A unified model of children's early mathematical learning as analogy-based programming language synthesis

Working Draft

Abstract

One of the enigmas of human cognitive development is that it is marked by moments of *discontinuity*. In landmark events during childhood, children experience discrete jumps in which they gain sudden access to fundamentally new bases of thought, enabling previously impossible expression. The classic example of these discontinuities is children's acquisition of number concepts, where the cross-cultural signature is their dramatic understanding of all number words in their count list after initial piecemeal learning (the *cardinal principle* or *CP* induction). Despite longstanding interest, a computational account of these conceptual transitions that answers the deepest nativist objections to discontinuity remains elusive. Here, we present such an explanation by mathematically formalizing Carey's theory of "Quinian bootstrapping." Building upon models based on program induction, we develop an enriched paradigm in which discontinuous conceptual change becomes the invention of a *new type system*, related by *structural analogy* to the previous representation system, but also, crucially, generalizing beyond it. We instantiate our method in the archetypal domain of children's early numerical development, modeling the acquisition of both natural and rational number. We find our model captures several previously unmodeled phenomena from the empirical literature—including the multiple stages of the often atomically-viewed CP induction, the discovery of discrete infinity, and the discovery of infinite divisibility—and provides more parsimonious explanations for cross-linguistic universals and intervention effects. More generally, our approach unifies disparate observations about children's early mathematical learning in a simple analogy-based symbolic model, laying the foundation for explaining similar discontinuous paradigm shifts in other developmental domains. It further makes key recommendations for STEM curriculum design, particularly fraction curriculum design, based on its crystallization of the mechanism by which symbolic analogies efficiently induce conceptual generalization. Finally, perhaps most provocatively, it provides a precise computational-level grounding of the uniquely human capacity for genuine *extrapolation*—the ability to think truly new thoughts—offering a key counterpoint to an AI zeitgeist fueled by models based on statistical interpolation.

Introduction

A longtime dream of research in computational cognitive science is achieving a formal model of children's cognitive development: How do children grow increasingly rich systems for representing and reasoning about the world as they age? This evolution is shaped by two knowledge systems developed over longer timescales. The first is *biologically innate* representation systems, such as functionally-specific brain regions that implement systems of *core knowledge*, developed over evolutionary time and common across cultures. The second is *cultural* knowledge systems, such as symbolic language and mathematics, developed by human civilizations over the course of historical time. Here, cross-cultural differences can lead to variation in the pace and outcomes of children's developmental trajectories, sometimes leading to adult populations with markedly different cognitive abilities (Frank et al., 2008).

In this paper, we focus on a particularly transformative kind of developmental change known as a *discontinuous conceptual* change. These transitions are landmark events after which previously inexpressible thoughts and theories, *incommensurable* with the initial cognitive repertoire, become accessible. An archetypal example is children's acquisition of number concepts, where the developmental signature across industrialized cultures is children's sudden understanding of all number words in their count list after initial incremental learning. Remarkably, evidence from cognitive-historical analysis suggests that the cognitive mechanisms underlying these childhood discontinuities share many similarities with those metacognitively used by scientists who have created discontinuity in *scientific knowledge*, like Kepler and Darwin (Carey, 2009). A unified model that crystallizes the mechanism operating in both of these settings would be a significant achievement.

Despite longstanding interest, such a satisfying computational explanation remains elusive. Past attempts to model conceptual change based on *program induction* (e.g., Piantadosi et al., 2012, Ellis et al., 2021) have made key progress, explaining how new structured representations can be learned. However, they have not addressed the deepest nativist objections to discontinuity, because they build in initial access to the mental primitives—states and functions—needed to express the conceptual system ultimately developed (Fodor, 1998).

Here we present an algorithmic-level explanation of how genuinely new conceptual systems can be constructed via a formalization of *Quinian bootstrapping*, the informal class of mechanisms posited by Carey to explain radical conceptual change (Carey, 2009). Taking inspiration from a modern approach to the language-of-thought (LoT) hypothesis, we model a child's mental representation system for a given domain at a given time as a formal *programming language*—a *domain-specific language* or *DSL*—and conceptual change over time as *DSL synthesis*. Unlike previous LoT-based models, we

1. enrich the semantic search space beyond just functions over *fixed* primitive states to include *new* representations of state or *types*, and

2. explicitly model *analogical mapping* as part of the learning signal via the discovery of shared *relational structure* between the current hypothesis and some external inputs, such as symbolic language, or representational substrate.

In this richer paradigm, discontinuous ("Quinian") conceptual change becomes the invention of a *new type system*, defined over the external symbols and related by analogy to the previous system, but also, crucially, generalizing beyond it.

We evaluate our approach in the classic domain of children's early mathematical development and model the acquisition of both *natural number* and *rational number* as novel conceptual types discontinuous with previously available forms. As referenced, after learning the meanings of number words up to three or four, children famously undergo a sudden and consistent leap to understanding all words in their count list around age 3.5 (the cardinal principle or CP transition). Later, they enrich this understanding after learning fractions (typically between ages 8 and 12). We show that our account reproduces Carey's proposed mechanisms for each stage of this learning trajectory, capturing several previously unmodeled phenomena. In the natural number domain, these newly modeled observations include (1) the multiple stages of the often atomically-viewed CP induction, (2) cross-linguistic diversity in learning trajectories caused by morphological differences in quantifier structure, and (3) the discovery of discrete infinity. In the rational number setting, they include (1) intrusions from the natural number domain in early stages of arithmetic learning, (2) curriculum effects on the rate of fraction learning, and (3) the discovery of infinite divisibility. In addition, our model provides more parsimonious explanations than existing methods for cross-linguistic universals, such as the steady three-number prerequisite for the CP induction, and intervention effects.

More generally, our work provides a computational-level grounding of a longstanding theory on the human capacity for genuine conceptual change—the ability to think truly new thoughts—as a cascading of links from old to new type systems guided by generative analogies. In the rest of the paper, we present a detailed example and related work; formalize the model; and discuss results in the two number domains.

Example and related work

In plainest terms, Quinian bootstrapping is a mechanism that generalizes an initial conceptual system (CS1) into a more expressive conceptual system (CS2), by *adding meaning to the relations* in a *placeholder structure*: externally-provided symbolic structure such as natural language (Carey, 2009). Concretely, we model each conceptual system as a set of interrelated *type systems*, which may be viewed as miniature DSLs that each define the semantics of a single *type*. A type is defined by a set of *states* that may be used to represent the world, as well as a set of *operations* that describe valid relations between states. Further, CS1 additionally includes a representation of placeholder structure as an *incomplete* type system, or a set of external symbols with undefined or partially defined meanings for both its states and operations.

In the natural number setting (Figure 1), CS1 is composed of two type systems, corresponding to the two core number systems: the parallel individuation (PI) system and the approximate number system (ANS). The placeholder structure is the *verbal counting sequence*, i.e. the placeholder symbols are the number words, and the key placeholder relation is the *next_word* relation. Prior to learning the meanings of number words, children have already memorized the counting sequence up to a highest number, but critically, this knowledge is limited to knowing that each number comes *after* the previous one, not that each is *one more*. Enriching this limited understanding of the *next* relation, or put differently, *adding meaning* to it, is key to reaching the CP induction, modeled as transforming the incomplete type system representing the placeholder structure into a complete one with fully defined semantics and interrelations with the other type systems.

How does this work? First, the meanings of the first three (sometimes four) number words are learned in terms of the parallel individuation system, driven by the utility of knowing these meanings for performing tasks in the world (details provided in the following section). At this stage, however, the *analogy* between the *next_word* relation in the symbol (word) domain and the *add_one* relation in the physical (meaning) domain has not yet been discovered.

Unlike the earlier one-knower and two-knower stages, however, the three-knower stage is uniquely ripe as a substrate for making this discovery. This is because there is a *suspicious coincidence*: two *next_word* relations in the symbol domain align with two *add_one* relations in the physical domain. To use the language of the field of programming languages, noticing such a suspicious coincidence reveals an opportunity for *compression* of the mental representation system. Namely, rather than memorizing the count sequence and *separately* representing the meanings of the first three number words as three totally *unrelated* PI states, the full conceptual system can be more efficiently represented as (1) the count sequence, (2) the meaning of one, and (3) meanings of two and three *equivalently* defined as evaluations of *add_one* performed *in lockstep* with the already remembered count list.

Naturally, following analogy discovery, the next and final step of Quinian bootstrapping is analogy *generalization*. Unifying the relationship between the meanings of (one, two) and (two, three) as described suggests that a similar semantic relationship may exist between the meanings of later consecutive pairs in the count list. However, naively extending the analogy in this way fails after the meaning of four, since this is the upper limit of parallel individuation, and the *add_one* relation previously used is an operation supported and thus limited by that system. Put differently, the ability to represent exact quantities above four does not exist in either of the two core number systems, so the meanings of all the counting words cannot be represented by using only these systems.

This dilemma brings us to the deepest contribution of Quinian bootstrapping: Despite the limitations of the domain-

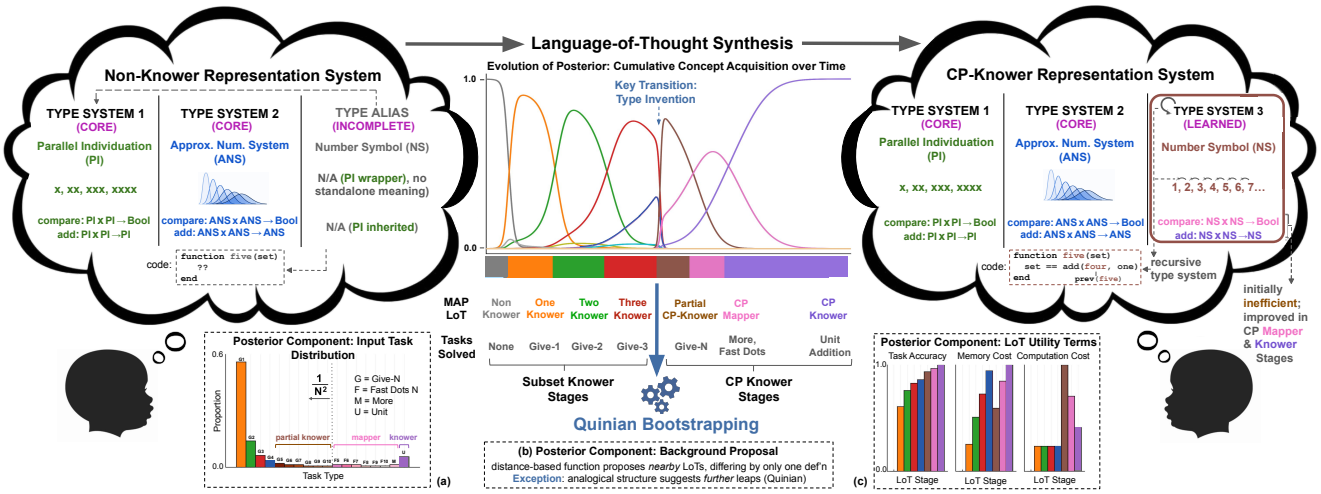

Figure 1: Overview of Natural Number Model. (a), (b), and (c) are all inputs to inference (unrelated to side of diagram).

specific core knowledge systems, truly novel conceptual invention can still occur via the deployment of *domain-general* cognitive constructs—like the ability to define *recursive* data types or representations of state—that *generalize* the relations supported by core cognition beyond their limits. While these constructs are inherently more difficult to access than core knowledge, partly due to their infinite variety and children's lack of practice with them, evidence suggests that they become more accessible when suggested by the *symbolic structure of the environment*, be it physical or *abstract*—e.g., in the form of the current mental representation system and its model of the externally-provided placeholder structure.

Precisely, the *recursive* structure of the analogy between the count list and the meanings of the first three number words suggests the invention of a new *recursive type* using the placeholder symbols. In particular, rather than attempting to *evaluate* the PI-based *add_one* operation to create an *iconic* mental representation of sets greater than size four (i.e. beyond normal working memory capacity), the mental representation of a number word like five may just be represented as "one more than four," or more generally, "one more than the previous word's meaning." Concretely, this is modeled by the invention of a new operation in the placeholder domain,

$$add : NS \times NS \rightarrow NS,$$

that generalizes the analogous relationship in the PI type system to all known number symbols (*NS*), but without requiring the imagination of *any* iconic or otherwise physically-grounded mental format. Instead, it is defined using just the lightweight *next_word* function. Further, the placeholder type definition is updated so that the symbols are no longer mere *wrappers* of other type systems for representing quantities, but a novel, recursively-defined type system that itself captures unique quantity representations—namely, quantities defined using *relation expressions* like *add(four, one)*. Critically, this new symbolic type system is fully *detached* from the physical realm, enabling reasoning about exact discrete quantities without explicitly imagining sets of different sizes (though it can always be *translated back* to physical sets by explicitly evaluating the recursion). As such, it is the key stepping stone towards efficiently reasoning about higher-order symbolic mathematics in later education.

Formally, we implement this two-step algorithm of (1) analogy discovery and (2) analogy generalization as a search for *relational compressibility* in the knowledge representation. We assume that, over the course of their lives, children and adults sometimes notice common relational patterns between parts of their knowledge representation and external signals. This recognition can then lead to knowledge *compression* and *generalization*. Unlike standard compression algorithms that search for common structure across the *semantics of individual functions* ("functional compression"), our model searches for common structure across the *relations between semantics*, be they functions or states, capturing how *grounded* a knowledge system is by an underlying network of analogies ("relational compression"). Overall, we wrap this analogy search method in a utility-based rational learning paradigm (described next), where relational compressibility decreases the *memory cost* of storing a knowledge representation and the *discovery cost* of proposing a new one from the current via generalization.

Main related work: Piantadosi et. al. 2012, 2023

To conclude this overview, we briefly describe the differences between our model and the closest previous work, the rational statistical inference (RSI) model of number learning by Piantadosi, 2023. Our model makes three changes with respect to RSI that let us better explain empirical data, and resolve Carey's critiques about why it is not Quinian bootstrapping.

First, we model each learning stage as being *dependent* on the previous one, in that the *structure* of the current LoT may influence search for new LoTs. In contrast, RSI models learning stages as being fully *independent*, so that the timing of the CP induction is driven by the low prior probability assigned to

Relational Calculus: Implements Domain-General Ability to Represent and Reason about Relational Patterns in a LoT

Base Rules

Name	Form	Interpretation	Code Syntax
(↔, type): Type Relate	$\alpha \leftrightarrow \beta$	$\forall\, a \in \alpha,\ \beta(a) = b \Rightarrow \alpha(b) = a$, and converse.	@relate α β
(<: , type): Type Hierarchy	$\alpha <: \beta$	$\forall\, a \in \alpha,\ \beta(a) = b \Rightarrow \alpha(b) = a$, but not converse.	@abstract α β
(↔, function): Functional Relate	$f_\alpha \leftrightarrow f_\beta$	$\forall\, a_i \in \alpha,\ \alpha(f_\beta(\beta(a_1), \ldots, \beta(a_n))) = f_\alpha(a_1, \ldots, a_n)$, and converse.	@relate f_α f_β
(<: , function): Functional Hierarchy	$f_\alpha <: f_\beta$	$\forall\, a_i \in \alpha,\ \alpha(f_\beta(\beta(a_1), \ldots, \beta(a_n))) = f_\alpha(a_1, \ldots, a_n)$, but not converse.	@abstract f_α f_β

Figure 2: Foundation of the Relational Calculus, or Domain-General Language (rDGL).

the *recurse* primitive in the search space. As such, our model provides a *structural* reason why the CP induction always occurs after at least the three-knower stage, and never, say, the two-knower stage. Namely, the relational shared structure between the count list and children's initial PI-based definitions of the first three numbers is a *suspicious coincidence* that draws attention to the key analogy, triggering compression and generalization to the full count list. In contrast, lightly tweaking the RSI prior assigned to *recurse* could induce the CP stage after the second or earlier stages, a pattern not observed, even in cultures where children experience the CP induction at much older ages (when the general ability to note recursion would be expected to have improved, modeled as a potentially higher prior given to *recurse*; Piantadosi et al., 2014).

