



General Supervised Learning as Change Propagation with Delta Lenses

Zinovy Diskin^(✉)

McMaster University, Hamilton, Canada
diskinz@mcmaster.ca

Abstract. Delta lenses are an established mathematical framework for modelling and designing bidirectional model transformations (Bx). Following the recent observations by Fong et al, the paper extends the delta lens framework with a new ingredient: learning over a parameterized space of model transformations seen as functors. We will define a notion of an asymmetric learning delta lens with amendment (ala-lens), and show how ala-lenses can be organized into a symmetric monoidal (sm) category. We also show that sequential and parallel composition of well-behaved (wb) ala-lenses are also wb so that wb ala-lenses constitute a full sm-subcategory of ala-lenses.

1 Introduction

The goal of the paper is to develop a formal model of *supervised learning* in a very general context of *bidirectional model transformation* or *Bx*, i.e., synchronization of two arbitrary complex structures (called *models*) related by a transformation.¹ Rather than learning parameterized functions between Euclidean spaces as is typical for machine learning (ML), we will consider learning mappings between model spaces and formalize them as parameterized functors between categories, $f: P \times \mathbf{A} \rightarrow \mathbf{B}$, with P being a parameter space. The basic ML-notion of a *training pair* $(A, B') \in \mathbf{A}_0 \times \mathbf{B}_0$ will be considered as an inconsistency between models caused by a change (*delta*) $v: B \rightarrow B'$ of the target model $B = f(p, A)$, $p \in P$, that was first consistent with A w.r.t. the transformation (functor) $f(p, _)$. An inconsistency is repaired by an appropriate change of the source structure, $u: A \rightarrow A'$, changing the parameter p to p' , and an *amendment* of the target structure $v^\circledast: B' \rightarrow B^\circledast$ so that $f(p', A') = B^\circledast$ is a consistent state of the parameterized two-model system.

The setting above without parameterization and learning (i.e., $p' = p$ always holds), and without amendment ($v^\circledast = \text{id}_{B'}$ always holds), is well known in the Bx literature under the name of *delta lenses*—mathematical structures, in

¹Term *Bx* refers to a wide area including file synchronization, data exchange in databases, and model synchronization in Model-Driven software Engineering (MDE), see [7] for a survey. In the present paper, Bx will mainly refer to Bx in the MDE context.

which consistency restoration via change propagation is modelled by functorial-like algebraic operations over categories [12,6]. There are several types of delta lenses tailored for modelling different synchronization tasks and scenarios, particularly, symmetric and asymmetric. In the paper, we only consider asymmetric delta lenses and will often omit explicit mentioning these attributes. Despite their extra-generality, (delta) lenses have been proved useful in the design and implementation of practical model synchronization systems with triple graph grammars (TGG) [5,2]; enriching lenses with amendment is a recent extension of the framework motivated and formalized in [11]. A major advantage of the lens framework for synchronization is its compositionality: a lens satisfying several equational laws specifying basic synchronization requirements is called *well-behaved* (*wb*), and basic lens theorems state that sequential and parallel composition of *wb* lenses is again *wb*. In practical applications, it allows the designer of a complex synchronizer to avoid integration testing: if elementary synchronizers are tested and proved to be *wb*, their composition is automatically *wb* too.

The present paper makes the following contributions to the delta lens framework for Bx. a) We motivate model synchronization enriched with learning and, moreover, with *categorical* learning, in which the parameter space is a category, and introduce the notion of a *wb asymmetric learning (delta) lens with amendment* (a *wb ala-lens*) (this is the content of Sect. 3). b) We prove compositionality of *wb ala-lenses* and show how their universe can be organized into a symmetric monoidal (*sm*) category (Theorems 1-3 in Sect. 4). All proofs (rather straightforward but notationally laborious) can be found in the long version of the paper [9]. One more compositional result is c) a definition of a *compositional bidirectional transformation language* (Def. 6) that formalizes an important requirement to model synchronization tools, which (surprisingly) is missing from the Bx literature. Background Sect. 2 provides a simple example demonstrating main concepts of Bx and delta lenses in the MDE context. Section 5 briefly surveys related work, and Sect. 6 concludes.

Notation. Given a category \mathbf{A} , its objects are denoted by capital letters A, A' , etc. to recall that in MDE applications, objects are complex structures, which themselves have elements a, a', \dots ; the collection of all objects of category \mathbf{A} is denoted by \mathbf{A}_0 . An arrow with domain $A \in \mathbf{A}_0$ is written as $u: A \rightarrow _$ or $u \in \mathbf{A}(A, _)$; we also write $\text{dom}(u) = A$ (and sometimes $u^{\text{dom}} = A$ to shorten formulas). Similarly, formula $u: _ \rightarrow A'$ denotes an arrow with codomain $u.\text{cod} = A'$. Given a functor $f: \mathbf{A} \rightarrow \mathbf{B}$, its object function is denoted by $f_0: \mathbf{A}_0 \rightarrow \mathbf{B}_0$.

