional assumptions are still needed.

A.1 What a POR reduction certifies

A POR reduction is best read as a structural simulation certificate. It says that solving one learning interface is sufficient for solving another, provided that four obligations are met: environments must be embedded, learners must be compiled, success criteria must be translated, and resource budgets must be calibrated. Each obligation rules out a common source of informal ambiguity.

The environment embedding specifies which instances of the reduced paradigm are being represented inside the solver paradigm. The learner compiler specifies how an algorithm designed for the solver paradigm is actually run in the reduced paradigm. The threshold map specifies why success for the solver objective implies success for the reduced objective. The resource overhead specifies the cost of the simulation. A proposed comparison that omits one of these pieces may still be intuitively useful, but it is not yet a POR reduction.

This interpretation is deliberately algorithm-independent. The compiler acts on every admissible learner for the solver paradigm, not just on a particular algorithm. This is what makes the relation a hardness relation rather than a statement about one method. If every solver for paradigm A can be compiled into a solver for paradigm B, then the difficulty of B is already present inside A under the stated comparison class and accuracy regime.

A.2 Understanding transfer and method transfer

The framework separates two notions that are often conflated in practice. First, there is *understanding transfer*: structural facts about one paradigm, such as upper bounds, lower bounds, monotonicity, or impossibility results, may imply corresponding facts about another paradigm. This is the role of the sample-complexity transfer theorem and its lower-bound consequence. Once a reduction is established, a theorem proved in one setting can be interpreted in the other without reproving the entire result from scratch.

Second, there is *method transfer*: an actual learning procedure designed for one paradigm may be reused in another. A POR reduction makes explicit what has to be preserved for this reuse to be valid. The target interaction must provide the transcript required by the method, the success criterion in the source setting must imply the desired criterion in the target setting, and all additional costs must be counted. This is why shared architectures or optimizers do not by themselves establish method transfer. The protocol, objective, and resources must also align.

This distinction is useful when diagnosing empirical claims. If a transfer method fails, the failure may come from representation, but it may also come from a broken protocol simulation, a mismatch between training and evaluation objectives, or an unaccounted resource such as replay memory, adaptation queries, or task identifiers. The POR language gives a checklist for locating which part of the claimed transfer is unsupported.

A.3 The role of comparison classes

Comparison classes are essential because hardness is not a property of a paradigm name alone. Transfer learning with unrelated source and target tasks is different from transfer learning with a shared representation. Continual learning with adversarial task switches is different from continual learning with slowly drifting tasks. Meta-learning over a broad family of unrelated tasks is different from meta-learning over a tightly structured task family.

The same named paradigm can therefore occupy different places in the hardness preorder depending on the comparison class. This is a feature rather than a defect. It prevents the framework from making overly broad

statements such as "transfer learning is always easier" or "continual learning is always harder." Instead, a comparison must state which environments are being compared and how they are embedded.

In applications, the comparison class is where domain assumptions enter. A shared representation assumption, a task-relatedness assumption, a smooth drift assumption, or a bounded-memory assumption should be encoded in the environment class or learner class before a reduction is claimed. This makes the statistical content of the comparison visible.

A.4 Accuracy regimes and meaningful objectives

Accuracy regimes prevent formally correct but uninformative reductions. Without restricting the thresholds under consideration, one could map every solver requirement to a threshold that is automatically satisfied with no data. Such a map would satisfy a formal implication but would not express a meaningful hardness relation.

Proposition A.1 (Vacuity from loose thresholds)**.** *Suppose the accuracy regime $\mathscr{T}_B$ for B contains a threshold $\bar{\tau}_B$ for which there exists a learner $T_B^0 \in \mathfrak{T}_B$ satisfying*

$$\Pr[\mathbf{J}_B(T_B^0, e; 0) \preceq \bar{\tau}_B] = 1 \qquad \textit{for all } e \in \mathcal{E}_B.$$

If $\mathcal{E}_A$ and $\mathscr{T}_A$ are nonempty, then $B \preceq_{\mathrm{ex}} A$ relative to $(\mathscr{T}_A, \mathscr{T}_B)$ for every paradigm A.

Description and significance. This proposition shows that if a permitted threshold can be met with zero samples, then B exactly reduces to every nonempty paradigm A. The result explains why accuracy regimes are essential: without excluding vacuous thresholds, the framework would certify comparisons that carry no substantive learning content.

Proof deferred to Appendix B.1.

The regime should therefore be chosen to reflect the scientific question. For scalar risk, a regime might contain only thresholds below a baseline. For continual learning, a regime might require simultaneous control of final-task risk and forgetting. For meta-learning, a regime might focus on post-adaptation query risk after a fixed support budget. Different regimes can lead to different valid reductions because they ask different questions about success.

This also explains the objective-strengthening examples in the main text. If one paradigm requires control of only one performance coordinate and another requires control of all coordinates, then the second objective can be strictly harder even when the protocol and learner class are unchanged. The difference comes from what success means, not from what data are observed.

A.5 Resource accounting

The resource component is the mechanism that keeps reductions honest. A simulation may require additional labeled samples, source tasks, support examples, adaptation queries, memory, replay buffers, communication, or computation. The present paper focuses on sample complexity and treats other resource restrictions through the admissible learner class, but the same principle applies more broadly: a reduction should state what is being spent.

This point is especially important for method transfer. For example, a continual-learning method that uses a large replay buffer may look comparable to a memoryless method if only labeled examples are counted. A meta-learning method may appear sample-efficient if support and query examples are charged under different conventions. A transfer-learning method may appear to improve target performance while relying on a large amount of source data. POR reductions force these choices into the statement of the comparison.