Second, the external learning signal in the RSI model comes exclusively from the distribution of individual number words. As such, it predicts that adding more low-number (1-3) counting input at the three-knower stage will not accelerate the CP induction, because it does not affect the LoT's accuracy (three-knowers already know all those word meanings). Counting intervention experiments, however, show this is false: added low number counting *does* accelerate the CP induction (Gibson et al., 2020). Our model predicts this, because more low-number counting input at the three-knower stage accelerates search via *further analysis* of the current hypothesis, causing faster discovery of the analogy leading to the CP leap.

Third, rather than modeling the evolution of a *single function*, namely, the counting algorithm, we model a child's domain representation as a *related set of functions and states*. By carefully modeling the initial primitive states and functions of core numerical cognition, and distilling domain-specific and domain-general biases, this richer representational substrate gives us a platform to *jointly* model how children learn all the *concepts* and *algorithms* of higher math. We instantiate the first step of this research program in this paper by modeling how children augment their hard-won understanding of natural number with the significantly richer notion of *rational number*. There, evidence from educational studies suggests that deeper understanding is also driven by the discovery of a key *symbolic-physical analogy*, as formalized by our approach.

Model

We now present a formal model of our learning system. We frame conceptual evolution as *search* for a mental representation system (LoT) L within a space $\mathcal{L}$, guided by (1) the *utility* of L for solving a distribution of reasoning tasks T (Figure 1a), and (2) the difficulty of *discovering* L from an initial mental representation system L_0. In detail, the utility $U(L;T)$ balances a task *accuracy* term, $A(L;T)$, with two *cost* terms, a *memory cost* $C_m(L)$ capturing how difficult L is to remember (based on description length and relational compressibility), and a *computation cost* $C_c(L;T)$ capturing how *efficiently* L solves the tasks T (based on the length of its *execution trace*):

$$U(L;T) = \gamma_1(t)A(L;T) - \gamma_2 C_m(L) - \gamma_3 C_c(L;T) - \gamma_4.$$

The time-dependent parameter $\gamma_1(t)$ controls the tradeoff between cost and accuracy, increasing over time to model growing tolerance for higher-cost representations if they are also highly accurate. Further, the *discovery cost*, $C_d(L;L_0,T,t)$, is used as part of a *background proposal function* (Figure 1b) that proposes new LoTs L from the current state L_0:

$$P(L \mid L_0;T,t) \propto \exp(-\alpha C_d(L;L_0,T,t)).$$

The added dependence on the task distribution T and time t reflects an increasing ability to propose more *distant* (e.g. CP-knower is more distant from three-knower than four-knower is) LoTs L from the starting state L_0 over time (t), which may be further accelerated by increasing exposure to particular kinds of inputs in the task distribution (T). This design explains the accelerated pace of learning caused by low-number counting interventions—a modification of the base task distribution T—and *curriculum* effects in the rational number domain.

These utility and discovery components are unified in a single learning paradigm using the formalism of rejection sampling. Namely, the discovery-cost-based background proposer is used to *suggest* new possible LoTs, and the utility function is used to decide whether to *accept* or *reject* the proposal:

$$a(L;T,U^*) = \min\{1, \exp(\beta[U(L;T) - U^*])\},\ \text{with}$$

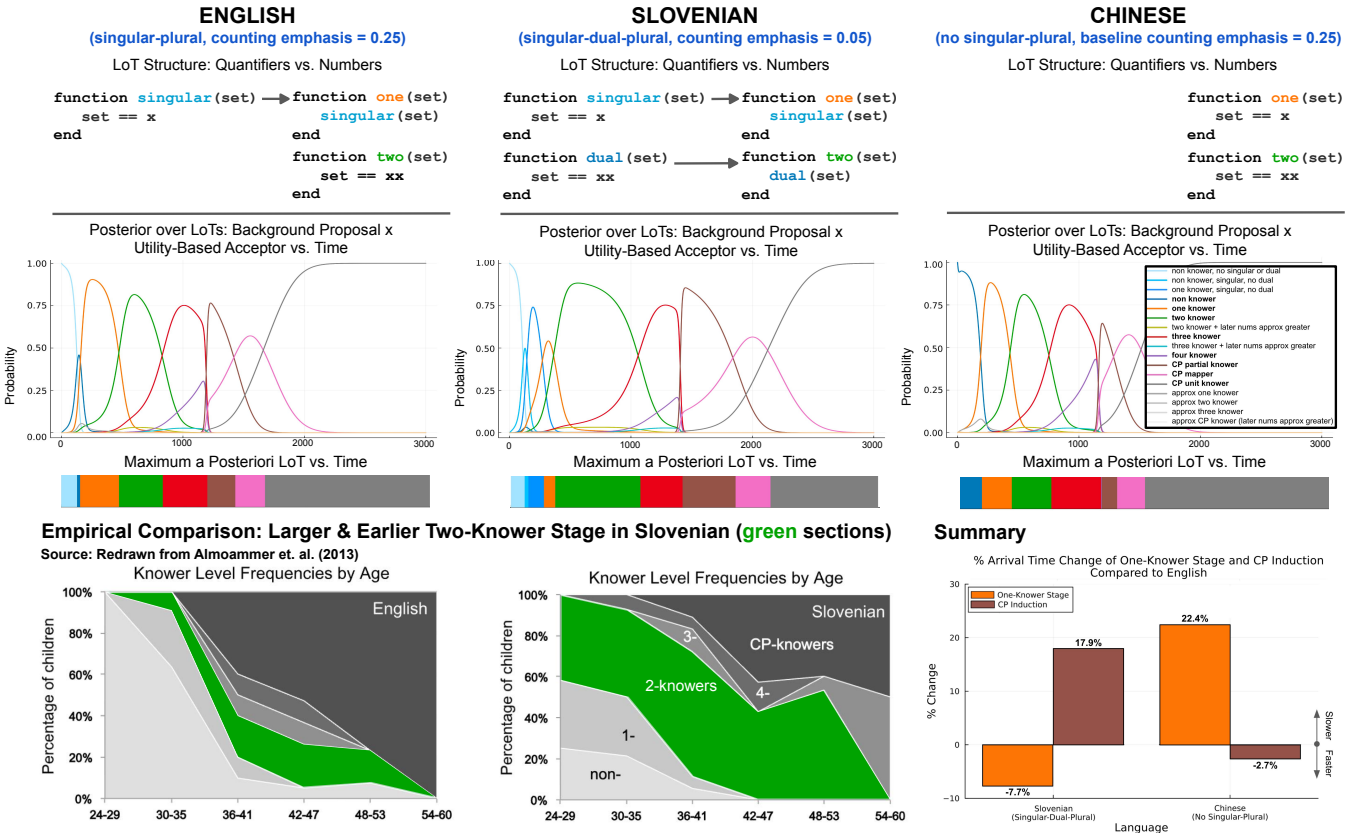

Figure 3: Cross-Linguistic Diversity Results. The direction, not magnitude, of change is most relevant in this prototype model.

$$U^* \geq \max_L U(L;T).$$

Lastly, the posterior is proportional to the product of the background proposal and the utility-based acceptance probability.

Importantly, the mathematical connection between this formalism and that of the RSI model is that the latter lacks the *previous-state-based* (i.e. L_0-based) *background proposal function*, meaning that the optimal LoT at a given time is simply the *maximum utility* one at that time, without dependence on the previous LoT. As such, the RSI model requires *tuning* of the memory cost parameter assigned to the CP stage in order to reproduce the observed invariant that the CP induction always happens after at least the three-knower stage, whereas our model avoids this parameter sensitivity via a *structural*, previous-state-based explanation. Correspondingly, in our full set of model comparison results (Figure 11 and Appendix B), we refer to our model as the *UB* model, for "utility + background proposal," and the baseline RSI model as the *U* model, for "utility-only."

Overall, our design is different in three ways from more standard formulations of utility-based models. (1) While standard formulations balance a single *complexity* cost term against an accuracy term, we split the complexity cost into *two* terms: a discovery cost capturing the difficulty of *proposing* a hypothesis, and a memory cost capturing the difficulty of *remembering* a hypothesis once proposed. This distinction captures the intuition that, while the CP induction is hard to *discover*, it is not that hard to *remember*, with fewer distinct definitions than the higher-knower stages due to relational understanding. (2) We add a computational cost term, modeling the observation that computational efficiency on tasks can also drive conceptual evolution (see results on modeling the multiple stages of the CP induction). (3) Traditional definitions of complexity cost are based on the description length of a hypothesis, where the final length is often computed after a *functional compression* step that shortens the length by extracting common structure in function semantics. Our model extends this by accounting for *relational compressibility*, or the existence of shared structure in the *relations between semantics*, capturing the notion that knowledge systems grounded by common relations are often easier to store (memory cost; Figure 1c) and make particular generalizations easier to find (discovery cost; Figure 1b).

Results

We now describe the results of instantiating our model.

Domain 1: Natural Number

In our current natural number implementation, the LoT search space $\mathcal{L}_N$ contains 75 LoTs, including the key observed stages

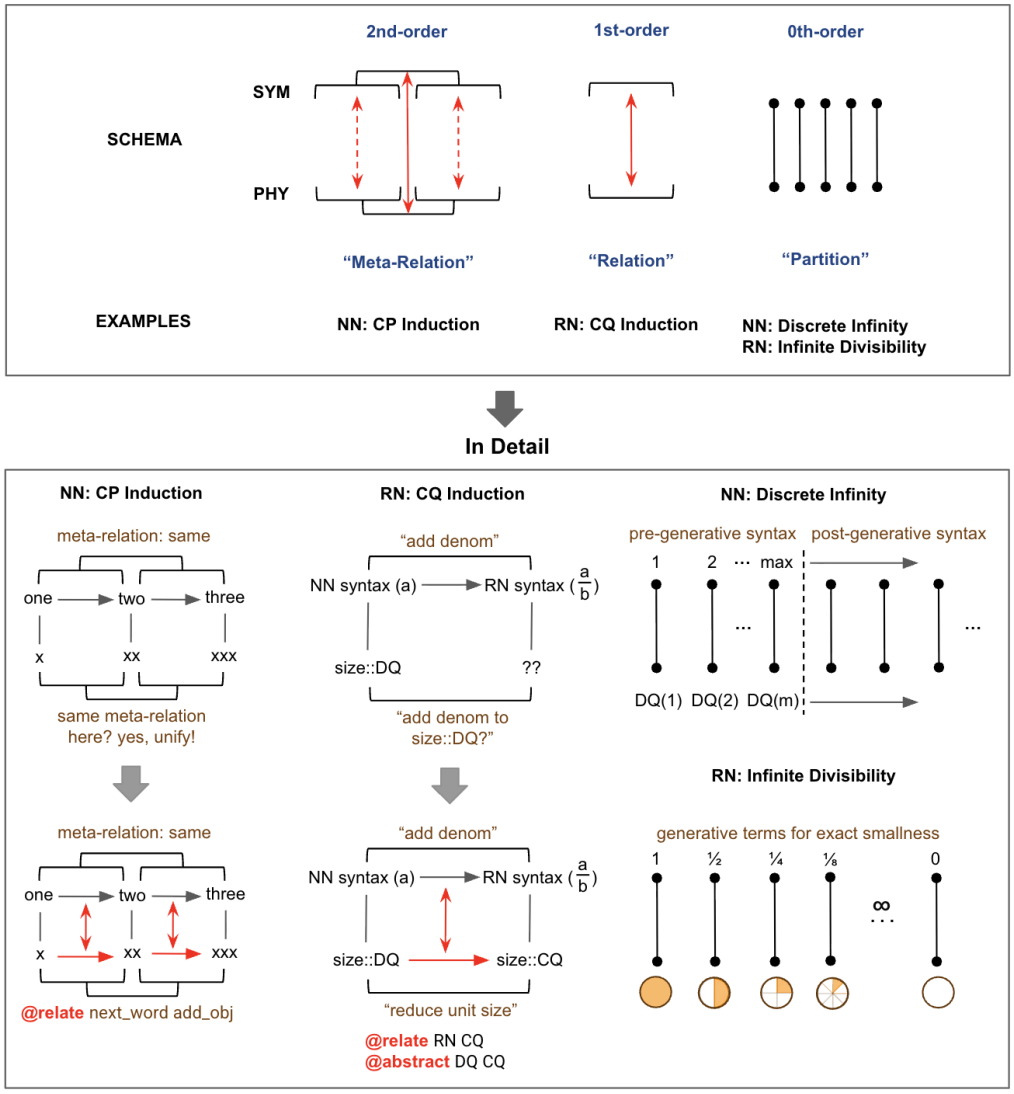

Figure 4: Summary of How Symbols and Symbolic Relations Suggest Conceptual Generalization. Legend: NN = Natural Number, RN = Rational Number, DQ = Discrete Quantity, CQ = Continuous Quantity.

(i.e. one-knower, two-knower, etc.), along with other definitions that children might assign to the number words one through ten based on parallel individuation (e.g. two means *x*), the ANS (e.g. one means *approximately one*), or a combination of both (e.g. one means *x*, and each subsequent number is approximately more than the last). A subset of these LoTs are listed in the center-right legend of Figure 3.

Inducing the Cardinal Principle The base result—the evolution of the incremental subset-knower stages leading to the CP induction—is shown in Figure 1. As the cost tolerance increases over time, higher-cost but also higher-accuracy LoTs emerge. The key Quinian transition takes place between the three-knower and *partial CP knower* stage once the discovery cost, modeled as *decreasing* in proportion to the time *t* as children *increasingly* analyze the three-knower LoT's structure, is low enough for the CP leap to be realistically proposed.

Multiple CP Stages: Partial Knower, Mapper, and Knower While traditionally viewed as an atomic event, Davidson et al., 2012 and others have shown that many children who solve the Give-N task, historically the CP-knower criterion, *cannot* answer other basic number questions. These other tasks include the More task, which asks which of a pair of higher numbers is larger ("Which is more, six or nine?"); the Fast Dots task, which asks children to estimate the number of dots in a flashed set, with the correctness criterion being stating an increasing list of numbers for increasingly large sets (indicating knowledge of the *later-greater* principle); and the Unit task, which tests children's understanding of the exact unit difference between consecutive number words ("If I add one to seven, do I get eight or nine?"). Previous models have failed to capture

these nuances, because they only model the acquisition of one part of number understanding, a single function representing the counting algorithm, as opposed to the full conceptual system also describing relative magnitude understanding (More and Fast Dots tasks) and exact arithmetic (Unit task).

Our richer modeling formalism lets us explain this behavioral data by representing the post-Give-N concepts as *additional function semantics* to learn upon recognizing natural numbers as a recursive type system. In the language of computer science, these extra functions—namely, the *compare* and *add* functions—form a *programming interface* for any type system for quantity representation, meaning that bestowing that label upon number words requires having to define these interface functions as well (indeed, both core number systems include built-in semantics for these functions).

Concretely, following a hypothesis proposed in Davidson et al., 2012, we model the initial, Give-N-based CP stage (partial CP-knower) as including *default* definitions for both functions that are *overly grounded* in the newly-understood counting procedure. Precisely, to *compare* two numbers above the PI limit, children at this stage *count up* from one for each number, and then give an answer. Similarly, to *add one* to a higher number, they also count up from one to the largest number, and then one more—despite being able to answer what the *next* word is immediately. Both of these methods are tremendously *computationally inefficient*, causing children to give up and guess, especially at higher numbers. Conceptual evolution takes place as children learn more efficient algorithms for comparison and addition (similar to Shrager and Siegler, 1998), by memorizing the later-greater principle (CP-mapper), and later refining it via the exact +1 principle (CP-knower). While memorizing these extra facts *increases the memory cost*, the *gain in accuracy* and *reduction in computational cost* (Figure 1c) eventually causes them to emerge.