A subcategory $\mathbf{B} \subset \mathbf{A}$ is called *wide* if it has the same objects. All categories we consider in the paper are small.

2 Background: Update propagation and delta lenses

Although Bx ideas work well only in domains conforming to the slogan *any implementation satisfying the specification is good enough* such as code generation (see [10] for discussion), and have limited applications in databases (only so called updatable views can be treated in the Bx-way), we will employ a simple

database example: it allows demonstrating the core ideas without any special domain knowledge required by typical Bx-amenable areas. The presentation will be semi-formal as our goal is to motivate the delta lens formalism that abstracts the details away rather than formalize the example as such.

2.1 Why deltas.

Bx-lenses first appeared in the work on file synchronization, and if we have two sets of strings, say, $B = \{\text{John, Mary}\}$ and $B' = \{\text{Jon, Mary}\}$, we can readily see the difference: $\text{John} \neq \text{Jon}$ but $\text{Mary} = \text{Mary}$. We thus have a structure in-between B and B' (which maybe rather complex if B and B' are big files), but this structure can be recovered by string matching and thus updates can be identified with pairs. The situation dramatically changes if B and B' are object structures, e.g., $B = \{o_1, o_2\}$ with $\text{Name}(o_1) = \text{John}$, $\text{Name}(o_2) = \text{Mary}$ and similarly $B' = \{o'_1, o'_2\}$ with $\text{Name}(o'_1) = \text{Jon}$, $\text{Name}(o'_2) = \text{Mary}$. Now string matching does not say too much: it may happen that o_1 and o'_1 are the same object (think of a typo in the dataset), while o_2 and o'_2 are different (although equally named) objects. Of course, for better matching we could use full names or ID numbers or something similar (called, in the database parlance, primary keys), but absolutely reliable keys are rare, and typos and bugs can compromise them anyway. Thus, for object structures that Bx needs to keep in sync, deltas between models need to be independently specified, e.g., by specifying a *same-ness relation* $u \subset B \times B'$ between models. For example, $u = \{o_1, o'_1\}$ says that $\text{John}@B$ and $\text{Jon}@B'$ are the same person while $\text{Mary}@B$ and $\text{Mary}@B'$ are not. Hence, model spaces in Bx are categories (objects are models and arrows are update/delta specifications) rather than sets (codiscrete categories).

2.2 Consistency restoration via update propagation: An Example

Figure 1 presents a simple example of delta propagation for consistency restoration. Models consist of objects (in the sense of OO programming) with attributes (a.k.a. labelled records), e.g., the source model A consists of three objects identified by their oids (object identifiers) #A, #J, #M (think about employees of some company) with attribute values as shown in the table: attribute *Expr.* refers to Experience measured by a number of years, and *Depart.* is the column of department names. The schema of the table, i.e., the triple S_A of attributes (*Name*, *Expr.*, *Depart.*) with their domains of values **String**, **Integer**, **String** resp., determines a model space \mathbf{A} . A model $X \in \mathbf{A}$ is given by its set of objects OID^X together with three functions Name^X , Expr^X , Depart^X from the same domain OID^X to targets **String**, **Integer**, **String** resp., which are compactly specified by tables as shown for model A . The target model space \mathbf{B} is given by a similar schema S_B consisting of two attributes. The \mathbf{B} -view $\text{get}(X)$ of an \mathbf{A} -model X is computed by selecting those oids $\#O \in \text{OID}^X$ for which $\text{Depart}^X(\#O)$ is an *IT-department*, i.e., an element of the set $IT \stackrel{\text{def}}{=} \{\text{ML, DB}\}$. For example, the upper part of the figure shows the IT-view B of model A .

We assume that all column names in schemas S_A , and S_B are qualified by schema names, e.g., $OID@S_A$, $OID@S_B$ etc, so that schemas are disjoint except elementary domains like *String* etc. Also disjoint are OID-values, e.g., $\#J@A$ and $\#J@B$ are different elements, but constants like John and Mary are elements of set *String* shared by both schemas. To shorten long expressions in the diagrams, we will often omit qualifiers and write $\#J = \#J$ meaning $\#J@A = \#J@B$ or $\#J@B = \#J@B'$ depending on the context given by the diagram; often we will also write $\#J$ and $\#J'$ for such OIDs. Also, when we write $\#J = \#J$ inside block arrows denoting updates, we actually mean a pair, e.g., $(\#J@B, \#J@B')$.

Given two models over the same schema, say, B and B' over S_B , an update $v: B \rightarrow B'$ is a relation $v \subset OID^B \times OID^{B'}$; if a schema contains several nodes, an update should provide a relation v_N for each node N in the schema.

Note an essential difference between the two parallel updates $v_1, v_2: B \rightarrow B'$ specified in the figure. Update v_1 says that John's name was changed to Jon (think of fixing a typo), and the experience data for Mary were also corrected (either because of a typo or, e.g., because the department started to use a new ML method for which Mary has a longer experience). Update v_2 specifies the same story for John but a new story for Mary: it says that Mary $\#M$ left the IT-view and Mary $\#M'$ is a new employee in one of IT-departments.