Future versions of the framework could make the resource vector more explicit, tracking samples, memory, computation, and adaptation rounds simultaneously. The affine-overhead form used here is a first step: it captures sample-budget calibration while keeping the main definitions simple.

A.6 Refinements within a fixed paradigm

The framework is equally useful for comparing *variants* of the same named paradigm. Two common axes are side information and memory budget.

Proposition A.2 (Replay-budget monotonicity in continual learning)**.** *Fix a continual-learning environment class and performance functional. For each replay budget $r \in \mathbb{N}$, let $\mathrm{Cont}(r)$ denote the continual-learning paradigm whose admissible learners may store at most r examples in replay memory. If $r \leq r'$, then*

$$\mathrm{Cont}(r) \succeq_{\mathrm{ex}} \mathrm{Cont}(r').$$

Description and significance. The result states that a continual learner with a larger replay budget can simulate any learner with a smaller replay budget. It is significant because replay memory becomes an ordered resource in the POR framework, allowing variants of continual learning to be compared by their admissible learner classes.

Proof deferred to Appendix B.12.

Proposition A.3 (Task-identifier side information in continual learning)**.** *Fix a continual-learning environment class, learner class, performance functional, and sample accounting rule. Let $\mathrm{Cont}^{\mathrm{id}}$ denote the version of the protocol that reveals task identifiers with each observed example, and let $\mathrm{Cont}^{\mathrm{blind}}$ denote the version that does not reveal these identifiers. Then*

$$\mathrm{Cont}^{\mathrm{blind}} \succeq_{\mathrm{ex}} \mathrm{Cont}^{\mathrm{id}}.$$

Description and significance. This proposition formalizes that task-identifier side information cannot make the continual-learning problem harder, since a task-aware learner may ignore the identifiers. The result clarifies how protocol-level side information induces a hardness ordering even inside a single named paradigm.

Proof deferred to Appendix B.13.

Remark A.4 (Design implications)*.* Propositions A.2 and A.3 illustrate a point that is easy to miss if one speaks only at the level of broad paradigm labels: many practically important design choices themselves induce hardness orderings. The POR language therefore provides a way to compare not only supervised versus transfer versus meta-learning, but also task-aware versus task-agnostic continual learning, or small-memory versus large-memory variants of the same setup.

Corollary A.5 (Quantitative monotonicity for continual-learning variants)**.** *Under the assumptions of Propositions A.2 and A.3, the corresponding minimax sample complexities satisfy the monotone inequalities*

$$N^{\star}_{\mathrm{Cont}(r')}(\mathcal{C};\tau,\delta) \leq N^{\star}_{\mathrm{Cont}(r)}(\mathcal{C};\tau,\delta) \qquad \textit{whenever } r \leq r',$$

and

$$N^{\star}_{(\mathrm{Cont}^{\mathrm{id}})}(\mathcal{C};\tau,\delta) \leq N^{\star}_{(\mathrm{Cont}^{\mathrm{blind}})}(\mathcal{C};\tau,\delta).$$

Description and significance. The corollary converts the replay and task-identifier monotonicity reductions into explicit minimax sample-complexity inequalities. Its significance is that practical design choices such as memory and side information lead to quantitative consequences, not only qualitative intuitions.

Proof deferred to Appendix B.14.

A.7 What the framework does not claim

The framework gives worst-case structural statements. It does not claim that two paradigms have the same optimization difficulty, empirical performance, or typical-case behavior. It also does not say that a harder paradigm is always less useful in practice. A harder paradigm may provide more information, more flexible objectives, or better inductive structure for a particular application.

Nor does the framework impose a universal total order over supervised, transfer, continual, and meta-learning. Some comparisons require additional assumptions, and some may be incomparable under natural regimes. This is why the formal relation is a preorder. The goal is to identify valid structural relations where they exist, and to explain precisely which obligation fails when they do not.

A.8 Possible extensions

Several extensions are natural. One direction is to develop approximate reductions that allow controlled slack in the performance threshold, not only in the sample budget. This would better match empirical practice, where one often accepts small losses in accuracy when moving a method across settings.

A second direction is to study randomized or distribution-dependent reductions. The current definition is uniform over the comparison class. In some applications, one may want reductions that hold for most environments under a meta-distribution, or reductions that adapt to observable structure in the task family.

A third direction is to combine POR reductions with sharper within-paradigm lower bounds. The framework is most useful when paired with concrete statistical results: once a lower bound is known for one paradigm, a reduction can move that lower bound to another. Conversely, a separation in sample complexity can rule out a proposed structural comparison. In this sense, the framework is not a replacement for problem-specific learning theory, but a language for transporting it across paradigms.

B Detailed Proofs

This appendix collects the detailed proofs for results stated in the main text and in Appendix A. The proofs use the notation and standing assumptions introduced in Section 2.

B.1 Proof of Proposition A.1

Proof. We verify the POR-reduction definition directly. Since $\mathcal{E}_A$ is nonempty, choose an arbitrary reference environment $e_A^0 \in \mathcal{E}_A$, and define the embedding

$$\Phi(e) := e_A^0 \qquad \text{for every } e \in \mathcal{E}_B.$$

Since $\mathscr{T}_A$ is nonempty and $\bar{\tau}_B \in \mathscr{T}_B$, define the threshold map

$$\Gamma(\tau_A) := \bar{\tau}_B \qquad \text{for every } \tau_A \in \mathscr{T}_A.$$