Cross-Linguistic Diversity A longstanding hypothesis in number learning is that the morphology of natural language might play a role in how children assign initial meanings to number words (Bloom and Wynn, 1997). In particular, the *quantifier structure* of language, such as singular-plural syntax, may provide a learning signal, via structures like singular articles and nouns—which are higher frequency in children's language input and thus learned earlier—being used in the same context as number words like "one." This predicts that children raised in languages with different quantifier structures may exhibit different paces of learning. In particular, children taught languages with *richer* quantifier structure might be expected to learn the meanings of the first few number words *sooner*. This prediction has been empirically corroborated with children's number learning data across several languages (e.g. Sarnecka et al., 2007, Villarroel et al., 2011), but has not been explicitly modeled in existing formalisms.

Our model naturally captures these effects, because it supports the addition of functions corresponding to singular, dual, and plural marker assignment to the knowledge representation, as well as a learning mechanism where the existence of definitions for these terms in the current hypothesis increases the probability of *discovering the same definitions* for number words, namely one (singular) and two (dual), as shown in Figure 3. By (1) adding these quantifier representations to the LoT for each natural language, (2) modifying the input task distribution to reflect better practice with early number words when their meanings are emphasized by hints in quantifier structure, and (3) incorporating variant emphases on *counting instruction* across different languages—approximated by the highest count of same-age children raised in different cultures—as a linear coefficient affecting the discovery cost of the CP generalization from the three-knower stage over time, we explain observed differences in number learning trajectories across several languages (Le Corre et al., 2016, Almoammer et al., 2013). We find that children raised in languages *without* singular-plural structure (Chinese) are *slower* to learn the meaning of one compared to those raised in languages *with* singular-plural structure (English), and children raised in languages with singular-*dual*-plural structure (Slovenian) learn the meanings of one and two *faster*, without the same-direction effects on the ultimate arrival of the CP induction.

Discovering Discrete Infinity Lastly, after they become CP-knowers, children still fail to understand the related concept of *discrete infinity*—namely, that it is always possible to *add one* to a number, and similarly that *numbers never end*—for roughly one to two years. Furthermore, evidence suggests that recognizing infinity is correlated with learning a *productive counting rule*, or *generative grammar*, for number words, meaning that children who know how to productively generate novel *number syntaxes* are more likely to recognize the concept that *discrete sizes* are conceptually infinite, as well (Chu et al., 2020, Sullivan et al., 2023). Our model explains this observation by formalizing knowledge of a generative number word grammar as a *richer semantics* for the *next_word* function, which was originally defined using a *finite list* prior to the CP induction, along with a learning mechanism (expressed in the background proposal function) by which the *existence* of such *unbounded symbols* in the knowledge representation increases the probability of recognizing the *analogous* existence of *unbounded meanings* for those symbols, as well. In other words, similar to how the *relational structure* of the symbolic count list promoted analysis that ultimately led children to notice *analogous relations* among the symbols' meanings (the CP induction), the *existence* of unbounded symbols also promotes analysis that suggests an *analogous partition* of the meaning space. This relationship between the two kinds of learning signals is crystallized in the schema in Figure 4, which classifies the stable order of the verbal counting sequence as a *second-order* relational signal, and generative number word syntax as a *zeroth-order* relational signal, where each signal suggests analogous structure among the words' meanings.

Domain 2: Rational Number

Next, in the years following their hard-won acquisition of the concept of natural number, children are introduced to a

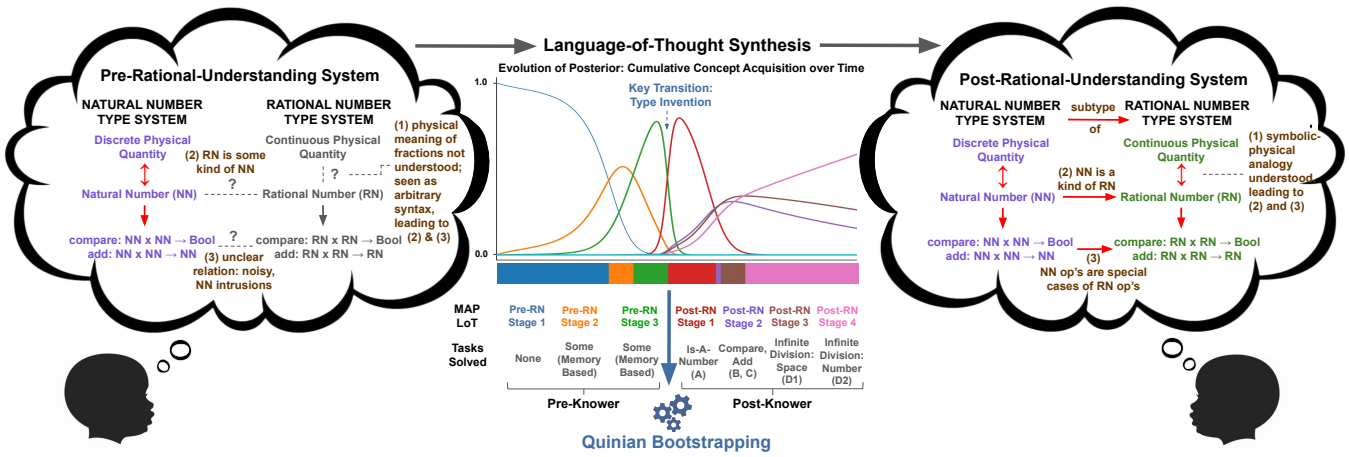

Figure 5: Overview of Rational Number Model.

second generalization of their number representation system during their formal education: the concept of *rational number*, typically taught between ages 8 and 12. Like the CP induction, empirical evidence suggests that understanding rational numbers also involves a *discontinuous* conceptual change, in which the initial conceptual system—anchored by natural numbers and their physical analog, discrete quantities—is expanded with a new universal anchor, *continuous quantities*, denoted by symbolic fraction syntax. Precisely, surveys of children's number understanding indicate that they originally do not see fractions as measures of continuous quantities, but instead as *special kinds* of natural numbers (Moss and Case, 1999, Smith et al., 2005). This conceptual inversion leads to several mistakes, especially caused by natural number intrusions—the existing and more relationally grounded knowledge representation—in performing fraction comparison and arithmetic (e.g. $\frac{1}{2} + \frac{1}{3} = \frac{2}{5}$). Children's misconceptions are corrected upon discovering the symbolic-physical analogy between fraction syntax and continuous quantities (Carey, 2009), at which point they first become systematically sensitive to the notion of continuous *unit size* and, more abstractly, the continuous *number line*. In the following sections, we detail how these observations are captured by a natural continuation of our analogy-driven modeling framework for learning to count. In this second stage of modeling, the key discontinuity, which we call the *Continuous Quantity* or *CQ* induction, becomes a *type system inversion* catalyzed by (1) a symbolic structure-based signal expressed in the background proposal function and (2) a physical training signal expressed in the utility function, as in the natural number setting.

Model Implementation *LoT and Task Spaces.* We implement LoT synthesis over a space $\mathcal{L}_R$ of 120 LoTs, each of which describes variant semantics for fractions and their operations. These variations include both the correct, taught semantics and several natural-number-based intrusions. The LoTs are evaluated on a set of tasks T, where the tasks fall into two categories reflecting the key choice of emphasis faced by educators designing curricula: (1) physical understanding tasks and (2) symbolic memorization tasks. The physical understanding tasks are sourced from a study by Smith et al., 2005 in which children were asked several questions regarding their conceptual knowledge of fractions. Results showed that most of the students exhibited approximately "all or nothing" patterns of responses, meaning they either fully grasped or fully did not grasp the physical meaning of fractions, with very few exhibiting true *transitional* patterns. This outcome provides suggestive evidence that the CQ induction indeed takes place as a discontinuous (Quinian) conceptual change, rather than more typical incremental learning. The concrete tasks included (1) the Is-A-Number task, in which children were asked if they believed there were numbers between zero and one, and what the two numbers in a fraction mean; (2) Comparison tasks, in which children were asked to compare the magnitude of two fractions, where the answer could be attained using just conceptual understanding as opposed to algorithm execution (e.g. "Which is more, $\frac{1}{2}$ or $\frac{1}{3}$?"); and (3) Infinite Divisibility tasks, in which children were asked about their understanding of the infinite divisibility of (a) space, (b) number, and (c) weight, key fraction-supported concepts. In addition to these conceptual understanding tasks, we also include two types of symbolic memorization tasks: (1) symbolic *notation* tasks (e.g. being able to translate "four fifths" into the correct fraction syntax, $\frac{4}{5}$), and (2) symbolic *algorithm* tasks, specifically for fraction arithmetic ("What is $\frac{1}{2} + \frac{1}{3}$?") and comparison ("Which is more, $\frac{2}{5}$ or $\frac{1}{3}$?"). Each element of a rational number LoT specifies a procedure for answering one of these task types, either correctly or incorrectly.

Utility. Unlike the natural number setting, the overall distribution of tasks in the rational number context is not primarily defined by an environmental "power law," since fractions are taught as part of formal education rather than prior to school.

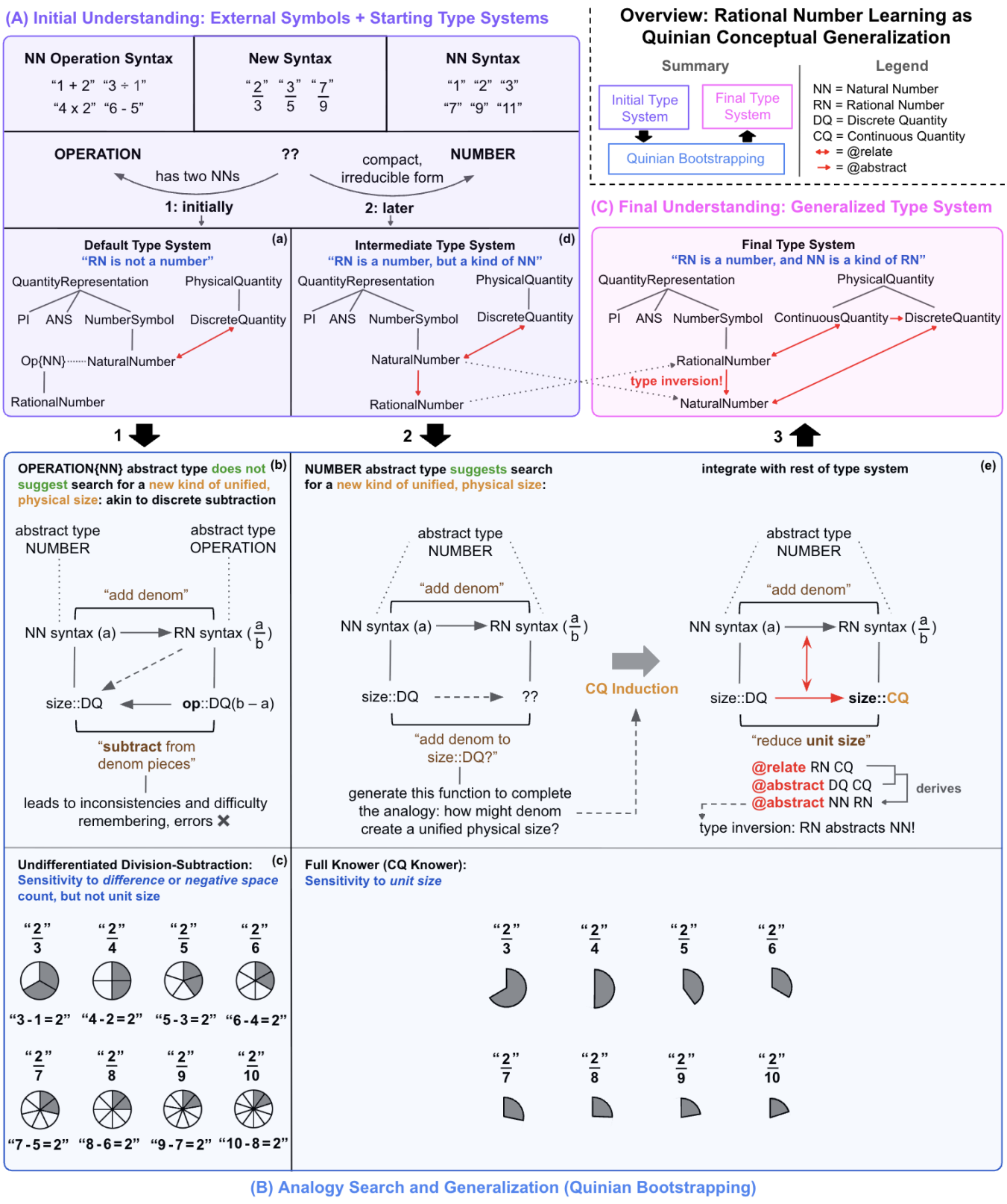

Figure 6: Overview of the Continuous Quantity (CQ) Induction Mechanism.

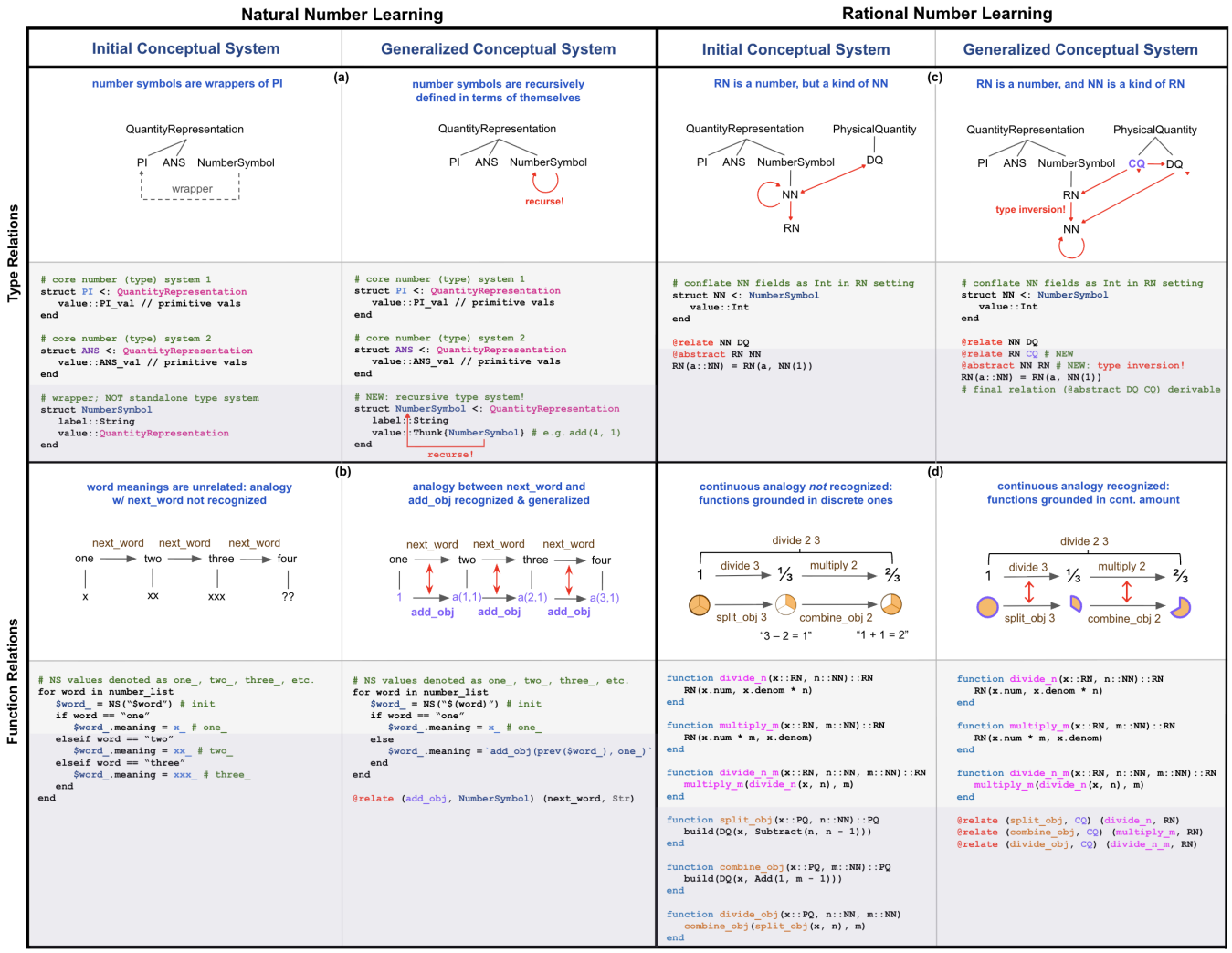

Figure 7: Implementation of Type and Function Relations. Code Background Legend: Light = Unchanged, Dark = Changed.