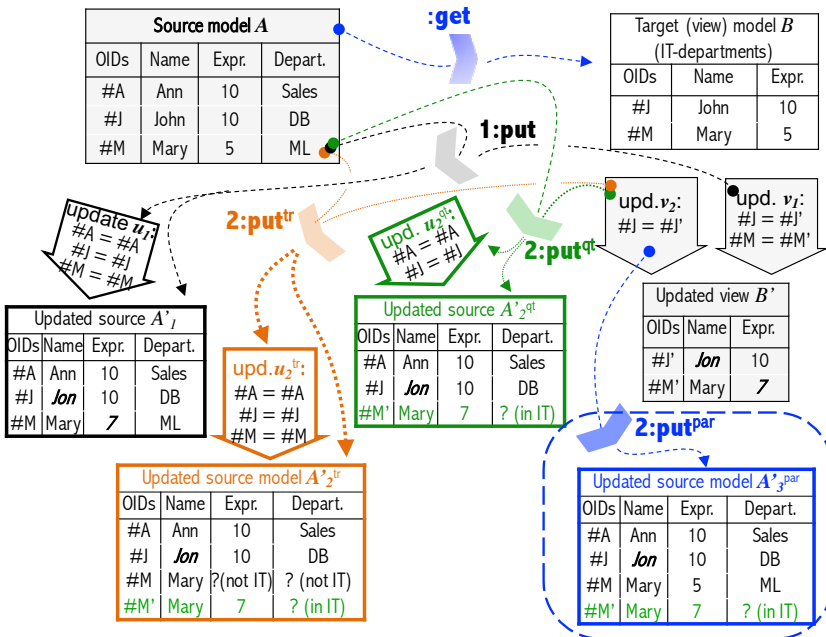


Fig. 1: Example of update propagation

2.3 Update propagation and update policies

The updated view B' is inconsistent with the source A and the latter is to be updated accordingly — we say that update v is to be propagated back to A . Propagation of v_1 is easy: we just update accordingly the values of the attributes as shown in the figure in the block arrow $u_1: A \rightarrow A'_1$ (of black colour). Importantly, propagation needs two pieces of data: the view update v_1 and the original state A of the source as shown in the figure by two data-flow lines into the chevron $1:\text{put}$; the latter denotes invocation of the backward propagation operation put (read “put view update back to the source”). The quadruple $1 = (v_1, A, u_1, A')$ can be seen as an *instance* of operation put , hence the notation $1:\text{put}$ (borrowed from the UML).

Propagation of update v_2 is more challenging: Mary can disappear from the IT-view because a) she quit the company, b) she transitioned to a non-IT department, and c) the view definition has changed, e.g., the new view must only show employees with experience more than 5 years. Choosing between these possibilities is often called choosing an (*update*) *policy*. We will consider the case of changing the view in Sect. 3, and in the current section discuss policies a) and b) (ignore for a while the propagation scenario shown in blue in the right lower corner of the figure that shows policy c)).

For policy a), further referred to as *quitting* and briefly denoted by qt , the result of update propagation is shown in the figure with green colour: notice the update (block) arrow u_2^{qt} and its result, model $A'_2{}^{\text{qt}}$, produced by invoking operation put^{qt} . Note that while we know the new employee Mary works in one of IT departments, we do not know in which one. This is specified with a special value ‘?’ (a.k.a. labelled null in the database parlance).

For policy b), further referred to as *transition* and denoted tr , the result of update propagation is shown in the figure with orange colour: notice update arrow u_2^{tr} and its result, model $A'_2{}^{\text{tr}}$ produced by put^{tr} . Mary $\#M$ is the old employee who transitioned to a new non-IT department, for which her expertise is unknown. Mary $\#M'$ is a new employee in one of IT-departments (we assume that the set of departments is not exhausted by those appearing in a particular state $A \in \mathbf{A}$). There are also updates whose backward propagation is uniquely defined and does not need a policy, e.g., update v_1 is such.

An important property of update propagations we have considered is that they restore consistency: the view of the updated source equals to the updated view initiated the update: $\text{get}(A') = B'$; moreover, this equality extends for update arrows: $\text{get}(u_i) = v_i$, $i = 1, 2$. Such extensions can be derived from view definitions if the latter are determined by so called monotonic queries (which encompass a wide class of practically useful queries including the Select-Project-Join class). For views defined by non-monotonic queries, in order to obtain get 's action on source updates $u: A \rightarrow A'$, a suitable policy is to be added to the view definition (see [1,14,12] for details and discussion). Moreover, normally get preserves identity updates, $\text{get}(\text{id}_A) = \text{id}_{\text{get}(A)}$, and update composition: for any $u: A \rightarrow A'$ and $u': A' \rightarrow A''$, equality $\text{get}(u; u') = \text{get}(u); \text{get}(u')$ holds.