The only point requiring care is the budget in the conclusion of the reduction. The learner T_B^0 is known to satisfy the B-threshold at zero samples. To use the same zero-sample behavior under an arbitrary external budget M, let $q_0(M) = 0$ for all M. By the budget-capping closure in Assumption 2.1, the capped learner $(T_B^0)^{q_0}$ is admissible for B, and

$$\mathbf{J}_B((T_B^0)^{q_0}, e; M) = \mathbf{J}_B(T_B^0, e; 0)$$

as random variables for every e and every M. Define the compiler by

$$\mathfrak{C}(T_A) := (T_B^0)^{q_0} \qquad \text{for every } T_A \in \mathfrak{T}_A.$$

Now fix arbitrary $e \in \mathcal{E}_B$, $T_A \in \mathfrak{T}_A$, $\tau_A \in \mathscr{T}_A$, $\delta \in (0,1)$, and $M \in \mathbb{N}$. Regardless of whether the antecedent success event for A holds, the compiled learner satisfies

$$\Pr[\mathbf{J}_B(\mathfrak{C}(T_A), e; M) \preceq \Gamma(\tau_A)] = \Pr\big[\mathbf{J}_B(T_B^0, e; 0) \preceq \bar{\tau}_B\big] = 1.$$

Thus the POR implication holds with overhead $a = 1$, $b = 0$. This is an exact reduction. Because the construction ignores A entirely, the reduction carries no meaningful hardness information; it is precisely the vacuity that accuracy regimes are meant to exclude. □

B.2 Proof of Proposition 3.1

Proof. The assumed implication runs the pre-compiled learner $\mathfrak{C}_0(T_A)$ with internal budget $\psi(M)$. The POR definition, however, supplies the compiled B-learner with the external budget $aM + b$. We therefore cap the compiled learner so that, whenever the external budget equals $aM + b$, it uses exactly the internal budget $\psi(M)$.

For an external budget $M' \in \mathbb{N}$, define

$$m(M') := \max\{m \in \mathbb{N} : am + b \le M'\},$$

with $m(M') = 0$ if the displayed set is empty. Define

$$q(M') := \begin{cases} \psi(m(M')), & M' \ge b, \\ 0, & M' < b. \end{cases}$$

We first check that q is a valid cap. If $M' < b$, then $q(M') = 0 \le M'$. If $M' \ge b$, then $am(M') + b \le M'$ by the definition of $m(M')$. Since $\psi(m) \le am + b$ for every m, we get

$$q(M') = \psi(m(M')) \le am(M') + b \le M'.$$

Thus $q(M') \le M'$ for all external budgets M'. By Assumption 2.1, the learner

$$\mathfrak{C}(T_A) := (\mathfrak{C}_0(T_A))^q$$

is admissible for B, and its performance at external budget M' equals the performance of $\mathfrak{C}_0(T_A)$ at internal budget $q(M')$.

Now set $M' = aM + b$. Since $a \ge 1$, the condition

$$am' + b \le aM + b$$

is equivalent to $m' \le M$, and hence $m(aM + b) = M$. Therefore

$$q(aM + b) = \psi(M).$$

For arbitrary $e, T_A, \tau_A, \delta, M$, if

$$\Pr[\mathbf{J}_A(T_A, \Phi(e); M) \preceq \tau_A] \ge 1 - \delta,$$

then the assumed budget-conversion implication gives

$$\Pr[\mathbf{J}_B(\mathfrak{C}_0(T_A), e; \psi(M)) \preceq \Gamma(\tau_A)] \ge 1 - \delta.$$

By the equality of the capped and uncapped performances at the relevant budgets, this is exactly

$$\Pr[\mathbf{J}_B(\mathfrak{C}(T_A), e; aM + b) \preceq \Gamma(\tau_A)] \ge 1 - \delta.$$

Thus the POR-reduction implication holds with overhead parameters a, b. □

B.3 Proof of Theorem 3.3

Proof. Write

$$N_A^\star := N_A^\star(\Phi(\mathcal{C}_B); \tau_A, \delta).$$

If $N_A^\star = \infty$, the desired inequality has an infinite right-hand side and is immediate under the extended-real convention. We therefore assume that $N_A^\star < \infty$.

Fix $\varepsilon > 0$. By the definition of the infimum in minimax sample complexity, there exists a learner $T_A \in \mathfrak{T}_A$ such that

$$\sup_{e' \in \Phi(\mathcal{C}_B)} N_A(T_A, e'; \tau_A, \delta) \le N_A^\star + \varepsilon.$$

Let $T_B := \mathfrak{C}(T_A)$. We claim that T_B solves every $e \in \mathcal{C}_B$ within budget

$$a\big(N_A^\star + \varepsilon\big) + b.$$

Fix $e \in \mathcal{C}_B$ and set

$$n_e := N_A(T_A, \Phi(e); \tau_A, \delta).$$

The previous display implies $n_e \leq N_A^\star + \varepsilon$, and in particular $n_e < \infty$. Since budgets range over $\mathbb{N}$, this finite infimum is attained. Thus, by the definition of n_e,

$$\Pr[\mathbf{J}_A(T_A, \Phi(e); n_e) \preceq \tau_A] \geq 1 - \delta.$$

Applying the POR-reduction implication at budget n_e gives

$$\Pr[\mathbf{J}_B(T_B, e; an_e + b) \preceq \Gamma(\tau_A)] \geq 1 - \delta.$$

Hence

$$N_B(T_B, e; \Gamma(\tau_A), \delta) \leq an_e + b \leq a\big(N_A^\star + \varepsilon\big) + b.$$

Taking the supremum over $e \in \mathcal{C}_B$ yields

$$\sup_{e \in \mathcal{C}_B} N_B(T_B, e; \Gamma(\tau_A), \delta) \leq aN_A^\star + b + a\varepsilon.$$

Finally, $N_B^\star$ is the infimum over all admissible B-learners, so

$$N_B^\star(\mathcal{C}_B; \Gamma(\tau_A), \delta) \leq aN_A^\star + b + a\varepsilon.$$