A given task distribution thus reflects a particular curriculum design choice. We test variations of this distribution, producing different accuracy terms $A(L;T)$ for each LoT L, in our experiments, particularly along a single symbolic-physical training axis relevant to curriculum design recommendations (see "Curriculum effects" section). The other terms in the utility function, the memory cost $C_m(L)$ and computational cost $C_c(L;T)$, are defined as in the natural number setting: They are based on the LoT *description length* and *execution trace length*, respectively. Finally, just as the CP induction reveals an opportunity for *relational compression* that reduces the memory costs of some LoTs, the CQ induction enables analogous relational compression and memory cost effects in the rational number setting, described in the following sections.

Background Proposal. Mirroring the natural number setting, the background proposal function $P(L \mid L_0;T,t)$ in the rational number model is primarily based on the edit distance between the current LoT L_0 and the proposed LoT L, and modulated by *relational compression* and *generalization* suggested by the symbolic structure of L_0 (Figure 6). Specifically, the relevant symbols are fraction symbols $\frac{a}{b}$, and the relevant symbolic relation is the *divide(a:NN, b:NN)* relation, which adds a denominator b to an initial natural number a to make $\frac{a}{b}$. Longstanding observations on children's rational number learning indicate that children initially see fraction symbols not as denoting numbers in themselves, but as *two numbers*, or more generally, some kind of *operation between two numbers* (Figure 6a). This stage has been called "undifferentiated division-subtraction," since fractions are viewed as having something to do with subtraction, the most similar prior operation (indeed, the special case of whole number division *does* actually center integer subtraction, adding to the confusion). As such, children initially appear to view the physical analog of fractions to be a representation of a subtraction-like operation between two discrete quantities, a items and b items (Figure 6b). In other words, they effectively show sensitivity

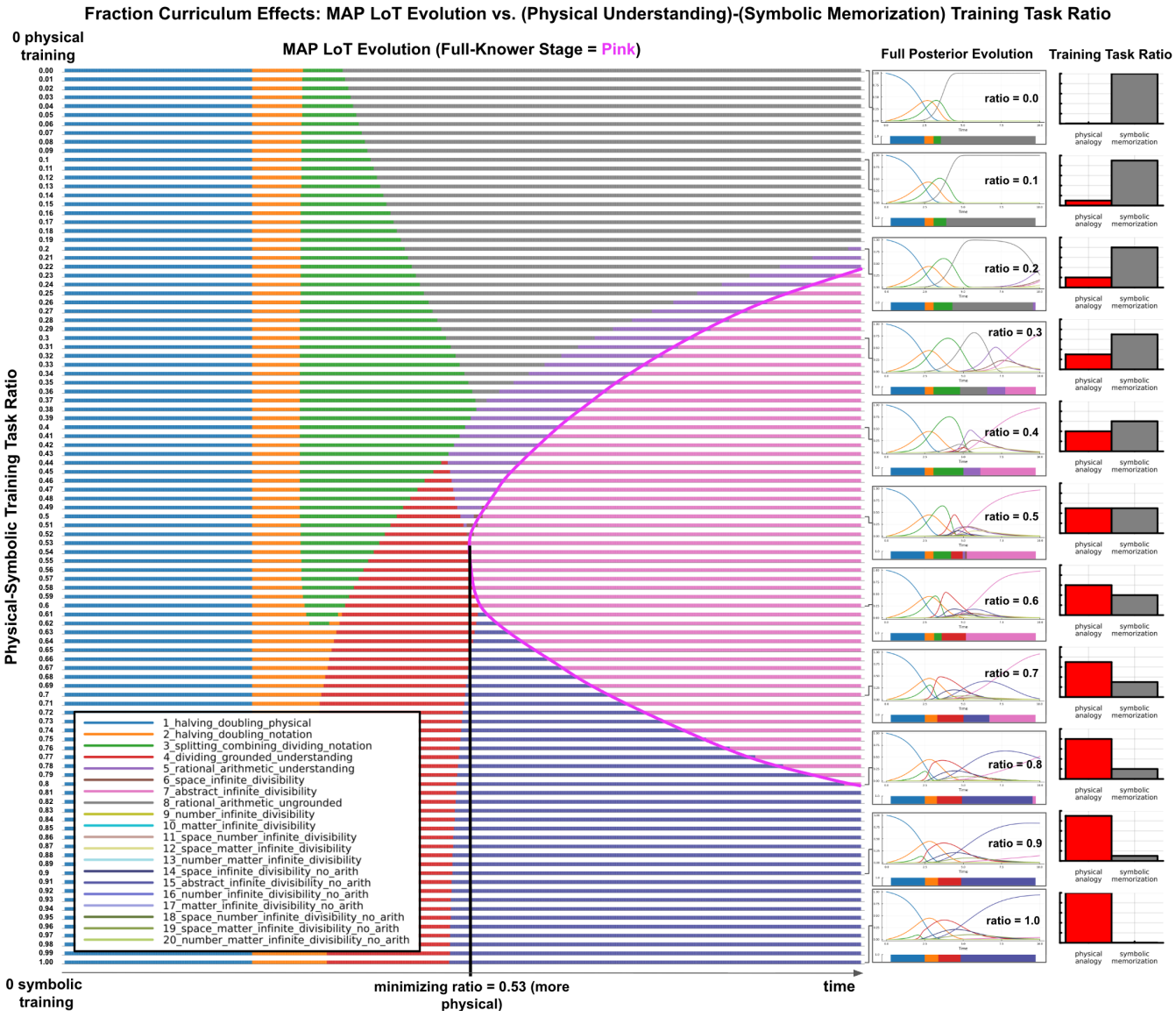

Figure 8: Empirical Prediction of Rational Number Model: Optimal Symbolic-Physical Training Ratio for Student Learning.

to the *difference* ("negative space"), $b - a$ items, rather than the reduction in continuous *unit size*, a property of the original set of a items (Figure 6c). While this conceptual system lets one answer simple pie-shading questions, it leads to incorrect answers on more aptly designed conceptual questions.

Eventually, the correct understanding is induced with the help of a symbolic signal, along with the standard utility signal of not being to solve most physical understanding tasks. In detail, the *compact, number-like syntax* of novel fraction notation, along with its use as the final, *irreducible* answer in equation syntax such as $4 \div 6 = \frac{2}{3}$—a syntactic role previously occupied only by natural numbers—helps suggest that fractions may be some kind of *unified number*, not just an operation (Figure 6d). This sets the stage for the final analogy completion: If fractions $\frac{a}{b}$ are some new kind of number, and all previously known numbers denote some kind of *unified physical size*, how might adding a denominator b to a set of a items produce some new kind of *unified size property*? This draws attention away from the operation-based *difference*, $b - a$ items, to the reduction in the *unit size* of each of the original a items (Figure 6e), a leap that this model predicts would be much more difficult in a cultural context where fractions were only denoted using ordinary natural number operation syntax, $a \div b$, instead of novel, compact number-like syntax.

Overall, the significance of this discovery of an analogy between adding a denominator b to a number a—a symbolic relation—and *splitting* the continuous unit size of a pieces so

each has to be combined b times to reach the original size—a physical relation—is twofold (Figure 7). First, it simplifies or *compresses* the existing knowledge representation, by defining a *translation function* (implementing the abstract *number line*) by which operations in the symbolic domain can be intuitively viewed in terms of continuous quantities in the physical domain. In contrast, these operations previously had more discordant physical meanings in terms of adding and subtracting discrete sets ("undifferentiated division-subtraction"; Figure 7d). Second, it enables *generalization* of the knowledge representation previously blocked by dependence on the recognition that fractions measure continuous quantities, specifically with respect to understanding the *infinite divisibility* of number and physical matter. Both of these effects are supported by the underlying *inversion* of the original type system spurred by the symbolic-physical analogy discovery: natural numbers are special kinds of rational numbers, just as discontinuous quantities are special kinds of continuous quantities, not the opposite (Figure 7c). Further, both these effects mirror those caused by the CP induction in the natural number setting. We analyze these and other comparisons between the natural and rational number models—including that the symbolic analogy signal in the CP induction is a *second-order* relation and that in the CQ induction is a *first-order* relation—in the discussion, and now present the results of evaluating our model.

Base Result: Inducing the Continuity Principle The bedrock result of the rational number model—the evolution of incremental symbolic- and undifferentiated-physical-knower stages, followed by the CQ induction, and then a subsequent flurry of quick conceptual advances—is shown in Figure 5. The base task distribution contains large, equal proportions of both symbolic algorithm memorization tasks and physical understanding tasks, the two most challenging tasks to learn, and the ones that we vary in the curriculum experiments described later in this section. In this base result, the MAP LoT evolves from the *non-knower* stage, to the *symbolic halving notation* stage, to the *general symbolic notation* stage, as the cost-tolerance gradually increases over time and selects for higher-cost but also higher-accuracy LoTs. At the general notation stage, children have memorized full fraction syntax, as well as acquired the partial understanding of their physical meanings sufficient to solve pie shading tasks (i.e. by counting up to the numerator in a pre-sliced, unshaded pie), but not deeper conceptual questions (e.g. about the existence of numbers between zero and one, or the continuous unit size meaning of the denominator). However, the symbolic scaffolding of this stage provides the substrate for the ultimate Quinian leap (CQ induction). We model this by letting the discovery cost of the initial CQ stage from this supporting stage *decrease* in proportion to the time t as children *increasingly* analyze the structure of their knowledge representation, ultimately noticing similarities between natural number syntax and symbolic fraction syntax and inducing the symbolic-physical analogy those similarities suggest. The stages following the CQ induction, symbolic arithmetic acquisition and infinite divisibility understanding, become significantly easier to learn after the analogy-induced type generalization due to relational compressibility. We describe the details of these structural signals in the next two sections. We note here that the order of these two post-CQ stages with respect to each other is not a *structural* prediction of our model, and instead is directly sensitive to their corresponding task proportions in the training distribution (i.e. increasing the amount of infinite divisibility tasks will make it emerge first, before operations). Rather, the key *invariant* predicted by our model is that each of these conceptual advances is *significantly easier* to achieve after experiencing the CQ leap, as we discuss next.

Analogy-Induced Compression: Modeling "Rederivation from First Principles" As referenced, the critical relational compression enabled by the CQ induction is the introduction of a translation function—instantiating the concept of the abstract number line—that *decreases redundancy* between the symbolic and physical parts of the knowledge representation by *deduplicating* shared structure. Intuitively, rather than storing $2x$, where x is the size of one knowledge domain (symbolic or physical), relational compression enables storing just $x + \delta$, where δ is the size of the translation functions that enable straightforward reconstruction of the second domain using the structure of the first. The crux of the translation function implementation is what we call the *symbolic-physical basis functions* (Figure 7d): the recognition that

1. symbolic *dividing* (changing the denominator) is analogous to physical *splitting*, and that
2. that symbolic *multiplying* (changing the numerator) is analogous to physical *combining*,

where splitting and combining each produce a new continuous size, not just a discrete set of pieces. These analogies let any higher-order symbolic algorithm composed of these basis functions to be translated into a corresponding, intuitive physical algorithm, a powerful aid for learning and remembering such procedures.

This aid is concretely used in our model in the acquisition of the symbolic operation semantics. While these semantics—which involve (1) finding a common denominator, (2) scaling the numerators and denominators of each fraction accordingly, and (3) performing the respective natural number operation over the numerators, for each of fraction addition, subtraction, and comparison—can be memorized prior to the CQ induction, they are understood then as *arbitrary symbolic manipulations*, since children have not yet understood how fractions correspond to continuous size. The absence of this symbolic-physical connection thus leads to difficulty remembering and a high error rate, as empirically observed. After the CQ induction, however, learning these semantics becomes much easier, because each step of the symbolic operation can be *translated to an equivalent step* in the more intuitive physical domain via the basis functions. Further, the steps become much easier to *remember* after being learned in the post-CQ stage because they can be *rederived from first principles*: re-synthesized

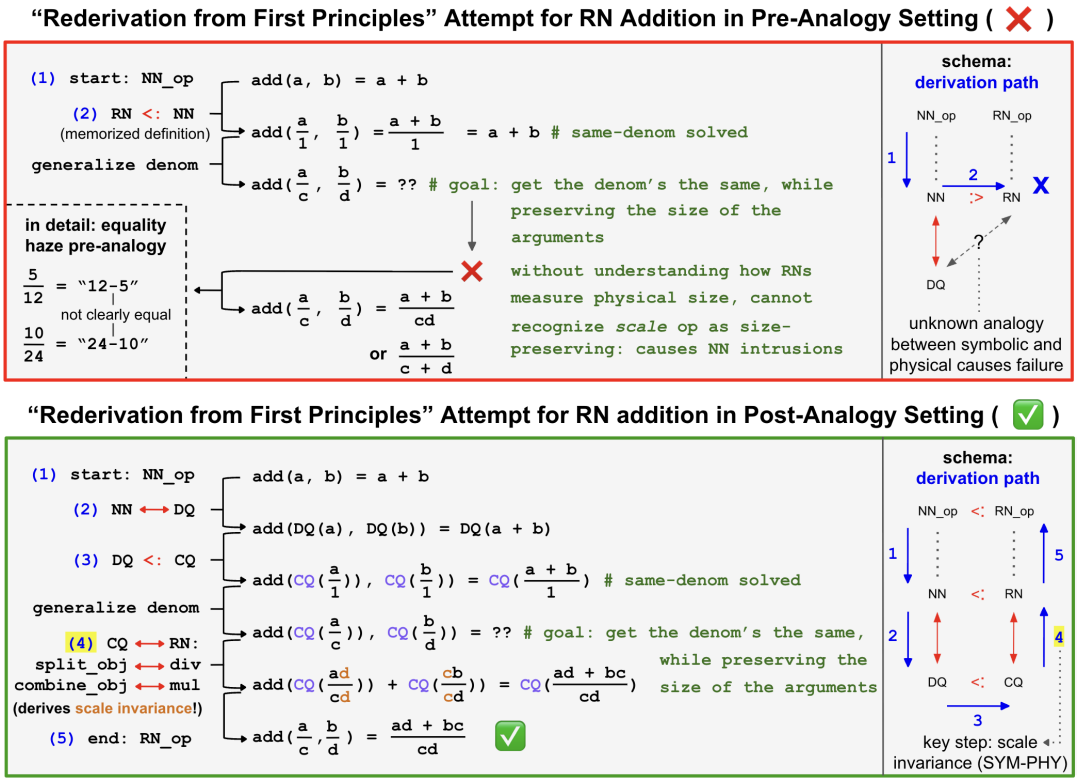

Figure 9: Implementation of "Rederivation from First Principles" using the rDGL.

from the program representations of the corresponding natural number operations and the translation functions (i.e. rDSL relations) between symbolic fraction syntax and continuous quantities, as well as between rational and natural number symbols, as shown in Figure 9. In other words, after the CQ induction, the implementations of these symbolic procedures do not need to be explicitly stored in the LoT, because they can instead be rederived on the fly using the existing relational knowledge (Figure A.5). This causes a significant reduction in memory cost, and hence the (proportional) forgetting rate.