2.4 Delta lenses

Our discussion of the example can be summarized in the following algebraic terms. We have two categories of *models* and *updates*, \mathbf{A} and \mathbf{B} , and a functor $\text{get}: \mathbf{A} \rightarrow \mathbf{B}$ incrementally computing \mathbf{B} -views of \mathbf{A} -models (we will often write $A.\text{get}$ for $\text{get}(A)$). We also suppose that for a chosen update policy, we have worked out precise procedures for how to propagate any view update backwards. This gives us a family of operations $\text{put}_A: \mathbf{A}(A, _) \leftarrow \mathbf{B}(A.\text{get}, _)$ indexed by \mathbf{A} -objects, $A \in \mathbf{A}_0$, for which we write $\text{put}_{A.v}$ or $\text{put}_A(v)$ interchangeably.

Definition 1 (Delta Lenses ([12])) Let \mathbf{A}, \mathbf{B} be two categories. An (*asymmetric delta*) *lens* from \mathbf{A} (the source of the lens) to \mathbf{B} (the target) is a pair $\ell = (\text{get}, \text{put})$, where $\text{get}: \mathbf{A} \rightarrow \mathbf{B}$ is a functor and put is a family of operations $\text{put}_A: \mathbf{A}(A, _) \leftarrow \mathbf{B}(A.\text{get}, _)$ indexed by objects of \mathbf{A} , $A \in \mathbf{A}_0$. Given A , operation put_A maps any arrow $v: A.\text{get} \rightarrow B'$ to an arrow $u: A \rightarrow A'$ such that $A'.\text{get} = B'$. The last condition is called (co)discrete Putget law:

$$(\text{Putget})_0 \quad (\text{put}_{A.v}).\text{cod}.\text{get}_0 = v.\text{cod} \text{ for all } A \in \mathbf{A}_0 \text{ and } v \in \mathbf{B}(A.\text{get}, _)$$

where get_0 denotes the object function of functor get . We will write a lens as an arrow $\ell: \mathbf{A} \rightarrow \mathbf{B}$ going in the direction of get .

Note that family put corresponds to a chosen update policy, e.g., in terms of the example above, for the same view functor get , we have two families of put -operations, put^{qt} and put^{tr} , corresponding to the two updated policies we discussed. These two policies determine two lenses $\ell^{\text{qt}} = (\text{get}, \text{put}^{\text{qt}})$ and $\ell^{\text{tr}} = (\text{get}, \text{put}^{\text{tr}})$ sharing the same get .

Definition 2 (Well-behavedness) A (*lens*) *equational law* is an equation to hold for all values of two variables: $A \in \mathbf{A}_0$ and $v: A.\text{get} \rightarrow T'$. A lens is called *well-behaved* (*wb*) if the following two laws hold:

$$(\text{Stability}) \quad \text{id}_A = \text{put}_{A.\text{id}_{A.\text{get}}} \text{ for all } A \in \mathbf{A}_0$$

$$(\text{Putget}) \quad (\text{put}_{A.v}).\text{get} = v \text{ for all } A \in \mathbf{A}_0 \text{ and all } v \in \mathbf{B}(A.\text{get}, _)$$

Remark 1. Stability law says that a *wb* lens does nothing if nothing happens on the target side (no actions without triggers). Putget requires consistency after the backward propagation is finished. Note the distinction between the Putget_0 condition included into the very definition of a lens, and the full Putget law required for the *wb* lenses. The former is needed to ensure smooth tiling of put -squares (i.e., arrow squares describing application of put to a view update and its result) both horizontally (for sequential composition) and vertically (not considered in the paper). The full Putget assures true consistency as considering a state B' alone does not say much about the real update and elements of B' cannot be properly interpreted. The real story is specified by delta $v: B \rightarrow B'$, and consistency restoration needs the full (PutGet) law as above.²

A more detailed trailer of lenses can be found in the long version [9].

²As shown in [6], the Putget_0 condition is needed if we want to define operations put separately from the functor get : then we still need a function $\text{get}_0: \mathbf{A}_0 \rightarrow \mathbf{B}_0$ and the codiscrete Putget law to ensure a reasonable behaviour of put .

3 Asymmetric Learning Lenses with Amendments

We will begin with a brief motivating discussion, and then proceed with formal definitions

3.1 Does Bx need categorical learning?

Enriching delta lenses with learning capabilities has a clear practical sense for Bx. Having a lens $(\text{get}, \text{put}): \mathbf{A} \rightarrow \mathbf{B}$ and inconsistency $A.\text{get} \neq B'$, the idea of learning extends the notion of the search space and allows us to update the transformation itself so that the final consistency is achieved for a new transformation get' : $A.\text{get}' = B'$. For example, in the case shown in Fig. 1, disappearance of Mary #M in the updated view B' can be caused by changing the view definition, which now requires to show only those employees whose experience is more than 5 years and hence Mary #M is to be removed from the view, while Mary #M' is a new IT-employee whose experience satisfies the new definition. Then the update v_2 can be propagated as shown in the bottom right corner of Fig. 1, where index `par` indicates a new update policy allowing for view definition (parameter) change.