Letting $\varepsilon \downarrow 0$ gives the claimed domination inequality. □

B.4 Proof of Corollary 3.4

Proof. Theorem 3.3 gives

$$N_B^\star(\mathcal{C}_B; \Gamma(\tau_A), \delta) \leq a\, N_A^\star(\Phi(\mathcal{C}_B); \tau_A, \delta) + b.$$

If $N_A^\star(\Phi(\mathcal{C}_B); \tau_A, \delta) = \infty$, the claimed lower bound is automatic. Otherwise, the right-hand side in the theorem is finite, so the theorem also implies that $N_B^\star$ is finite. We may then subtract b and divide by the positive integer a, obtaining

$$N_A^\star(\Phi(\mathcal{C}_B); \tau_A, \delta) \geq \frac{N_B^\star(\mathcal{C}_B; \Gamma(\tau_A), \delta) - b}{a}.$$

Thus any lower bound on the reduced-paradigm quantity for B transfers to the embedded quantity for A, with the infinite case covered by the extended-real convention. □

B.5 Proof of Corollary 3.5

Proof. Exactness means that the POR-reduction has overhead parameters $a = 1$ and $b = 0$. Substituting these values into Theorem 3.3 gives

$$N_B^\star(\mathcal{C}_B; \Gamma(\tau_A), \delta) \leq N_A^\star(\Phi(\mathcal{C}_B); \tau_A, \delta),$$

which is the desired statement. □

B.6 Proof of Corollary 3.6

Proof. Assume, for contradiction, that a POR-reduction exists with the candidate data (Φ, Γ, a, b). Then Theorem 3.3 must hold for every comparison class, threshold, and confidence level. In particular, it would imply

$$N_B^\star(\mathcal{C}_B; \Gamma(\tau_A), \delta) \leq a\, N_A^\star(\Phi(\mathcal{C}_B); \tau_A, \delta) + b$$

for the class, threshold, and confidence level appearing in the corollary. This contradicts the displayed strict inequality in the statement. Therefore no reduction with those candidate data exists. If every possible embedding and threshold map admits such a witness, the same contradiction rules out every POR-reduction with the specified overhead parameters. □

B.7 Proof of Proposition 3.7

Proof. Because $\tau_P \preceq \tau'_P$, every outcome that satisfies the stricter threshold τ_P also satisfies the weaker threshold τ'_P:

$$\{\mathbf{J}_P(T, e; M) \preceq \tau_P\} \subseteq \{\mathbf{J}_P(T, e; M) \preceq \tau'_P\}.$$

Consequently, for every fixed learner T and environment e, every budget that makes T an (M, τ_P, δ)-solver also makes it an (M, τ'_P, δ)-solver. Hence

$$N_P(T, e; \tau_P, \delta) \geq N_P(T, e; \tau'_P, \delta).$$

Taking the supremum over $e \in \mathcal{C}_P$ preserves the inequality, and taking the infimum over $T \in \mathfrak{T}_P$ preserves it again. Therefore

$$N^\star_P(\mathcal{C}_P; \tau_P, \delta) \geq N^\star_P(\mathcal{C}_P; \tau'_P, \delta).$$

□

B.8 Proof of Proposition 3.8

Proof. Fix a learner $T \in \mathfrak{T}_P$. Since $\mathcal{C}_P \subseteq \mathcal{C}'_P$, the worst-case sample complexity over $\mathcal{C}_P$ is bounded by the worst-case sample complexity over the larger class:

$$\sup_{e \in \mathcal{C}_P} N_P(T, e; \tau_P, \delta) \leq \sup_{e \in \mathcal{C}'_P} N_P(T, e; \tau_P, \delta).$$

This inequality holds for every admissible learner T. Taking the infimum over T on both sides gives

$$N^\star_P(\mathcal{C}_P; \tau_P, \delta) \leq N^\star_P(\mathcal{C}'_P; \tau_P, \delta).$$

□

B.9 Proof of Proposition 3.9

Proof. It is enough to verify reflexivity and transitivity for the reduction relation $\preceq_{\mathrm{POR}}$, since $A \succeq_{\mathrm{POR}} B$ is defined to mean $B \preceq_{\mathrm{POR}} A$.

For reflexivity, fix a regime-equipped paradigm A. Use the identity environment map on $\mathcal{E}_A$, the identity compiler on $\mathfrak{T}_A$, the identity threshold map on $\mathscr{T}_A$, and overhead $a = 1, b = 0$. The antecedent and consequent in the POR implication are then the same event, so $A \preceq_{\mathrm{ex}} A$.

For transitivity, suppose

$$C \preceq_{\mathrm{POR}} B \quad \text{and} \quad B \preceq_{\mathrm{POR}} A.$$

Write the two reductions as

$$(\Phi_{CB}, \mathfrak{C}_{CB}, \Gamma_{CB}, a_{CB}, b_{CB})$$

for the reduction from C to B, and

$$(\Phi_{BA}, \mathfrak{C}_{BA}, \Gamma_{BA}, a_{BA}, b_{BA})$$

for the reduction from B to A. Define the composed data by

$$\Phi_{CA} := \Phi_{BA} \circ \Phi_{CB}, \qquad \mathfrak{C}_{CA} := \mathfrak{C}_{CB} \circ \mathfrak{C}_{BA}, \qquad \Gamma_{CA} := \Gamma_{CB} \circ \Gamma_{BA}.$$