Discovering Infinite Divisibility Next, for the final conceptual advance of rational number learning, the mechanism driving it mirrors that in the discovery of discrete infinity: learning a generative (infinite) set of *symbols* for exact gradations of size—*large* sizes in the natural number setting and *small* sizes in the rational number setting—ultimately suggests that these sizes are *conceptually* infinite, as well. Fraction syntax provides such an infinite set of symbols, and thus provides the key *zeroth-order* symbolic signal that drives the analogous recognition of infinitely large sets. An important difference, however, between grasping infinite degrees of smallness (fractions) and largeness (whole numbers) is that, while the latter has been shown to take place roughly simultaneously across symbolic and physical domains (Sullivan et al., 2023), the former appears to develop in a more "zigzag" fashion. Precisely, across the three substrate realms of physical *space*, symbolic *number*, and physical *weight*, several children in a study by Smith et al., 2005 recognized only the infinite divisibility of space, rather than all three at once. This empirical observation suggests a *co-constructive* effect, where learning fraction symbols makes it easier to recognize the infinite divisibility of all three domains, but most significantly in the domain of space, perhaps because it is *visual* and more accessible. Recognizing the infinite divisibility of space then makes it practical to recognize the same property in the more abstract domains of number and weight, which appear nearly impossible to propose independently, as evidenced by the fact that almost no children learned them prior to space or to each other. Our model naturally captures this co-constructive effect due to its integration of dependence on the structure of the current LoT in its proposal function.

Curriculum Effects: Symbolic Memorization vs. Physical Analogy Training The curriculum design choice that has long puzzled early mathematics educators with respect to fractions is how much to emphasize symbolic (memorization) training versus physical (understanding) training. We conduct an exploration of this tradeoff using our model of rational number learning, to produce empirical predictions of children's learning rate upon exposure to different training decisions. Results of this exploration are shown in Figures 8 and A.6, where we define symbolic training as symbolic algorithm tasks (arithmetic and comparison) and physical training as the

Is-A-Number conceptual tasks. A structural prediction of our model is that increasing the number of physical understanding tasks will accelerate full understanding, because it promotes *analysis* of the current LoT to accelerate the CQ leap, which then makes it easier to learn symbolic operations due to the revealed *relational redundancy* between symbolic and physical domains. Further, when the total number of symbolic and physical tasks is fixed, the model predicts an *optimal ratio* of symbolic to physical training that maximizes learning rate. While the exact minimum is sensitive to parameters that may be fit to future experimental data, this parabolic prediction structure is an invariant of this model, while the ablated utility-only model does not predict such a minimum. Instead, because it does not encode directional dependence on the current LoT structure, it is mainly sensitive just to the overall *sum* of tasks, which is invariant in this setting, leading to flat prediction structure except near the zero-task edge cases.

Memorization Strength Effects Finally, a second empirical prediction that differs between our model and its primary ablation, the model-structure-*independent* or utility-only model, is that our model predicts that kids who are very good at memorization will actually be *slower* at reaching the CQ induction and achieving full conceptual understanding. This is because fast memorizers quickly acquire LoTs with high utility scores, from which there is *less* utility incentive to accept new LoTs even when proposed, because the transition rate depends on the steepness of the relative utility differential. This result is shown in Figure A.7. In contrast, the utility-only model, lacking any dependence on the current LoT, predicts that fast memorizers will *also* more quickly arrive at the CQ induction and full rational number understanding. We are excited to test this structural prediction, which likely extends to other domains of learning as well, in future empirical work.

Discussion

To end, discontinuities in human learning have animated philosophers, educators, scientists, and parents for centuries. In this work, we present an algorithmic-level explanation of these enigmatic transitions, answering a longstanding nativist objection to discontinuity left unresolved by past accounts. Moreover, our paradigm also unifies several disparate observations about children's early learning in the archetypal domain of numerical development using a simple, symbolic cognitive model. Formally, inspired by Carey's theory of Quinian bootstrapping, our program induction framework casts discontinuity as the invention of a *new type system* for cognition, related by *structural analogy* to the previous representation system, but also, crucially, generalizing beyond it. We instantiate this method in the setting of children's acquisition of the concepts of both natural and rational number, each of which is (consecutively) discontinuous with previously available forms. We find our model captures several previously unmodeled phenomena from the empirical literature—including the multiple stages of the often atomically-viewed CP induction, the discovery of discrete infinity, and the discovery of infinite divisibility—and provides more parsimonious explanations for cross-linguistic universals and intervention effects. In the following sections, we expand on the implications of our contribution for research in philosophy, cognitive science, education design, and computer science.

Philosophical Implications: Against Radical Nativism

Infamously, Fodor defended a particular view of nativism known as *radical concept nativism*, which has two tenets: (1) Virtually all concepts at the level of individual lexical items already *exist* in the mind, in *atomic* or full-fledged form (they are "built in"), and (2) all these pre-existing concepts become *accessible* once they are atomically *triggered* by the corresponding stimuli in the external environment. Quinian bootstrapping was proposed by Carey as a counterpoint to Fodor's radical nativism. While her theory acknowledges that the abstract *capacity* to express all these concepts always exists in the mind, it disagrees with how these concepts become *accessible* to the learner, describing a mechanism by which initially inaccessible concepts become accessible via a *construction process* if they are suggested by *fine-grained structure* among external cultural symbols and their corresponding mental representations (Figure 10). Put differently, Quinian bootstrapping offers a different account of *access*, one that privileges *structure dependence* on the current mental representation, not just structure-independent, external-task-triggered emergence.

Concretely, how does our computational formalization of Quinian bootstrapping instantiate this high-level mechanism? Our key idea is to separate *domain-specific cognition*—initial systems like core knowledge (e.g. PI, ANS) that are immediately accessible, but have limited expressivity—from *domain-general cognition*—an infinite grammar capturing the arbitrary complexity of symbolic thought, including new abstractions of state (e.g. Quinian refactorings of types), but which is difficult to access. Prior approaches manage this dichotomy by mixing domain-specific primitives with a *hand-selected* set of domain-general ones to create a single, problem-specific search space, where the latter primitives can be discovered at *any point* in learning, *independent* of the current hypothesis (Piantadosi et al., 2012). In contrast, we argue that a model in which domain-general constructs *become accessible* when they are *suggested by the symbolic structure of the cultural environment*—especially cultural inventions like the number list or fraction syntax—is a more parsimonious explanation for empirical learning data, and avoids the initial primitive "build-in" noted by Fodor. Poetically, culture provides symbols to *grip onto* as one ventures into the infinite unknown beyond initial, domain-specific conceptual systems, where this venturing takes the form of (1) recognizing a specific analogy between some symbols and their initial meanings, and (2) generalizing that analogy to the remainder of the symbols to construct a wider, previously inaccessible set of meanings.

Overall, our crystallization of the *mechanism* and *outcome* of Quinian bootstrapping as the *analogy-driven* invention of

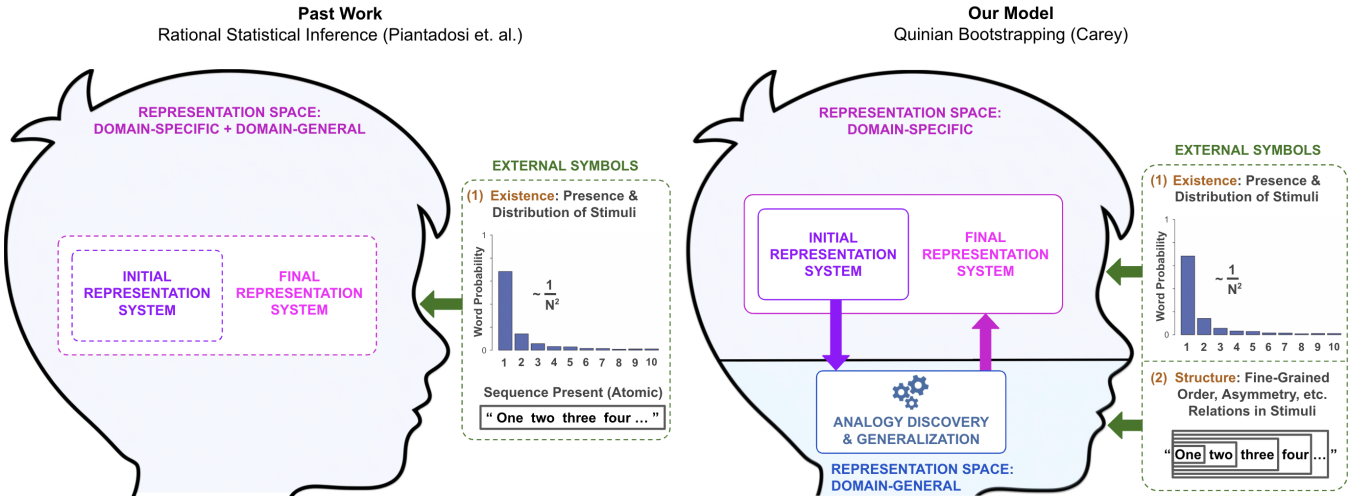

Figure 10: Comparison of Philosophical Paradigms: Fodorian (Piantadosi et. al.) vs. Quinian (Carey; our model) discovery. Solid pink-purple lines indicate discontinuity, while dashed versions indicate continuity within an initially accessible space.

a *new type system* for domain-specific reasoning, defined over external symbols and generalizing the previous system through domain-general reform, provides a new tool for thinking about foundational questions about the origins of thought. It further enables *formal analysis* that suggests more effective *didactic techniques* for inducing instances of such radical conceptual change in humans, as discussed in the next section.

Empirical Predictions in Number Learning

Our model makes several empirical predictions about children's number learning, some of which have already been corroborated by existing experimental work. These include the prediction that low-number counting interventions accelerate the CP induction in three-knowers (Gibson et al., 2020), and that educational curricula emphasizing symbolic-physical analogy training induce faster understanding of fraction concepts than standard curricula (Moss and Case, 1999). These two predictions are instances of a single more abstract prediction that falls out of our model, and is distinct from the predictions made by the main prior model, RSI. Precisely, our model predicts an *asymmetry* with respect to the order of knower-level stages, forecasting that learning certain concepts earlier *makes it easier* to learn certain other concepts later on. In contrast, the RSI model, without any dependence on the previous LoT, predicts that learning certain concepts early *does not* make it easier to learn other concepts—the order of learning is irrelevant, meaning that different orders have *symmetric*, null effects on the ultimate learning rate. In other words, learning in the RSI system is driven entirely by the *distribution* of external tasks, a fundamentally stage-order-*independent* signal, while learning in our model is driven also by *structural dependence* between different symbolic LoTs, which exhibit a clear, graphical order.

Concretely, our model specifically predicts that additional training on symbolic-physical tasks will accelerate arrival at the full-knower stage more than symbolic-only or physical-only training, because recognizing analogical structure between symbolic and physical domains makes it easier to generalize properties of one domain ("-only" properties) to properties of the other domain. Mechanically, this is because it defines a *translation function* that suggests how properties in one domain, e.g. symbolic properties, may be translated to become properties in the other, e.g. physical properties. For example, recognizing the symbolic-physical analogy between the number list and discrete quantities makes it easier to recognize the later-greater principle, a symbolic-only property that enables answering symbolic More tasks ("Which is more, six or nine?" without a physical aid), because it exposes more "training data" in the form of extra number-labeled physical sets. This thus accelerates the conceptual unification of perceived set magnitude order with abstract numeral order. In contrast, under a structure-independent paradigm like RSI, extra symbolic-physical analogy training *and* extra More training will both have the *same*, null effect on the arrival of the full-knower stage. The only exception to this general pattern is that extra training on the very last, "hardest" task that children are observed to solve—in the natural number domain, the Unit task corresponding to the symbolic, exact +1 principle—*will* indeed accelerate full-knower acquisition. This is because all the conceptual stages are all independent of each other in the RSI model, so only the timing of the very final component's acquisition determines the overall learning time. A full comparison of these distinct empirical predictions made by the two models is shown in Figure 11, and additional explanation of these outcomes is provided in Appendix A1.

The second general prediction that falls out of our model is that children who are very good at memorization will often be *slower* at achieving full conceptual understanding. This is

Summary of Model Comparison

Modeling Correlational Data:
Parameter Sensitivity Differences in Predicting Knower-Level Order

	Natural Number Domain	Rational Number Domain
Pre-Analogy Discovery	**Observation**: CP induction emerges after at least 3-knower **UB model**: Less sensitivity **U model**: More sensitivity	**Observation**: CQ induction emerges after learning frac. syntax **UB model**: Less sensitivity **U model**: More sensitivity
Post-Analogy Discovery	**Observation**: Later-greater, exact +1, & inf. knowl. emerge after CP **UB model**: Less sensitivity **U model**: More sensitivity	**Observation**: Infinite divisibility knowl. emerges after CQ **UB model**: Less sensitivity **U model**: More sensitivity

Overall: UB model is **less sensitive** to parameters than U model in predicting observed empirical knower-level order, due to structural dependence on prev. LoT in the model.

Modeling Interventional Data:
Empirical Prediction Differences in Accelerating Knower-Level Timing

	Natural Number Domain	Rational Number Domain
Pre-Analogy Discovery	**Goal**: Accel. CP-knower **UB model**: Low-number training better than high-number **U model**: High-number training better; low-number 0 eff.	**Goal**: Accel. CQ-knower **UB model**: Sym.-phys. training better than symbol-only **U model**: Sym.-phys. training better than symbol-only
Post-Analogy Discovery	**Goal**: Accel. full-knower **UB model**: Sym.-phys. training better than symbol-only **U model**: Sym.-only training has eq. / more eff. vs. sym.-phys.	**Goal**: Accel. full-knower **UB model**: Exists optimal ratio of sym.-phys. vs. sym.-only **U model**: No optimal ratio: Almost all ratios have eq., best eff.

Overall: UB model makes different empirical predictions than U model in almost all categories, predicting that **extra symbolic-physical analogy training** (incl. low-number training) accelerates learning most.

Figure 11: Model Comparison. UB = Utility + Background Proposal (Our Model), U = Utility-Only (Piantadosi et. al.).

because quickly arriving at high task accuracies by memorizing symbolic-only heuristics, without grasping the underlying physical analogy, *decreases* the utility (accuracy) incentive to search for the deeper, more generalizable solution. In contrast, the RSI paradigm predicts that the fastest memorizers will *also* always be fastest at achieving full conceptual understanding. While this prediction needs to be empirically verified, it aligns with the intuitive notion of a distinction between "fast" and "slow" thinkers, where the former prevail at rapid procedure memorization, and the latter succeed on noticing conceptual insights that enable out-of-distribution generalization.

In sum, our model makes several empirical predictions about children's number learning that are structurally distinct from those made by the previous model. This nicely supports the design of experimental paradigms that *tease apart* the two potential learning mechanisms, with ramifications for math education design. In particular, if future intervention experiments reinforce evidence from existing studies in favor of our model, it would further support the implementation of curricula that emphasize *symbolic-physical analogy training* over pure symbolic algorithm memorization, including through lower-level changes such as replacing pie shading tasks with less subtraction-reminiscent physical forms in the fraction setting, among other recommendations (see Appendix A1 for a full list). In future work, we aim to extend our model and test such predictions in the domain of higher-level math and physics as well, including learning concepts like the analogy between algebra and geometry, as symbolized by yet another cultural invention: the *Cartesian plane*, which generalizes the one-dimensional number line.