To manage the extended search possibilities, we parameterize the space of transformations as a family of mappings $\text{get}_p: \mathbf{A} \rightarrow \mathbf{B}$ indexed over some parameter space $p \in \mathbf{P}$. For example, we may define the IT-view to be parameterized by the experience of employees shown in the view (including *any* experience as a special parameter value). Then we have two interrelated propagation operations that map an update $B \rightsquigarrow B'$ to a parameter update $p \rightsquigarrow p'$ and a source update $A \rightsquigarrow A'$. Thus, the extended search space allows for new update policies that look for updating the parameter as an update propagation possibility. The possibility to update the transformation appears to be very natural in at least two important Bx scenarios: a) model transformation design and b) model transformation evolution (cf. [21]), which necessitates the enrichment of the delta lens framework with parameterization and learning. Note that all transformations get_p , $p \in \mathbf{P}$ are to be elements of the same lens, and operations `put` are *not* indexed by p , hence, formalization of learning by considering a family of ordinary lenses would *not* do the job.

Categorical vs. codiscrete learning Suppose that the parameter p is itself a set, e.g., the set of departments forming a view can vary depending on some context. Then an update from p to p' has a relational structure as discussed above, i.e., $e: p \rightarrow p'$ is a relation $e \subset p \times p'$ specifying which departments disappeared from the view and which are freshly added. This is a general phenomenon: as soon as parameters are structures (sets of objects or graphs of objects and attributes), a parameter change becomes a structured delta and the space of parameters gives rise to a category \mathbf{P} . The search/propagation procedure returns an arrow $e: p \rightarrow p'$ in this category, which updates the parameter value from p to p' . Hence, a general model of supervised learning should assume \mathbf{P} to be a category (and we say that learning is *categorical*). The case of the parameter

space being a set is captured by considering a codiscrete category \mathbf{P} whose only arrows are pairs of its objects; we call such learning *codiscrete*.

3.2 Ala-lenses

The notion of a *parameterized functor* (*p-functor*) is fundamental for ala-lenses, but is not a lens notion per se and is thus placed into Appendix Sect. A.1. We will work with its exponential (rather than equivalent product-based) formulation but will do uncurrying and currying back if necessary, and often using the same symbol for an arrow f and its uncurried version \check{f} .

Definition 3 (ala-lenses) Let \mathbf{A} and \mathbf{B} be categories. An *ala-lens* from \mathbf{A} (the *source* of the lens) to \mathbf{B} (the *target*) is a pair $\ell = (\text{get}, \text{put})$ whose first component is a p-functor $\text{get}: \mathbf{A} \xrightarrow{\mathbf{P}} \mathbf{B}$ and the second component is a triple of (families of) operations $\text{put} = (\text{put}_{p,A}^{\text{upd}}, \text{put}_{p,A}^{\text{req}}, \text{put}_{p,A}^{\text{self}})$ indexed by pairs $p \in \mathbf{P}_0$, $A \in \mathbf{A}_0$; arities of the operations are specified below after we introduce some notation. Names req (for 'request') and upd (for 'update') are chosen to match the terminology in [17].

Categories \mathbf{A} , \mathbf{B} are called *model spaces*, their objects are *models* and their arrows are (*model*) *updates* or *deltas*. Objects of \mathbf{P} are called *parameters* and are denoted by small letters p, p', \dots rather than capital ones to avoid confusion with [17], in which capital P is used for the entire parameter set. Arrows of \mathbf{P} are called *parameter deltas*. For a parameter $p \in \mathbf{P}_0$, we write get_p for the functor $\text{get}(p): \mathbf{A} \rightarrow \mathbf{B}$ (read "get \mathbf{B} -views of \mathbf{A} "), and if $A \in \mathbf{A}_0$ is a source model, its get_p -view is denoted by $\text{get}_p(A)$ or $A.\text{get}_p$ or even A_p (so that $_{-p}$ becomes yet another notation for functor get_p). Given a parameter delta $e: p \rightarrow p'$ and a source model $A \in \mathbf{A}_0$, the model delta $\text{get}(e): \text{get}_p(A) \rightarrow \text{get}_{p'}(A)$ will be denoted by $\text{get}_e(A)$ or e_S (rather than A_e as we would like to keep capital letters for objects only). In the uncurried version, $\text{get}_e(A)$ is nothing but $\check{\text{get}}(e, \text{id}_S)$.