The proposed overhead is

$$a_{CA} := a_{CB} a_{BA}, \qquad b_{CA} := a_{CB} b_{BA} + b_{CB}.$$

Fix $e \in \mathcal{E}_C$, $T_A \in \mathfrak{T}_A$, $\tau_A \in \mathscr{T}_A$, $\delta \in (0, 1)$, and $M \in \mathbb{N}$. Assume

$$\Pr[\mathbf{J}_A(T_A, \Phi_{CA}(e); M) \preceq \tau_A] \geq 1 - \delta.$$

Since $\Phi_{CA}(e) = \Phi_{BA}(\Phi_{CB}(e))$, the reduction $B \preceq_{\mathrm{POR}} A$ gives

$$\Pr[\mathbf{J}_B(\mathfrak{C}_{BA}(T_A), \Phi_{CB}(e); a_{BA}M + b_{BA}) \preceq \Gamma_{BA}(\tau_A)] \geq 1 - \delta.$$

Now apply the reduction $C \preceq_{\mathrm{POR}} B$ with learner $\mathfrak{C}_{BA}(T_A)$, threshold $\Gamma_{BA}(\tau_A)$, and budget $a_{BA}M + b_{BA}$. This yields

$$\Pr[\mathbf{J}_C(\mathfrak{C}_{CA}(T_A), e; a_{CB}(a_{BA}M + b_{BA}) + b_{CB}) \preceq \Gamma_{CA}(\tau_A)] \geq 1 - \delta.$$

The displayed budget equals $a_{CA}M + b_{CA}$, so the composed data define a POR-reduction $C \preceq_{\mathrm{POR}} A$. If both original reductions are exact, then $a_{CB} = a_{BA} = 1$ and $b_{CB} = b_{BA} = 0$, so the composed reduction is exact as well. □

B.10 Proof of Corollary 3.10

Proof. By definition, $A \equiv_{\mathrm{POR}} B$ means both $A \succeq_{\mathrm{POR}} B$ and $B \succeq_{\mathrm{POR}} A$. Reflexivity of $\equiv_{\mathrm{POR}}$ follows from reflexivity of $\succeq_{\mathrm{POR}}$. Symmetry is built into the definition. For transitivity, if $A \equiv_{\mathrm{POR}} B$ and $B \equiv_{\mathrm{POR}} C$, then transitivity of $\succeq_{\mathrm{POR}}$ gives $A \succeq_{\mathrm{POR}} C$ and $C \succeq_{\mathrm{POR}} A$, so $A \equiv_{\mathrm{POR}} C$.

It remains to check that the induced order on equivalence classes is well defined. Suppose $A \equiv_{\mathrm{POR}} A'$, $B \equiv_{\mathrm{POR}} B'$, and $A \succeq_{\mathrm{POR}} B$. Then

$$A' \succeq_{\mathrm{POR}} A \succeq_{\mathrm{POR}} B \succeq_{\mathrm{POR}} B',$$

so $A' \succeq_{\mathrm{POR}} B'$. Thus the comparison does not depend on the chosen representatives. Reflexivity and transitivity descend from the preorder. Antisymmetry holds because mutual comparability of two classes means their representatives are POR-equivalent, and hence the two classes are the same. □

B.11 Proof of Proposition 3.11

Proof. We verify each template against the POR definition.

For (1), use the stated embedding Φ, compiler $\mathfrak{C}$, threshold map Γ, and overhead $a = 1, b = 0$. Fix e, T, τ_A, δ, M. If the A-success event has probability at least $1 - \delta$, then the assumed event inclusion, up to null events, implies

$$\Pr[\mathbf{J}_B(\mathfrak{C}(T), e; M) \preceq \Gamma(\tau_A)] \geq \Pr[\mathbf{J}_A(T, \Phi(e); M) \preceq \tau_A] \geq 1 - \delta.$$

This is exactly $B \preceq_{\mathrm{ex}} A$, or equivalently $A \succeq_{\mathrm{ex}} B$.

For (2), the environments, protocols, learners, and resource accounting are the same. Therefore the identity embedding and identity compiler are admissible. With the given Γ, the event-inclusion argument from (1) applies verbatim, so the reduction is exact.

For (3), the two paradigms differ only in that B's protocol reveals extra side information. Given any learner T for the less informative paradigm A, define $\mathfrak{C}(T)$ to run the same decision rule under B's protocol while discarding the additional side information. After discarding that information, the transcript seen by $\mathfrak{C}(T)$ under B has the same law as the transcript seen by T under A. The objective, environment class, resource accounting, and regime are the same, so the success events agree in distribution under the identity threshold map. This gives an exact reduction.

For (4), every learner admissible for A is also admissible for B. The identity map on learners is therefore a valid compiler from $\mathfrak{T}_A$ to $\mathfrak{T}_B$. Since all other components are the same, the success events are identical under the identity embedding and identity threshold map. The sample budget is unchanged, so the reduction is exact. □

B.12 Proof of Proposition A.2

Proof. The paradigms $\mathrm{Cont}(r)$ and $\mathrm{Cont}(r')$ have the same environment class, observation protocol, performance functional, and sample accounting rule. Their learner classes differ only in the replay-memory

constraint. If $r \le r'$, any learner that stores at most r examples also stores at most r' examples, so

$$\mathfrak{T}_{\mathrm{Cont}(r)} \subseteq \mathfrak{T}_{\mathrm{Cont}(r')}.$$

Proposition 3.11(4), applied with $A = \mathrm{Cont}(r)$ and $B = \mathrm{Cont}(r')$, gives

$$\mathrm{Cont}(r) \succeq_{\mathrm{ex}} \mathrm{Cont}(r').$$

□

B.13 Proof of Proposition A.3

Proof. The two continual-learning variants have the same environments, learner class, performance functional, sample accounting rule, and accuracy regime. The only difference is the observation protocol: $\mathrm{Cont}^{\mathrm{id}}$ gives the learner task identifiers, while $\mathrm{Cont}^{\mathrm{blind}}$ does not. A learner designed for the blind protocol can be run under the identifier-revealing protocol by ignoring the identifiers. The resulting transcript has the same law as in the blind protocol. Therefore Proposition 3.11(3) applies and yields

$$\mathrm{Cont}^{\mathrm{blind}} \succeq_{\mathrm{ex}} \mathrm{Cont}^{\mathrm{id}}.$$