Beyond Number: Discontinuities in Other Domains

Both of the overarching predictions described in the previous section—that educational emphasis on the analogy between two domains is more effective than training on single-domain properties, and that being too good at memorization may slow down arrival at deeper conceptual understanding—likely also extend to other learning domains beyond number. The most natural extensions of our model are to developmental domains where acquiring new language is also correlated with a discontinuous conceptual leap. Our ongoing work explores this direction in the domain of spatial cognition, where we center our model on the classic "left of blue" reasoning setting developed by Elizabeth Spelke and collaborators. There, children have been shown to fail on a spatial reorientation task requiring them to encode the location of a prize relative to colored wall until around age six, when they learn the relevant *spatial language*, in the form of relative phrases like "left of the blue wall." Spelke has hypothesized that learning such relative spatial constructs, where the reference object may be encoded using non-geometric properties like color, overcomes an innate barrier between a core *geometric* module and core *non-geometric* module for representing space. This initial barrier causes younger children to fail on the task, because they can-

not yet *conjoin* geometric and non-geometric representations. Since children (1) learn *intrinsic* spatial constructs, like "my left hand," years before they learn relative ones, and (2) have also been shown to spontaneously *comprehend* provided relative phrases while still failing nonverbal tasks requiring they *produce* such relative encodings on their own, we model discontinuity in this domain as partially stemming from recognizing the *analogy* between intrinsic and relative spatial *syntaxes* (`my_left(a::Obj)` and `left_of(a::Obj, b::Obj)`) and the corresponding *meanings* of those syntaxes—a first-order relational signal, as in the rational number domain (Figure 4).

In addition to language-guided development in childhood, our paradigm also likely applies to modeling discontinuities in adults' intuitive theories, as well as to settings where the key structural learning signal originates not necessarily from symbolic analogies in explicit language, but other kinds of symbolic structure in the environment. One example of such a domain we have started to explore is modeling discontinuities in people's intuitive theories of *social relations*. In particular, we hypothesize that neurodevelopmental disorders like autism may be related to an inability to undergo such change-of-basis-style conceptual change in one's social theories, potentially caused by impairments in the underlying *background proposal function*. This then leads to compensation measures in the form of memorizing low-level heuristics for social interaction that generalize poorly to new settings. We look forward to exploring these and other extensions further in future work.

Beyond Development: A Model of Genuine, Human-Like Extrapolation, with Implications for AI Design

Finally, beyond developing a formal framework for understanding and predicting human behavior, our work also lays the groundwork for a new path towards designing more human-like *AI systems*. Precisely, today's AI landscape is dominated by large language models (LLMs), which have achieved remarkable success in recent years by scaling up training of the underlying *transformer architecture* to datasets with trillions of tokens. Despite the impressive language and reasoning fluency of these models, however, they remain risky to deploy in truly safety-critical contexts due to lingering, non-human-like hallucinations. In addition, doubts have emerged about whether the existing scaling routes, which involve training these models on ever increasing amounts of data, will be able to overcome this existential unreliability.

Given these challenges, one possibility is to search for answers at the opposite end of the training spectrum: How do children so efficiently learn, and learn from, language? Could decades of empirical data on children's language acquisition reveal insights that could translate into more efficient and trustworthy AI systems? Our work takes a first step in this direction by modeling how children use the rich structure of natural language to *crisply generalize* initially limited meanings *beyond the bounds* of their existing conceptual systems. Put differently, it provides a precise computational-level grounding of the uniquely human capacity for genuine *extrapolation*—the ability to think truly new thoughts—as opposed to merely *abstraction*, or extracting common patterns in *already existing semantics* to express those same meanings more *concisely*, as targeted by past work in program synthesis (Ellis et al., 2021). Our model accomplishes this by exploiting higher-order relational structure in natural language as a *hint* that analogous relational structure may also exist between the underlying meanings of that language. We formalize these analogies as *type casting functions* that *compress* the overall knowledge representation by first *inventing* and then *unifying* shared structure between language and meanings. While LLMs and other statistical learning techniques have long taken advantage of such higher-order signals in language, past program induction techniques have not, using only zeroth-order signals instead (Wong et al., 2021; Figure 4). As such, our model is the first to combine (1) the soundness guarantees and data efficiency of symbolic program induction with (2) the richness of relations in natural language as a learning signal, making a first leap towards explaining even some of the wonder of children's marvelously few-shot language acquisition.

In the future, we are excited to explore whether *hybrid* techniques, combining the extrapolative strengths of symbolic type system induction with the associative flexibility of interpolation-based transformer architectures—especially smaller models, trained on a volume and content of data closer to that consumed by an individual child—may combine the best of these two complementary styles of human learning.

References

Almoammer, A., Sullivan, J., Donlan, C., Marušič, F., Žaucer, R., O'Donnell, T., & Barner, D. (2013). Grammatical morphology as a source of early number word meanings. *Proceedings of the National Academy of Sciences*, *110*(46), 18448–18453. https://doi.org/10.1073/pnas.1313652110

Bloom, P., & Wynn, K. (1997). Linguistic cues in the acquisition of number words. *Journal of Child Language*, *24*(3), 511–533. https://doi.org/10.1017/S0305000997003188

Carey, S. (2009). *The origin of concepts* (1st). Oxford University Press, Inc.

Chu, J., Cheung, P., Schneider, R. M., Sullivan, J., & Barner, D. (2020). Counting to infinity: Does learning the syntax of the count list predict knowledge that numbers are infinite? *Cognitive Science*, *44*(8), e12875. https://doi.org/https://doi.org/10.1111/cogs.12875

Davidson, K., Eng, K., & Barner, D. (2012). Does learning to count involve a semantic induction? *Cognition*, *123*(1), 162–173. https://doi.org/https://doi.org/10.1016/j.cognition.2011.12.013

Dechter, E., Malmaud, J., Adams, R. P., & Tenenbaum, J. B. (2013). Bootstrap learning via modular concept discovery. In F. Rossi (Ed.), *IJCAI 2013, proceedings of the 23rd international joint conference on artificial intelligence, beijing, china, august 3-9, 2013* (pp. 1302–1309). IJCAI/AAAI. http://www.aaai.org/ocs/index.php/IJCAI/IJCAI13/paper/view/6890

Ellis, K., Wong, C., Nye, M., Sablé-Meyer, M., Morales, L., Hewitt, L., Cary, L., Solar-Lezama, A., & Tenenbaum, J. B. (2021). Dreamcoder: Bootstrapping inductive program synthesis with wake-sleep library learning. *Proceedings of the 42nd ACM SIGPLAN International Conference on Programming Language Design and Implementation*, 835–850. https://doi.org/10.1145/3453483.3454080

Fodor, J. (1998, February). Bibliography. In *Concepts: Where cognitive science went wrong*. Oxford University Press. https://doi.org/10.1093/0198236360.001.0001

Frank, M. C., Everett, D. L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from pirahã language and cognition. *Cognition*, *108*(3), 819–824. https://doi.org/https://doi.org/10.1016/j.cognition.2008.04.007

Fuson, K. C. (1998). *Children's counting and concepts of number*. Springer Verlag.

Gelman, R., & Gallistel, C. (1978). *The child's understanding of number*. Harvard Univ. P.

Gibson, D. J., Gunderson, E. A., & Levine, S. C. (2020). Causal effects of parent number talk on preschoolers' number knowledge. *Child Development*, *91*(6), e1162–e1177. https://doi.org/10.1111/cdev.13423

Le Corre, M., Li, P., Huang, B. H., Jia, G., & Carey, S. (2016). Numerical morphology supports early number word learning: Evidence from a comparison of young mandarin and english learners. *Cognitive Psychology*, *88*, 162–186. https://doi.org/https://doi.org/10.1016/j.cogpsych.2016.06.003

Moss, J., & Case, R. (1999). Developing children's understanding of the rational numbers: A new model and an experimental curriculum. *Journal for Research in Mathematics Education*, *30*(2), 122–147. Retrieved January 31, 2026, from http://www.jstor.org/stable/749607

Piantadosi, S. T. (2023). The algorithmic origins of counting. *Child Development*, *94*(6), 1472–1490. https://doi.org/https://doi.org/10.1111/cdev.14031

Piantadosi, S. T., Jara-Ettinger, J., & Gibson, E. (2014). Children's learning of number words in an indigenous farming-foraging group. *Developmental Science*, *17*(4), 553–563. https://doi.org/https://doi.org/10.1111/desc.12078

Piantadosi, S. T., Tenenbaum, J. B., & Goodman, N. D. (2012). Bootstrapping in a language of thought: A formal model of numerical concept learning. *Cognition*, *123*(2), 199–217. https://doi.org/https://doi.org/10.1016/j.cognition.2011.11.005

Sarnecka, B. W., Kamenskaya, V. G., Yamana, Y., Ogura, T., & Yudovina, Y. B. (2007). From grammatical number to exact numbers: Early meanings of 'one', 'two', and 'three' in english, russian, and japanese. *Cognitive Psychology*, *55*(2), 136–168. https://doi.org/https://doi.org/10.1016/j.cogpsych.2006.09.001

Shrager, J., & Siegler, R. S. (1998). Scads: A model of children's strategy choices and strategy discoveries. *Psychological Science*, *9*(5), 405–410. https://doi.org/10.1111/1467-9280.00076

Smith, C. L., Solomon, G. E., & Carey, S. (2005). Never getting to zero: Elementary school students' understanding of the infinite divisibility of number and matter. *Cognitive Psychology*, *51*(2), 101–140. https://doi.org/https://doi.org/10.1016/j.cogpsych.2005.03.001

Sullivan, J., Cramer-Benjamin, S., Alvarez, J., & Barner, D. (2023). Everything is infinite: Children's beliefs about endless space, time, and number. *Open Mind*, *7*, 715–731. https://doi.org/10.1162/opmi_a_00104

Villarroel, J. D., Miñón, M., & Nuño, T. (2011). The origin of counting: A study of the early meaning of 'one', 'two' and 'three' among basque- and spanish-speaking children. *Educational Studies in Mathematics*, *76*, 345–361. https://doi.org/https://doi.org/10.1007/s10649-010-9291-0

Wong, C., Ellis, K., Tenenbaum, J. B., & Andreas, J. (2021). Leveraging language to learn program abstractions and search heuristics. *International Conference on Machine Learning*, 11193–11204.

Appendix A: Model Details

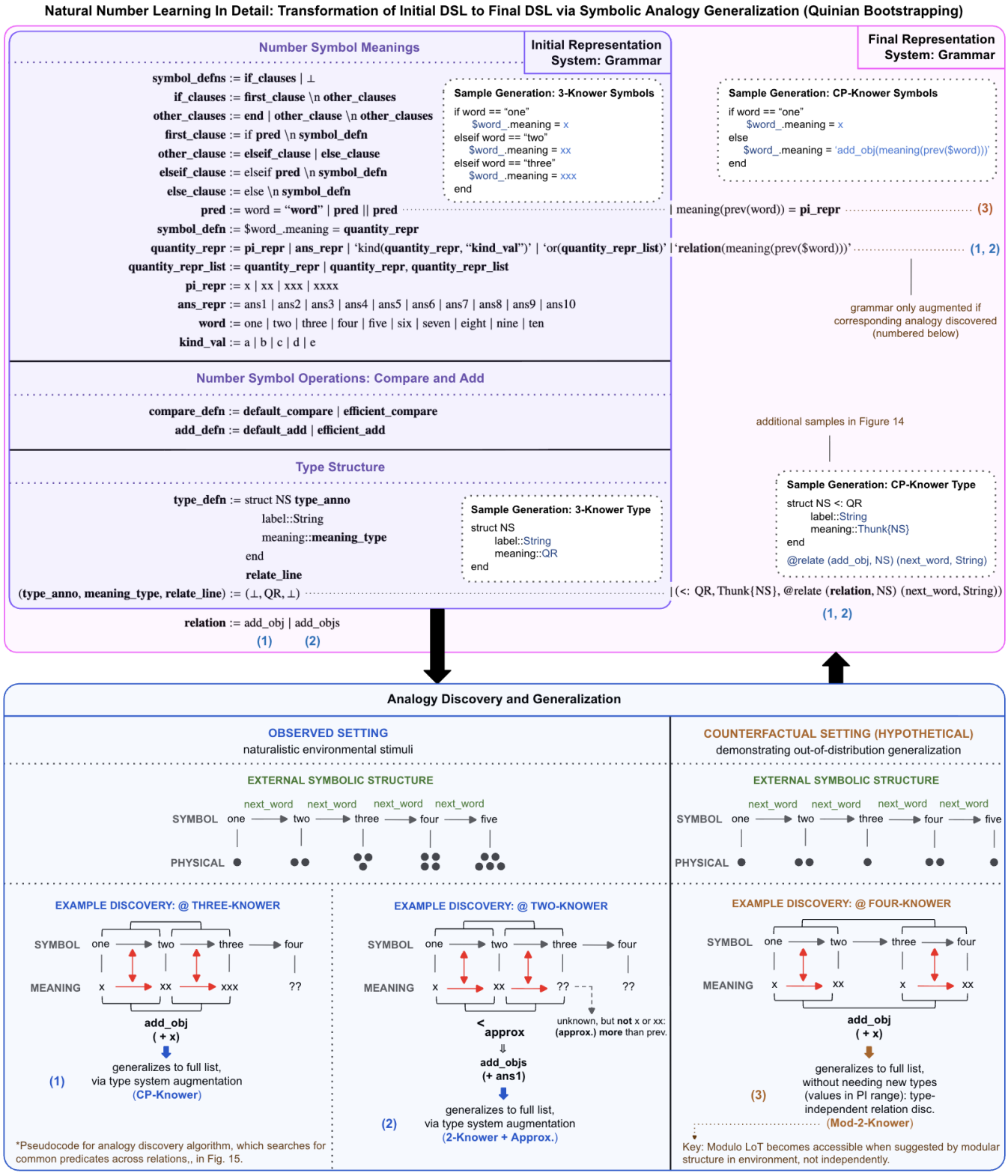

Figure A.1: Details of Natural Number DSL Transformation. In the current implementation, we use a sample of 75 LoTs from the grammar.

Figure A.2: Full Natural Number LoT Template.

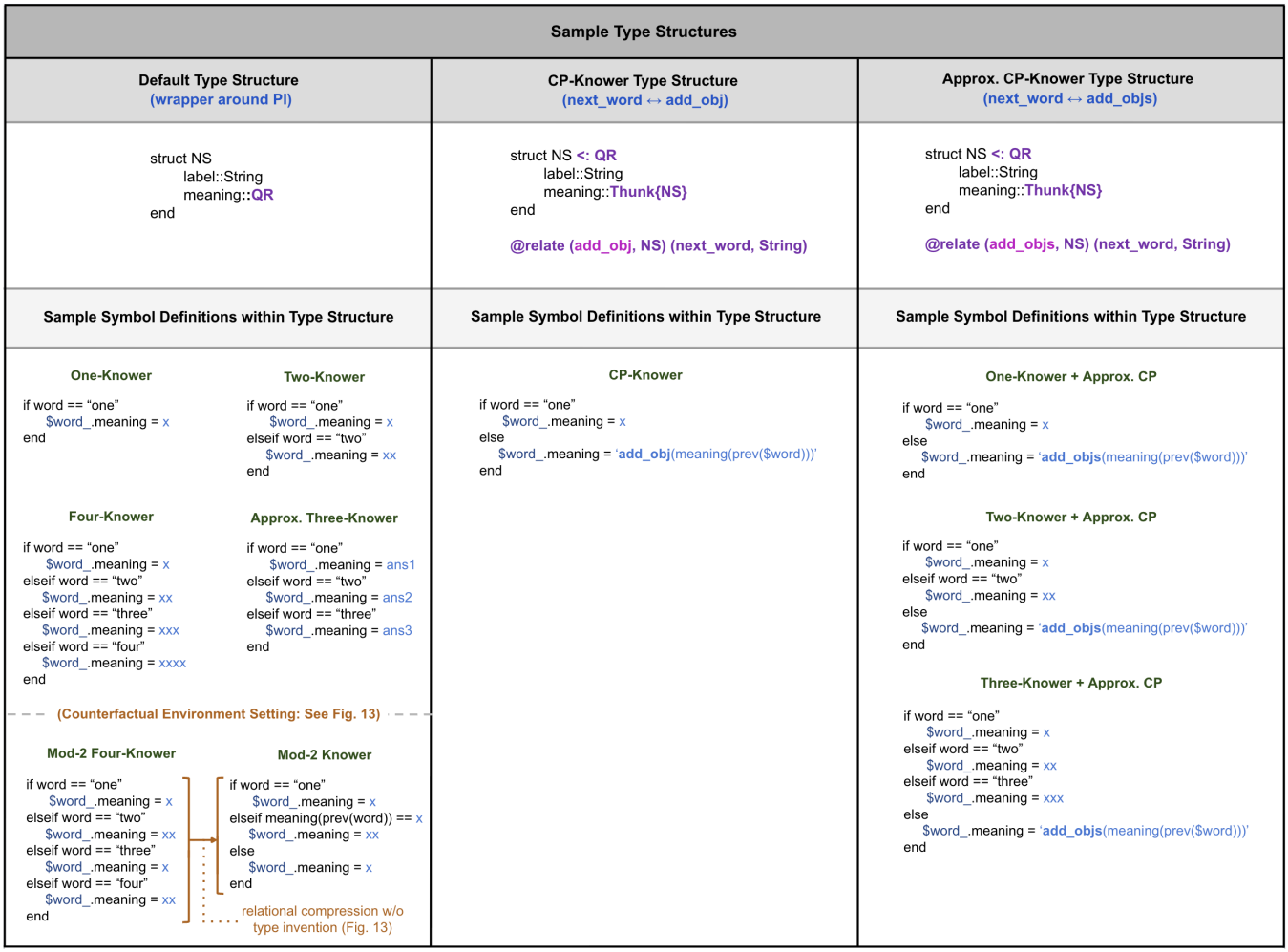

Figure A.3: Sample Generations from the Natural Number LoT Grammar.