Since get_e is a natural transformation, for any delta $u: A \rightarrow A'$ we have a commutative square $e_S; u_{p'} = u_p; e_{A'}$ (whose diagonal is $\check{\text{get}}(e, u)$). We will denote the diagonal of this square by $u.\text{get}_e$ or $u_e: A_p \rightarrow A'_{p'}$. Thus, we use notation

$$(1) \quad \begin{aligned} A_p &\stackrel{\text{def}}{=} A.\text{get}_p \stackrel{\text{def}}{=} \text{get}_p(A) \stackrel{\text{def}}{=} \text{get}(p)(A) \\ u_e &\stackrel{\text{def}}{=} u.\text{get}_e \stackrel{\text{def}}{=} \text{get}_e(u) \stackrel{\text{def}}{=} \text{get}(e)(u) \stackrel{\text{def}}{=} e_S; u_{p'} \stackrel{\text{nat}}{=} u_p; e_{A'}: A_p \rightarrow A'_{p'} \end{aligned}$$

Now we describe operations put . They all have the same indexing set $\mathbf{P}_0 \times \mathbf{A}_0$, and the same domain: for any index p, A and any model delta $v: A_p \rightarrow B'$ in \mathbf{B} , the value $\text{put}_{p,A}^{\mathbf{x}}(p, A)$, $\mathbf{x} \in \{\text{req}, \text{upd}, \text{self}\}$ is defined and unique:

$$(2) \quad \begin{aligned} \text{put}_{p,A}^{\text{upd}}: p \rightarrow p' &\quad \text{is a parameter delta from } p, \\ \text{put}_{p,A}^{\text{req}}: A \rightarrow A' &\quad \text{is a model delta from } A, \\ \text{put}_{p,A}^{\text{self}}: B' \rightarrow A'_{p'} &\quad \text{is a model delta from } B' \\ &\quad \text{called the } \textit{amendment} \text{ and denoted by } v^{\textcircled{a}}. \end{aligned}$$

Note that the definition of put^{self} involves an equational dependency between all three operations: for all $A \in \mathbf{A}_0$, $v \in \mathbf{B}(A.\text{get}, _)$, we require

$$(\text{Putget})_0 \quad (\text{put}_A^{\text{req}}.v).\text{cod}.\text{get}_{p'} = (v; \text{put}_A^{\text{self}}).\text{cod} \text{ where } p' = (\text{put}_A^{\text{upd}}.v).\text{cod}$$

We will write an ala-lens as an arrow $\ell = (\text{get}, \text{put}): \mathbf{A} \xrightarrow{\mathbf{P}} \mathbf{B}$.

A lens is called *(twice) codiscrete* if categories \mathbf{A} , \mathbf{B} , \mathbf{P} are codiscrete and thus $\text{get}: \mathbf{A} \xrightarrow{\mathbf{P}} \mathbf{B}$ is a parameterized function. If only \mathbf{P} is codiscrete, we call ℓ a *codiscretely learning delta lens*, while if only model spaces are codiscrete, we call ℓ a *categorically learning codiscrete lens*.

Diagram in Fig. 2 shows how a lens' operations are interrelated. The upper part shows an arrow $e: p \rightarrow p'$ in category \mathbf{P} and two corresponding functors from \mathbf{A} to \mathbf{B} . The lower part is to be seen as a 3D-prism with visible front face $AA_{p'}A'_pA'$ and visible upper face $AA_pA_{p'}$, the bottom and two back faces are invisible, and the corresponding arrows are dashed. The prism denotes an algebraic term: given elements are shown with black fill and white font while derived elements are blue (recalls being mechanically computed) and blank (double-body arrows are considered as "blank"). The two pairs of arrows originating from A and A' are not blank because they denote pairs of nodes (the UML says *links*) rather than mappings/deltas between nodes.

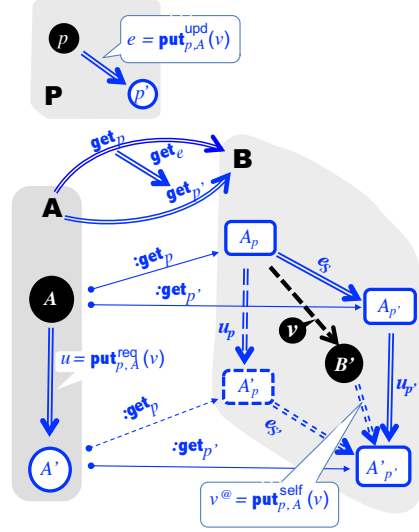


Fig. 2: Ala-lens operations

Equational definitions of deltas $e, u, v^@$ are written up in the three callouts near them. The right back face of the prism is formed by the two vertical derived deltas $u_p = u.\text{get}_p$ and $u_{p'} = u.\text{get}_{p'}$, and the two matching them horizontal derived deltas $e_S = \text{get}_e(A)$ and $e_{A'} = \text{get}_e(A')$; together they form a commutative square due to the naturality of $\text{get}(e)$ as explained earlier.

Definition 4 (Well-behavedness) An ala-lens is called *well-behaved (wb)* if the following two laws hold for all $p \in \mathbf{P}_0$, $A \in \mathbf{A}_0$ and $v: A_p \rightarrow B'$:

- (Stability) if $v = \text{id}_{A_p}$ then all three propagated updates $e, u, v^@$ are identities:
 $\text{put}_{p,A}^{\text{upd}}(\text{id}_{A_p}) = \text{id}_p$, $\text{put}_{p,A}^{\text{req}}(\text{id}_{A_p}) = \text{id}_S$, $\text{put}_{p,A}^{\text{self}}(\text{id}_{A_p}) = \text{id}_{A_p}$
- (Putget) $(\text{put}_{p,A}^{\text{req}}.v).\text{get}_e = v; v^@$ where $e = \text{put}_{p,A}^{\text{upd}}(v)$ and $v^@ = \text{put}_{p,A}^{\text{self}}(v)$

Remark 2. Note that Remark 1 about the Putget law is again applicable.