□

B.14 Proof of Corollary A.5

Proof. For replay budgets, Proposition A.2 gives the exact reduction

$$\mathrm{Cont}(r') \preceq_{\mathrm{ex}} \mathrm{Cont}(r) \qquad \text{whenever } r \le r'.$$

Applying Corollary 3.5 with the identity embedding and identity threshold map gives

$$N^{\star}_{\mathrm{Cont}(r')}(\mathcal{C};\tau,\delta) \le N^{\star}_{\mathrm{Cont}(r)}(\mathcal{C};\tau,\delta).$$

Similarly, Proposition A.3 gives

$$\mathrm{Cont}^{\mathrm{id}} \preceq_{\mathrm{ex}} \mathrm{Cont}^{\mathrm{blind}}.$$

The exact-domination corollary therefore implies

$$N^{\star}_{(\mathrm{Cont}^{\mathrm{id}})}(\mathcal{C};\tau,\delta) \le N^{\star}_{(\mathrm{Cont}^{\mathrm{blind}})}(\mathcal{C};\tau,\delta).$$

□

B.15 Proof of Proposition 3.12

Proof. First, $P_{\mathrm{all}} \succeq_{\mathrm{ex}} P_{\mathrm{one}}$. The two paradigms have the same environments, protocol, learners, and sample accounting. Moreover,

$$\{\mathbf{J}_{\mathrm{all}}(T, D_\sigma; M) \preceq (0,0)\} \subseteq \{\mathbf{J}_{\mathrm{one}}(T, D_\sigma; M) \preceq 0\}.$$

Thus Proposition 3.11(2) gives the exact reduction $P_{\mathrm{one}} \preceq_{\mathrm{ex}} P_{\mathrm{all}}$.

We now compute the two minimax quantities.

For P_{one}, consider the learner that memorizes the label of input 1 if input 1 is observed, and otherwise guesses that label uniformly at random. The event that input 1 is never observed in M samples has probability 2^{-M}. Conditional on this event, the learner is wrong with probability $1/2$; otherwise it is correct. Hence its failure probability is $2^{-(M+1)}$, uniformly over σ, and its success probability is

$$1 - 2^{-(M+1)}.$$

No learner can have a better worst-case success probability. To see this, place the uniform prior on σ_1 while fixing σ_2. On the event that input 1 is not observed, the entire transcript is independent of σ_1. Therefore,

conditioned on that event, any estimator of σ_1 has average success probability at most $1/2$. Since the worst-case success probability of a learner is no larger than its average success probability under any prior, every learner has worst-case failure at least $2^{-(M+1)}$. The memorize-and-guess learner is therefore minimax optimal for P_{one}.

It follows that success probability at least $1-\delta$ is possible exactly when

$$2^{-(M+1)} \leq \delta,$$

or equivalently

$$M \geq \log_2 \frac{1}{2\delta}.$$

Since M is an integer,

$$N^{\star}_{P_{\text{one}}}(\mathcal{C}; 0, \delta) = \left\lceil \log_2 \frac{1}{2\delta} \right\rceil.$$

For P_{all}, use the learner that memorizes all observed labels and guesses each unobserved label uniformly. Since $\delta < 1/2$, a zero-sample budget cannot be sufficient; indeed, with no samples the best success probability for two unknown bits is $1/4$. For $M \geq 1$, either both inputs have been observed or exactly one input has been observed. The second case occurs precisely when all M samples fall on the same input, which has probability $2 \cdot 2^{-M} = 2^{1-M}$. Conditional on this case, the unobserved label must be guessed, causing failure with probability $1/2$. Thus the memorize-and-guess learner has failure probability

$$\frac{1}{2} 2^{1-M} = 2^{-M},$$

and success probability $1 - 2^{-M}$.

The same prior argument proves optimality. Put the uniform prior on $\sigma = (\sigma_1, \sigma_2)$. Whenever one input is unobserved, its label is still uniform and independent of the transcript, so any learner has conditional average probability at most $1/2$ of predicting both labels correctly. Hence no learner can have worst-case failure below 2^{-M} for $M \geq 1$. Therefore success probability at least $1-\delta$ is possible exactly when

$$2^{-M} \leq \delta,$$

and

$$N^{\star}_{P_{\text{all}}}(\mathcal{C}; (0,0), \delta) = \left\lceil \log_2 \frac{1}{\delta} \right\rceil.$$

Finally, for $\delta \in (0, 1/2)$,

$$\left\lceil \log_2 \frac{1}{\delta} \right\rceil = \left\lceil \log_2 \frac{1}{2\delta} + 1 \right\rceil = \left\lceil \log_2 \frac{1}{2\delta} \right\rceil + 1.$$

Thus the all-coordinate objective has strictly larger minimax sample complexity.