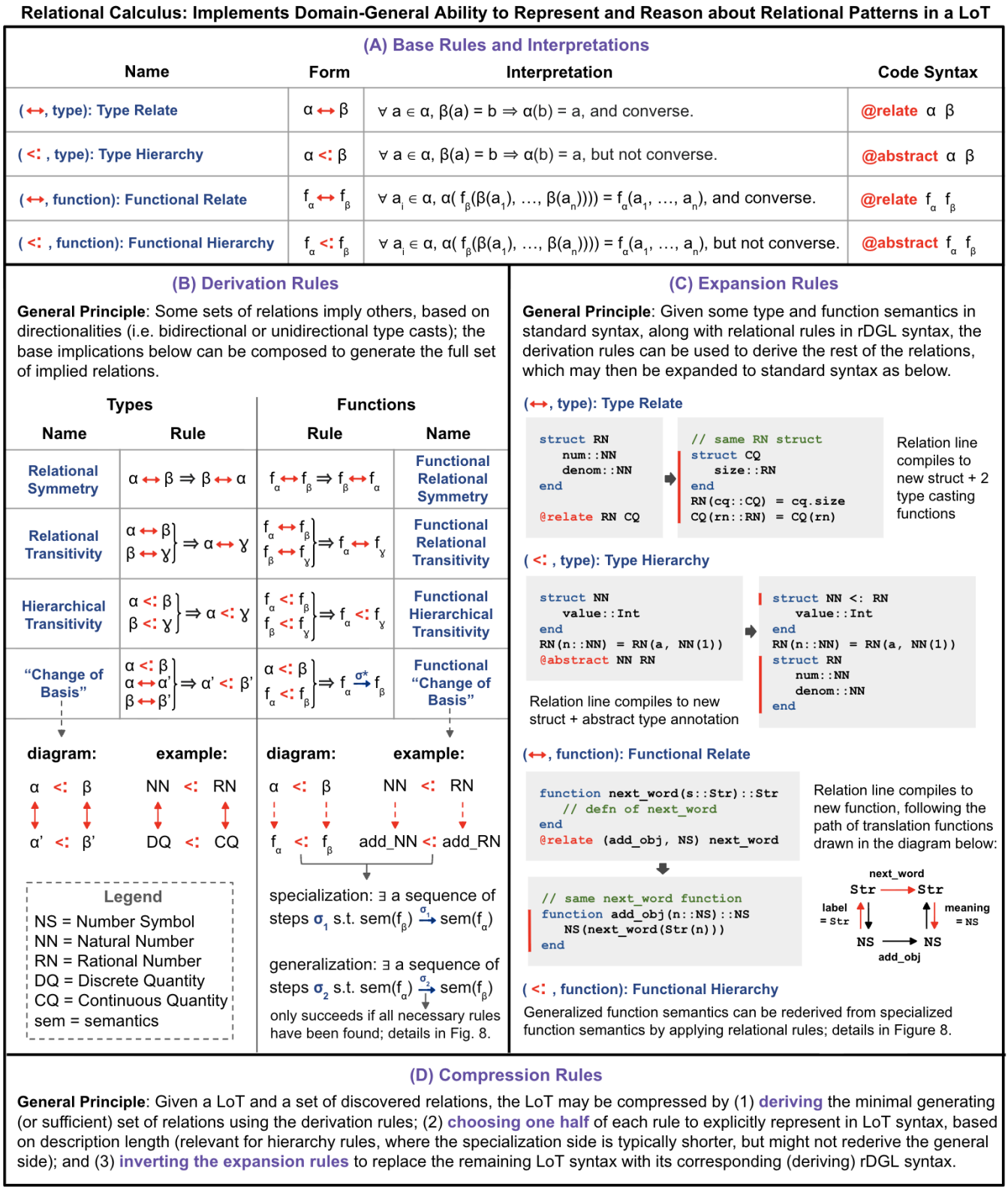

Figure A.4: Details ("Operational Semantics") of the Relational Calculus or Domain-General Language (rDGL).

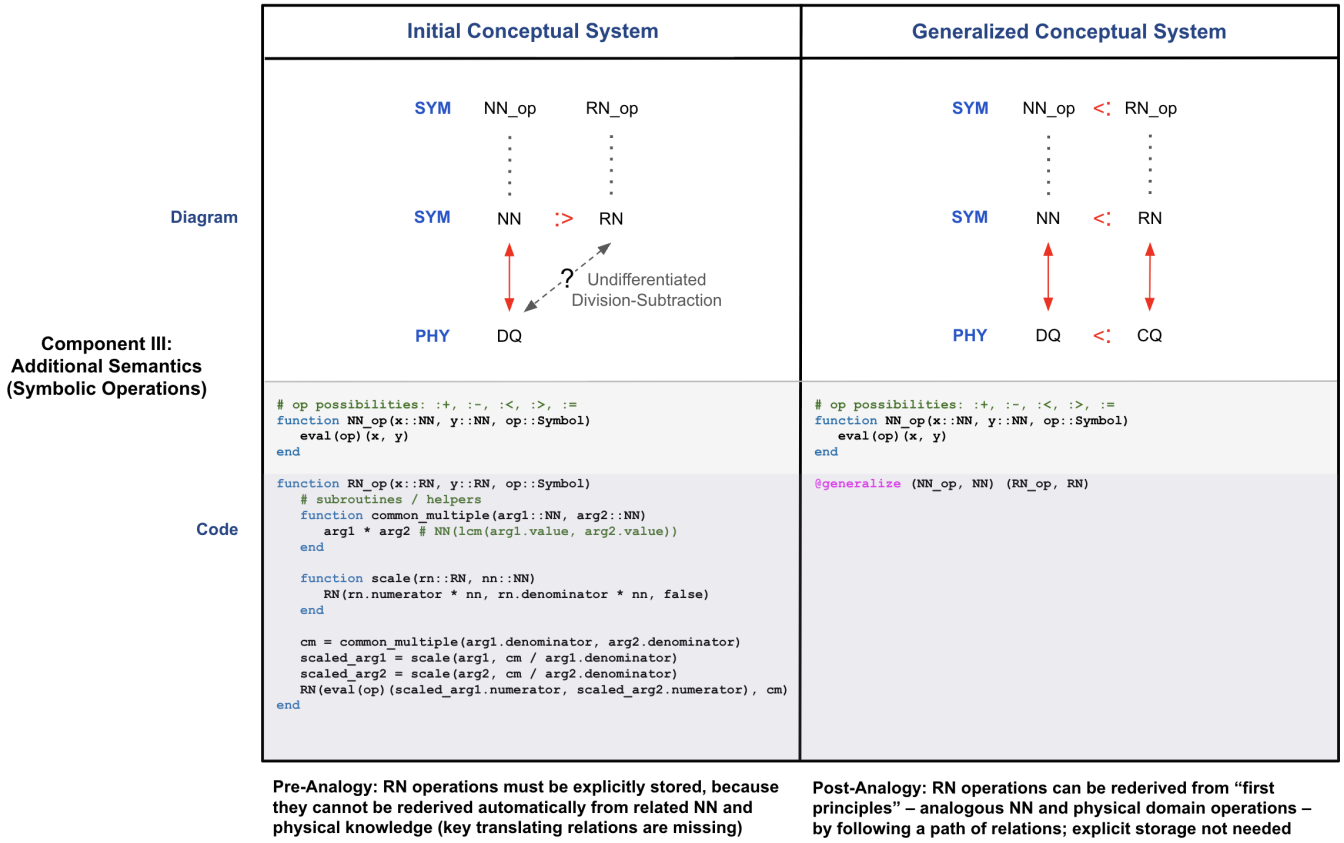

Figure A.5: Implications of Rederivation from First Principles: Storage (Memory Cost) Savings in Post-Analogy Setting.

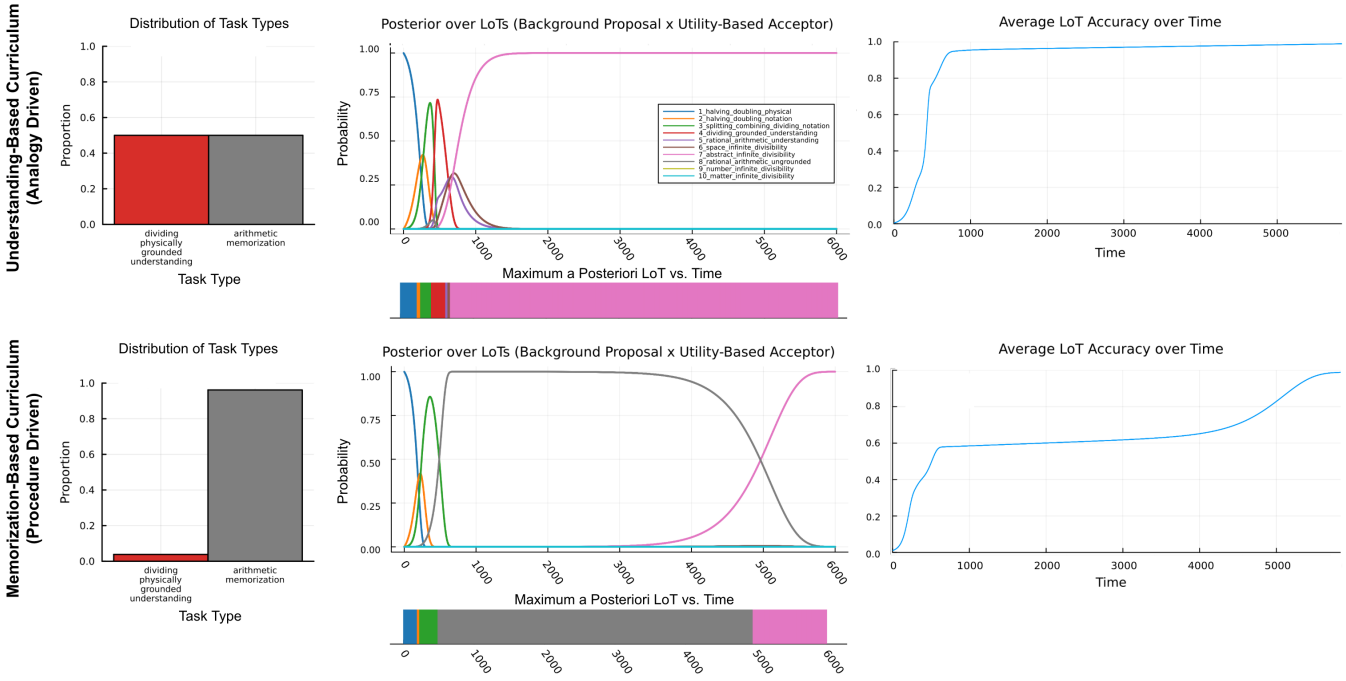

Figure A.6: Curriculum Effects Predicted by the Rational Number Model.

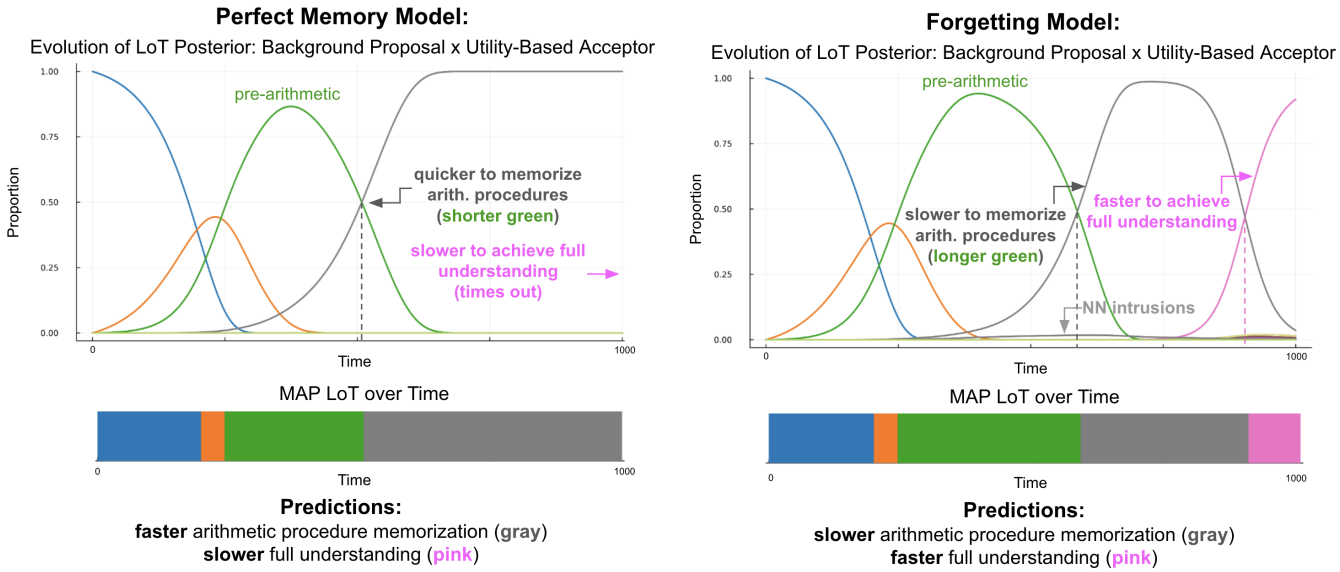

Figure A.7: Memorization Strength Effects in the Rational Number Model.

A1. Additional Recommendations for Mathematics Education Design

In addition to the two high-level predictions about children's mathematics learning discussed above—namely, that (1) symbolic-physical analogy training is more effective than symbolic-only or physical-only training, and (2) children who are too good at memorization may be slower at achieving full conceptual understanding—our formal model also enables *mechanistic analysis* that suggests several lower-level recommendations for teachers. Past work in other learning domains has shown that the most effective didactic aids are ones that are *least ambiguous*, meaning they filter out most alternative explanations, and *most easily interpretable*, as measured by description length in an appropriate language of thought. Our model suggests that the first criterion, ambiguity, is a key obstacle to the effectiveness of traditional curricula for teaching fractions. Specifically, our work formalizes the background knowledge systems that children bring to bear on their fraction lessons as

1. core knowledge, especially the Approximate Number System, since it supports (fuzzy) notions of continuous size, and
2. symbolic natural number knowledge, learned through an earlier instance of Quinian bootstrapping.