Example 1 (Identity lenses). Any category \mathbf{A} gives rise to an ala-lens $\text{id}_{\mathbf{A}}$ with the following components. The source and target spaces are equal to \mathbf{A} , and

the parameter space is $\mathbf{1}$. Functor `get` is the identity functor and all `puts` are identities. Obviously, this lens is `wb`.

Example 2 (Iso-lenses). Let $\iota: \mathbf{A} \rightarrow \mathbf{B}$ be an isomorphism between model spaces. It gives rise to a `wb ala-lens` $\ell(\iota): \mathbf{A} \rightarrow \mathbf{B}$ with $\mathbf{P}^{\ell(\iota)} = \mathbf{1} = \{*\}$ as follows. Given any A in \mathbf{A} and $v: \iota(A) \rightarrow B'$ in \mathbf{B} , we define $\text{put}_{*,A}^{\ell(\iota).\text{req}}(v) = \iota^{-1}(v)$ while the two other `put` operations map v to identities.

Example 3 (Bx lenses). Examples of `wb aa-lenses` modelling a `Bx` can be found in [11]: they all can be considered as `ala-lenses` with a trivial parameter space $\mathbf{1}$.

Example 4 (Learners). Learners defined in [17] are codiscretely learning codiscrete lenses with amendment, and as such satisfy (the amended) `Putget` (Remark 1). Looking at the opposite direction, `ala-lenses` are a categorification of learners as detailed in Fig. 8 on p. 194.

4 Compositionality of ala-lenses

This section explores the compositional structure of the universe of `ala-lenses`; especially interesting is their sequential composition. We will begin with a small example demonstrating sequential composition of ordinary lenses and showing that the notion of update policy transcends individual lenses. Then we define sequential and parallel composition of `ala-lenses` (the former is much more involved than for ordinary lenses) and show that `wb ala-lenses` can be organized into an `sm-category`. Finally, we formalize the notion of a compositional update policy via the notion of a compositional bidirectional language.

4.1 Compositionality of update policies: An example

Fig. 3 extends the example in Fig. 1 with a new model space \mathbf{C} whose schema consists of the only attribute `Name`, and a view of the `IT-view`, in which only employees of the `ML` department are to be shown. Thus, we now have two functors, `get1: A → B` and `get2: B → C`, and their composition `Get: A → C` (referred to as the *long get*). The top part of Fig. 3 shows how it works for model A considered above.

Each of the two policies, `policy qt` (green) and `policy tr` (orange), in which person’s disappearance from the view are interpreted, resp., as quitting the company and transitioning to a department not included into the view, is applicable to the new view mappings `get2` and `Get`, thus giving us six lenses shown in Fig. 4 with solid arrows; amongst them, lenses, \mathcal{L}^{qt} and \mathcal{L}^{tr} are obtained by applying *policy pol* to the (long) functor `Get`; and we will refer to them *long lenses*. In addition, we can compose lenses of the same colour as shown in Fig. 4 by dashed arrows (and we can also compose lenses of different colours (ℓ_1^{qt} with ℓ_2^{tr} and ℓ_1^{tr} with ℓ_2^{qt}) but we do not need them). Now an important question is how long and composed lenses are related: whether \mathcal{L}^{pol} and $\ell_1^{\text{pol}}; \ell_2^{\text{pol}}$ for $\text{pol} \in \{\text{qt}, \text{tr}\}$, are equal (perhaps up to some equivalence) or different?