It remains to rule out exact equivalence. Suppose, toward contradiction, that there were an exact reverse reduction $P_{\text{all}} \preceq_{\text{ex}} P_{\text{one}}$. The regimes are singletons, so the threshold map must send 0 to $(0,0)$. For any environment embedding $\Phi : \mathcal{C} \to \mathcal{C}$, class monotonicity gives

$$N^{\star}_{P_{\text{one}}}(\Phi(\mathcal{C}); 0, \delta) \leq N^{\star}_{P_{\text{one}}}(\mathcal{C}; 0, \delta) < N^{\star}_{P_{\text{all}}}(\mathcal{C}; (0,0), \delta).$$

This contradicts Corollary 3.6 with overhead $(1,0)$. Hence no exact reverse reduction exists, and the two paradigms are not exact-POR equivalent relative to the stated regimes. □

B.16 Proof of Proposition 3.13

Proof. We first prove the upper bound for $P^{(m)}_{\text{one}}$. Use the learner that memorizes the label of input 1 if it is observed and otherwise guesses that label uniformly. The probability that input 1 is not observed in M samples is $(1-1/m)^M$, so the failure probability is

$$\frac{1}{2}\left(1 - \frac{1}{m}\right)^M.$$

At $M = 2m$,

$$\frac{1}{2}\left(1 - \frac{1}{m}\right)^{2m} \leq \frac{1}{2}e^{-2} < \frac{1}{8}.$$

Thus

$$N^{\star}_{P^{(m)}_{\text{one}}}(\mathcal{C}_m; 0, 1/8) \leq 2m.$$

We next prove the lower bound for $P^{(m)}_{\text{all}}$. Let Z be the number of inputs in $\mathcal{X}_m$ not observed after M samples. Place the uniform prior on $\sigma \in \{0,1\}^m$. Conditional on any transcript with $Z > 0$, at least one label remains completely unobserved and is still uniform given the transcript. Therefore any learner has conditional average probability at most $1/2$ of predicting every label correctly. Consequently, for every learner,

$$\text{worst-case success} \leq \text{average success under the prior} \leq 1 - \frac{1}{2}\Pr[Z > 0].$$

If $\Pr[Z > 0] \geq 1/2$, then the worst-case success probability is at most $3/4$, which is below the required $7/8$. It remains to show that $\Pr[Z > 0] \geq 1/2$ whenever

$$M \leq \frac{m}{4}\log m.$$

Let I_j be the indicator that input j is unobserved, so $Z = \sum_{j=1}^{m} I_j$. Then

$$\mathbb{E}Z = m\left(1 - \frac{1}{m}\right)^{M}.$$

Since $m \geq 3$, $1/m \leq 1/2$, and the elementary inequality $\log(1-u) \geq -2u$ holds for $u \in (0, 1/2]$. Therefore

$$\mathbb{E}Z \geq m\exp\left(-\frac{2M}{m}\right) \geq m\exp\left(-\frac{1}{2}\log m\right) = \sqrt{m}.$$

For the second moment,

$$\mathbb{E}Z^2 = \sum_{j=1}^{m}\mathbb{E}I_j + \sum_{i\neq j}\mathbb{E}[I_i I_j].$$

Here

$$\mathbb{E}[I_i I_j] = \left(1 - \frac{2}{m}\right)^{M} \leq \left(1 - \frac{1}{m}\right)^{2M},$$

because $1 - 2/m \leq (1 - 1/m)^2$. Hence

$$\mathbb{E}Z^2 \leq \mathbb{E}Z + (\mathbb{E}Z)^2.$$

By Cauchy's inequality,

$$\Pr[Z > 0] \geq \frac{(\mathbb{E}Z)^2}{\mathbb{E}Z^2} \geq \frac{\mathbb{E}Z}{1 + \mathbb{E}Z} \geq \frac{1}{2},$$

where the last step uses $\mathbb{E}Z \geq \sqrt{m} \geq 1$. Thus no learner can achieve success probability $7/8$ using $M \leq (m/4)\log m$ samples, proving

$$N^{\star}_{P^{(m)}_{\text{all}}}(\mathcal{C}_m; \mathbf{0}_m, 1/8) \geq \frac{m}{4}\log m.$$

Finally, suppose that for every m there were a reverse reduction

$$P^{(m)}_{\text{all}} \preceq_{\text{POR}} P^{(m)}_{\text{one}}$$

with the same affine overhead parameters a, b, independent of m. Since the regimes are singletons, the threshold map must send 0 to $\mathbf{0}_m$. For any embedding $\Phi : \mathcal{C}_m \to \mathcal{C}_m$, class monotonicity and the upper bound above imply

$$N^{\star}_{P^{(m)}_{\text{one}}}(\Phi(\mathcal{C}_m); 0, 1/8) \leq N^{\star}_{P^{(m)}_{\text{one}}}(\mathcal{C}_m; 0, 1/8) \leq 2m.$$

Theorem 3.3 would therefore force

$$N^{\star}_{P^{(m)}_{\text{all}}}(\mathcal{C}_m; \mathbf{0}_m, 1/8) \le 2am + b.$$

But the lower bound just proved gives

$$N^{\star}_{P^{(m)}_{\text{all}}}(\mathcal{C}_m; \mathbf{0}_m, 1/8) \ge \frac{m}{4} \log m,$$

and $(m/4) \log m > 2am + b$ for all sufficiently large m. This contradiction rules out uniform affine-overhead reverse reductions. □

B.17 Proof of Proposition 4.4

Proof. The desired hardness statement Trans $\succeq_{\text{ex}}$ Sup is equivalent to the exact reduction

$$\text{Sup} \preceq_{\text{ex}} \text{Trans}.$$

We construct the four pieces of POR data. By Assumption 4.1, every supervised environment D can be embedded as the transfer environment

$$\Phi_{\text{Sup}\to\text{Trans}}(D) := (\varnothing; D).$$

Given a transfer learner T_{Trans}, define the compiled supervised learner $\mathfrak{C}(T_{\text{Trans}})$ as follows: on samples from D, run T_{Trans} on the transfer instance $(\varnothing; D)$, provide no source samples, forward each supervised sample as a target sample, and output the predictor produced by T_{Trans}. Use the identity threshold map on the scalar risk regime and overhead $a = 1, b = 0$.