The goal of fraction education is to bootstrap children's core, approximate understanding of continuous quantities, as supported by the ANS, to *exact* and *infinite* notions of continuous size, as symbolized by fraction notation (the CQ induction). This is analogous to how the CP induction bootstraps children's core representations of discrete sets, supported by parallel individuation, to ultimately *infinite* notions of exact discrete size, as denoted by natural numbers. In the fraction setting, this bootstrapping is stymied by demonstrative examples that can be "explained away" using symbolic natural number concepts, instead of *uniquely* suggesting a novel reconceptualization of ANS-based continuity. In other words, examples with *alternative* interpretations, using just natural numbers, reinforce the initially much more *accessible* explanation that fractions are some special kind of natural number, and hence discrete quantity, rather than the reverse. Our formalization suggests that a major cause of this ambiguity is demonstrations of the physical meaning of fractions that are compatible with a "discrete operation" semantics for fractions, such as *pre-cut pies* (Figure 6c). In these tasks, children can shade and label pies correctly just by *counting up* pieces until they reach the numerator, where the unshaded "negative space" corresponds to the result of *subtraction*, and where they never need to grasp the deeper concept of continuity. This explanation adds to the long list of other reasons that pies are a poor tool for teaching fractions, including that they make it hard to visualize both (1) fraction operations and (2) fractions over one.

A solution recommended by our model is to instead teach fractions using physical media where the "negative space," corresponding to the difference between the denominator and numerator, is *not the same medium* as the part corresponding to the numerator, such as liquid in a measuring glass (where the negative space is empty, i.e. air). This draws attention away from a discrete-subtraction-based interpretation of fractions, in which children display sensitivity to the number of *remaining* pieces, to one centered on the original medium (liquids are also preferred here since they lack rigid shape). In addition, our work suggests that physical fraction *comparison* tasks using such media where the *numerator is held the same*, and *only the denominator varies*, are most useful for triggering the CQ induction. This is because changes to the numerator can be seen as whole number operations (i.e. discrete addition and subtraction), but changes to the denominator cannot. Instead, the latter represent a concept not present in the natural number system: changing the *unit size* of a single piece. As such, recognizing that the same number of pieces can have different *total (continuous) sizes*, based on the *unit size* of each piece, is key to inducing the CQ leap. This is accelerated by tasks in which children are asked to compare the sizes of several sets each with the same number of items (a), where the unit size of each piece is different, in accordance with a given fraction label ($\frac{a}{b}$). This is visually shown in Figure 6e, where $a = 2$, and $b \in \{3, 4, \ldots, 10\}$: Each set has exactly two pieces, but different overall volumes due to decreasing unit sizes.

Overall, these recommendations are based on a single, more general didactic principle suggested by our model: To induce a discontinuous conceptual change in a child, consider the existing knowledge systems that they bring to bear, and select demonstrative examples that minimize the possibility of the child falling back to explanations using the prior knowledge systems (i.e., examples that minimize ambiguity). In the fraction setting, (1) these past knowledge systems are the ANS and symbolic natural number concepts, (2) the easily accessible but incorrect explanations are based on discrete (natural number) subtraction, and (3) the demonstrative examples that minimize the possibility of these errors are less subtraction-reminiscent physical fraction forms and tasks (e.g. denominator comparison with a constant numerator). Our formalization of all these pieces using the tools of programming languages makes it easier to reason about possible interventions and their impacts on children's learning. As one final example, another such recommendation suggested by our model is to extend methods for emphasizing the symbolic-physical analogy between fraction notation and continuous quantities to higher-order symbolic operations, like fraction addition, too. Though some children accurately memorize these symbolic algorithms as arbitrary sequences of natural number operations around a separator line, emphasizing what each of these operations corresponds to in terms of physical operations—e.g. by explicitly mapping them to transformations along a number line—may help induce deeper understanding.

In sum, our model makes the following recommendations: (1) Use physical media where the "negative space" is not the same medium as the part corresponding to the actual fraction, e.g. liquids; (2) emphasize tasks where children must compare physical fractions with the same numerator, but different denominators; and (3) map symbolic algorithms like fraction addition onto the corresponding physical operations. In future work, we are interested in exploring how our model might *differentially* recommend teaching students at different fraction understanding stages, with different memorization abilities, and other differing properties.

Appendix B: Model Comparison Details

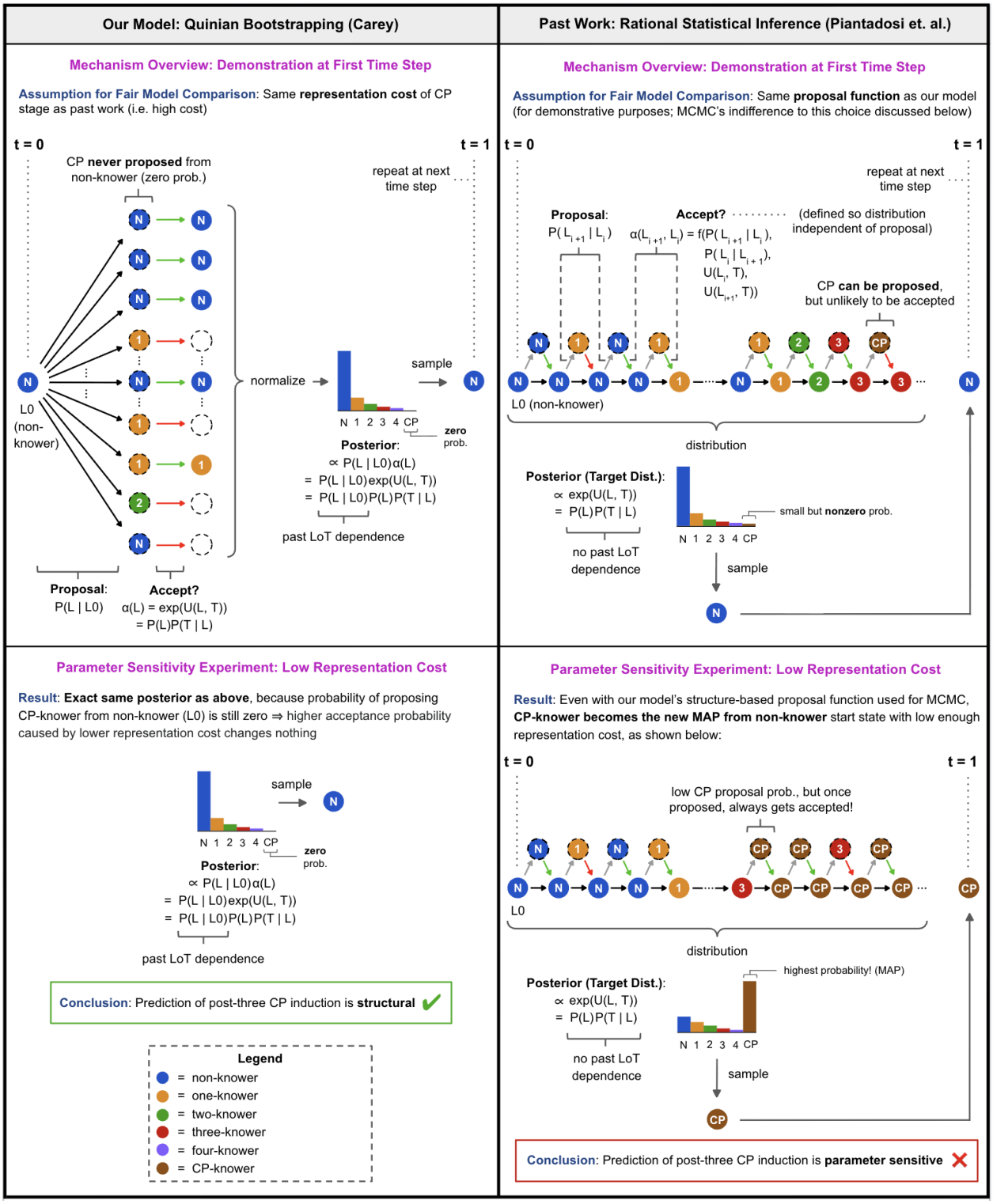

Figure B.1: Mechanism Underlying Parameter Sensitivity Differences between Models. (**Left**) Samples dependent on start state, but not each other ("in parallel"), implying no ability to propose CP-knower at the start ($t = 0$), regardless of the CP cost parameter. (**Right**) Samples locally dependent on each other, but *not* start state ("in series"), meaning it *is* possible to propose CP at the start, due to nonzero probability of *hopping from non-knower to three-knower*, which *can* indeed propose it. This causes sensitivity to the CP cost: If set low enough, CP will become the new MAP, even from non-knower as the previous stage.

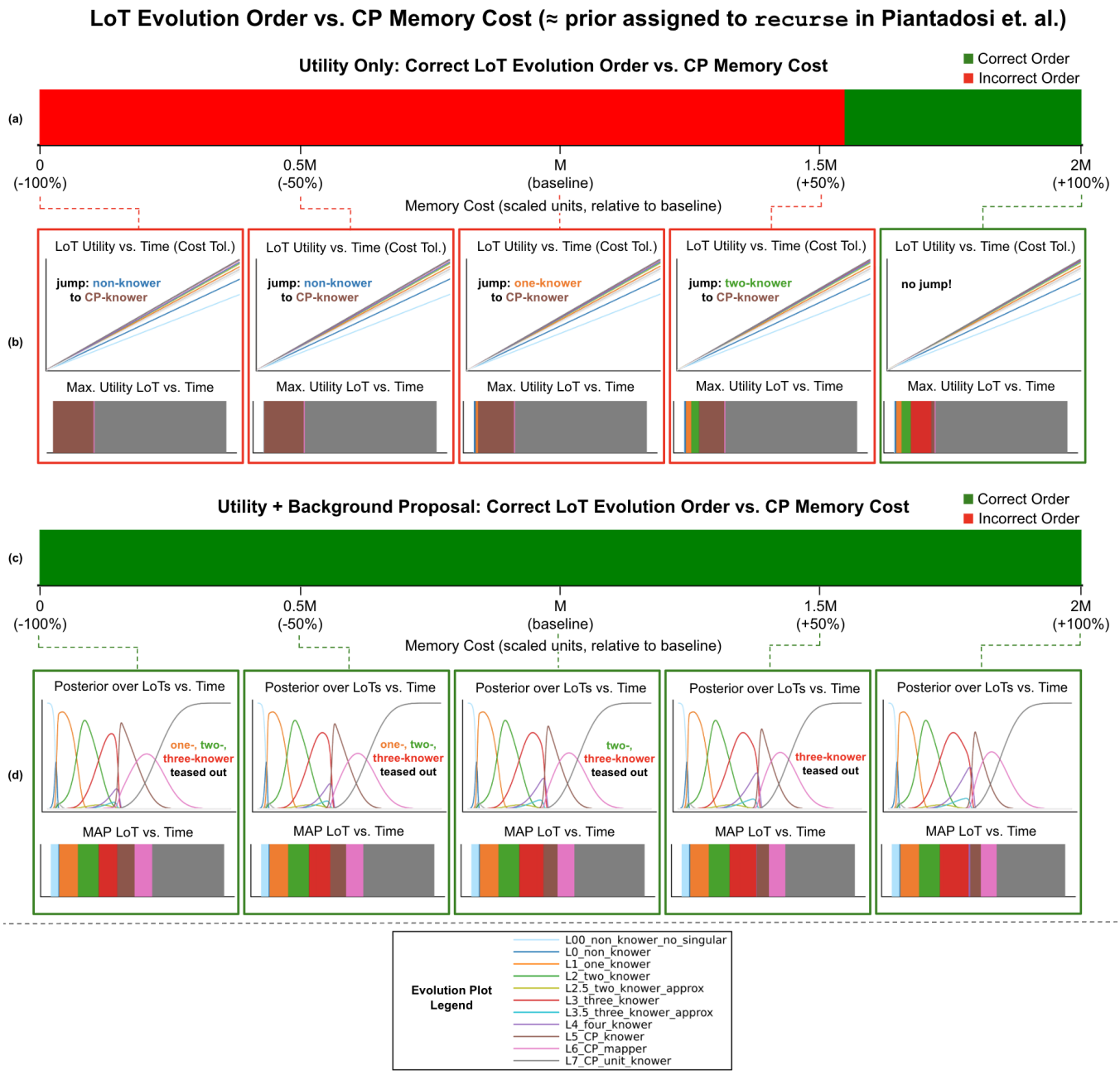

Figure B.2: Natural Number: Comparison of Parameter Sensitivity Differences Between Models, Part I (Horizontal).

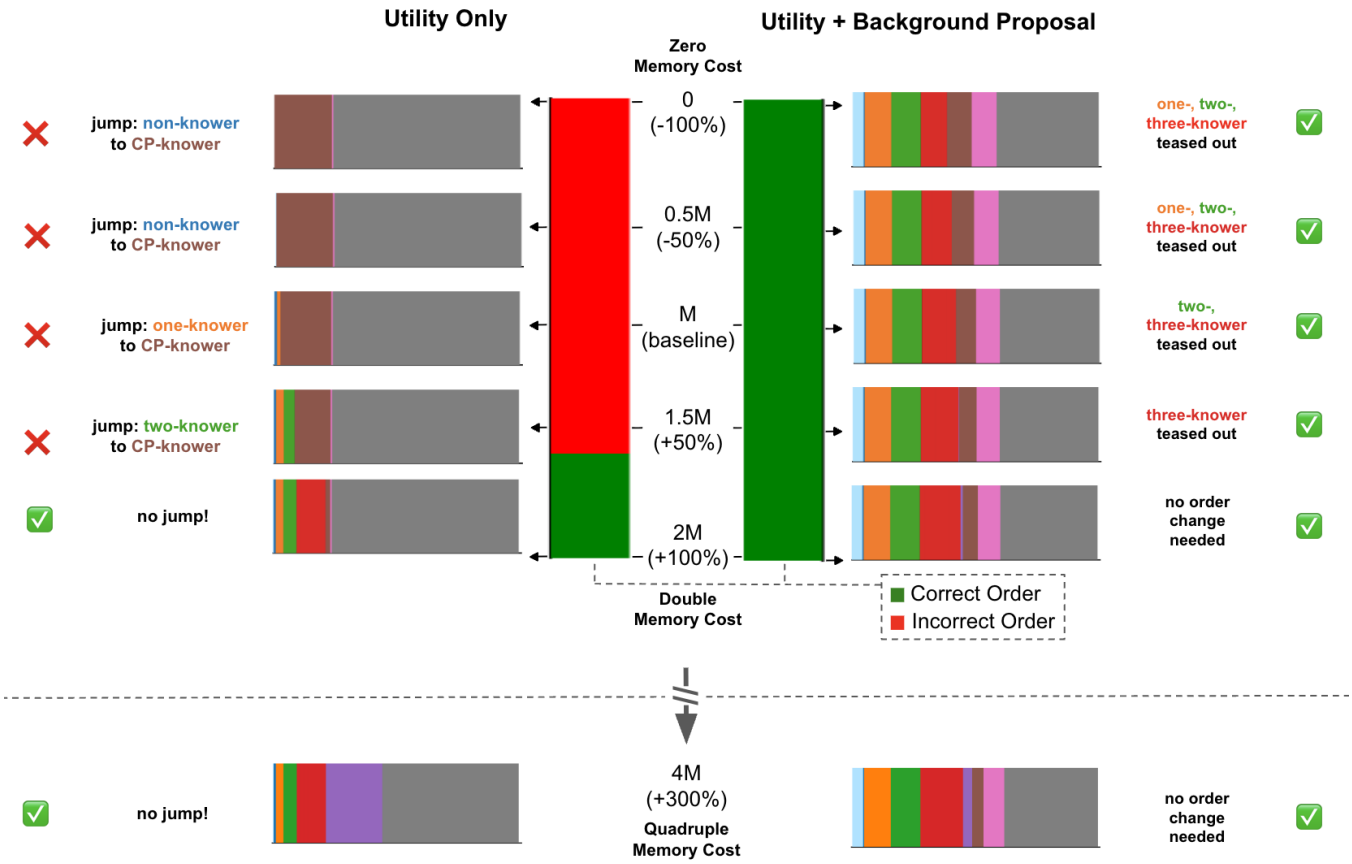

Figure B.3: Natural Number: Comparison of Parameter Sensitivity Differences Between Models, Part II (Vertical).