Under this construction, for every D, T_{Trans}, and budget M, the target-risk random variable in the embedded transfer problem is the same as the supervised-risk random variable of the compiled learner:

$$\mathbf{J}_{\text{Sup}}(\mathfrak{C}(T_{\text{Trans}}), D; M) = \mathbf{J}_{\text{Trans}}(T_{\text{Trans}}, \Phi_{\text{Sup}\to\text{Trans}}(D); M).$$

Therefore success in the transfer objective at threshold τ implies success of the compiled supervised learner at the same threshold and the same budget. This verifies an exact POR-reduction. □

B.18 Proof of Proposition 4.5

Proof. The statement Cont $\succeq_{\text{ex}}$ Trans is equivalent to

$$\text{Trans} \preceq_{\text{ex}} \text{Cont}.$$

Let

$$e = (D_1, \ldots, D_K; D_\star)$$

be a transfer-learning environment. By Assumption 4.2, the embedding

$$\Phi_{\text{Trans}\to\text{Cont}}(e) := (D_1, \ldots, D_K, D_\star)$$

is a valid continual-learning environment, and the continual objective contains the final-task risk $R_{D_\star}$ as one of its coordinates.

Given a continual learner T_{Cont}, define the compiled transfer learner $\mathfrak{C}(T_{\text{Cont}})$ by feeding the transfer source tasks and then the target task to T_{Cont} as the sequential stream $(D_1, \ldots, D_K, D_\star)$, and then returning the final predictor. Let Γ project a continual-learning threshold vector to the coordinate that bounds the final-task risk. Use overhead $a = 1, b = 0$.

If

$$\mathbf{J}_{\text{Cont}}(T_{\text{Cont}}, \Phi_{\text{Trans}\to\text{Cont}}(e); M) \preceq \tau_{\text{Cont}},$$

then, in particular, the final-task risk coordinate is at most $\Gamma(\tau_{\text{Cont}})$. That coordinate is exactly the transfer objective of the compiled learner. Hence

$$\{\mathbf{J}_{\text{Cont}}(T_{\text{Cont}}, \Phi_{\text{Trans}\to\text{Cont}}(e); M) \preceq \tau_{\text{Cont}}\} \subseteq \{\mathbf{J}_{\text{Trans}}(\mathfrak{C}(T_{\text{Cont}}), e; M) \preceq \Gamma(\tau_{\text{Cont}})\}.$$

The event inclusion verifies an exact POR-reduction by the special-case embedding template. □

B.19 Proof of Corollary 4.6

Proof. Proposition 4.4 gives

$$\text{Trans} \succeq_{\text{ex}} \text{Sup},$$

and Proposition 4.5 gives

$$\text{Cont} \succeq_{\text{ex}} \text{Trans}.$$

By the transitivity part of Proposition 3.9,

$$\text{Cont} \succeq_{\text{ex}} \text{Sup}.$$

□

B.20 Proof of Proposition 4.7

Proof. The desired statement is equivalent to

$$\text{Sup} \preceq_{\text{ex}} \text{Meta}.$$

By Assumption 4.3, every supervised environment D has a singleton-task meta-learning embedding

$$\Phi_{\text{Sup}\to\text{Meta}}(D) := \Pi_D = \delta_D.$$

Given a meta learner T_{Meta}, define the compiled supervised learner $\mathfrak{C}(T_{\text{Meta}})$ by partitioning its observed samples from D into the support/query interface required by the meta protocol and then outputting the adapted predictor produced by T_{Meta}. The closure assumption ensures that this use of supervised samples is admissible and that sample units are aligned. Use the identity threshold map and overhead $(1, 0)$.

Under the singleton task distribution Π_D, every episode is drawn from the same underlying task D. Therefore the post-adaptation query risk of the meta learner on Π_D is the same population risk on D controlled by the compiled supervised learner. Thus the success events agree under the embedding at the same sample budget. This verifies the exact POR-reduction $\text{Sup} \preceq_{\text{ex}} \text{Meta}$, equivalently $\text{Meta} \succeq_{\text{ex}} \text{Sup}$. □

B.21 Proof of Corollary 4.9

Proof. Each inequality is an instance of exact sample-complexity domination.

First, Proposition 4.4 gives $\text{Sup} \preceq_{\text{ex}} \text{Trans}$ with embedding $\Phi_{\text{Sup}\to\text{Trans}}$ and the identity threshold map. Applying Corollary 3.5 gives

$$N^{\star}_{\text{Sup}}(\mathcal{C}_{\text{Sup}}; \tau_{\text{Sup}}, \delta) \leq N^{\star}_{\text{Trans}}(\Phi_{\text{Sup}\to\text{Trans}}(\mathcal{C}_{\text{Sup}}); \tau_{\text{Sup}}, \delta).$$

Second, Corollary 4.6 gives the composed exact reduction from supervised learning to continual learning. The composed environment embedding is

$$\Phi_{\text{Trans}\to\text{Cont}} \circ \Phi_{\text{Sup}\to\text{Trans}},$$

and the relevant threshold map is the final-task projection. Thus, for every continual threshold τ_{Cont} whose final-task projection equals τ_{Sup}, exact domination gives the displayed continual inequality.

Third, Proposition 4.7 gives $\text{Sup} \preceq_{\text{ex}} \text{Meta}$ with embedding $\Phi_{\text{Sup}\to\text{Meta}}$ and the identity scalar-risk threshold map. Applying Corollary 3.5 gives the displayed meta-learning inequality. □