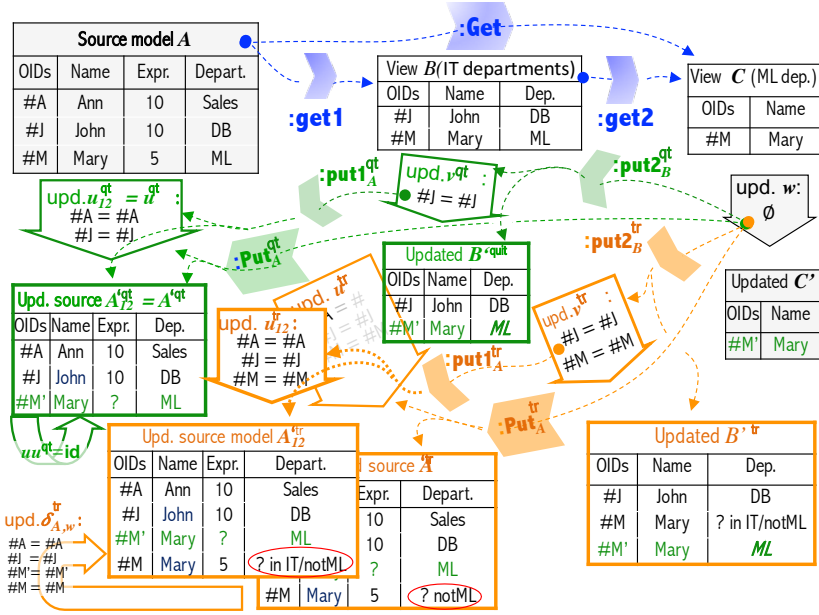


Fig. 3: Example cont'd: functoriality of update policies

Fig. 3 demonstrates how the mechanisms work with a simple example. We begin with an update w of the view C that says that Mary #M left the ML department, and a new Mary #M' was hired for ML. Policy qt interprets Mary's disappearance as quitting the company, and hence this Mary doesn't appear in view B^{qt} produced by $put2^{qt}$ nor in view A_{12}^{qt} produced from B^{qt} by $put1^{qt}$, and updates v^{qt} and u_{12}^{qt} are written accordingly. Obviously, Mary also does not appear in view A^{qt} produced by the long lens's Put^{qt} . Thus, $put1^{qt}(put2^{qt}(w)) = Put_A^{qt}(w)$, and it is easy to understand that such equality will hold for any source model A and any update $w: C \rightarrow C'$ due to the nature of our two views $get1$ and $get2$. Hence, $\mathcal{L}^{qt} = \ell_1^{qt}; \ell_2^{qt}$ where $\mathcal{L}^{qt} = (Get, Put^{qt})$ and $\ell_i^{qt} = (get_i, put_i^{qt})$.

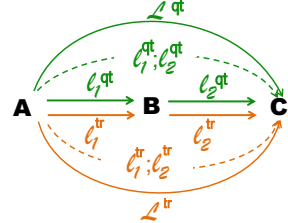


Fig. 4: Lens combination schemas for Fig. 3

The situation with policy tr is more interesting. Model A_{12}^{tr} produced by the composed lens $\ell_1^{tr}; \ell_2^{tr}$, and model A^{tr} produced by the long lens $\mathcal{L}^{tr} = (Get, Put^{tr})$ are different as shown in the figure (notice the two different values for Mary's department framed with red ovals in the models). Indeed, the composed lens has more information about the old employee Mary—it knows that Mary was in the IT view, and hence can propagate the update more accurately. The comparison update $\delta_{A,w}^{tr}: A^{tr} \rightarrow A_{12}^{tr}$ adds this missing information so that equality $u^{tr}; \delta_{A,w}^{tr} = u_{12}^{tr}$ holds. This is a general phenomenon: functor composition loses information and, in general, functor $Get = get1; get2$ knows less than the pair $(get1, get2)$. Hence, operation Put back-propagating updates over Get (we will

also say *inverting* Get) will, in general, result in less certain models than composition $\text{put1} \circ \text{put2}$ that inverts the composition $\text{get1}; \text{get2}$ (a discussion and examples of this phenomenon in the context of vertical composition of updates can be found in [8]). Hence, comparison updates such as $\delta_{A,w}^{\text{tr}}$ should exist for any A and any $w: A.\text{Get} \rightarrow C'$, and together they should give rise to something like a natural transformation between lenses, $\delta_{\mathbf{A},\mathbf{B},\mathbf{C}}^{\text{tr}}: \mathcal{L}^{\text{tr}} \Rightarrow \ell_1^{\text{tr}}; \ell_2^{\text{tr}}$. To make this notion precise, we need a notion of natural transformation between “functors” put , which we leave for future work. In the present paper, we will consider policies like qt , for which strict equality holds.

4.2 Sequential composition of ala-lenses

Let $\kappa: \mathbf{A} \rightarrow \mathbf{B}$ and $\ell: \mathbf{B} \rightarrow \mathbf{C}$ be two ala-lenses with parameterized functors $\text{get}^\kappa: \mathbf{P} \rightarrow [\mathbf{A}, \mathbf{B}]$ and $\text{get}^\ell: \mathbf{Q} \rightarrow [\mathbf{B}, \mathbf{C}]$ resp. Their *composition* is the following ala-lens $\kappa; \ell$. Its parameter space is the product $\mathbf{P} \times \mathbf{Q}$, and the get -family is defined as follows. For any pair of parameters (p, q) (we will write pq), $\text{get}_{pq}^{\kappa; \ell} = \text{get}_p^\kappa; \text{get}_q^\ell: \mathbf{A} \rightarrow \mathbf{C}$. Given a pair of parameter deltas, $e: p \rightarrow p'$ in \mathbf{P} and $h: q \rightarrow q'$ in \mathbf{Q} , their $\text{get}^{\kappa; \ell}$ -image is the Godement product $*$ of natural transformations, $\text{get}^{\kappa; \ell}(eh) = \text{get}^\kappa(e) * \text{get}^\ell(h)$ (we will also write $\text{get}_e^\kappa \parallel \text{get}_h^\ell$)

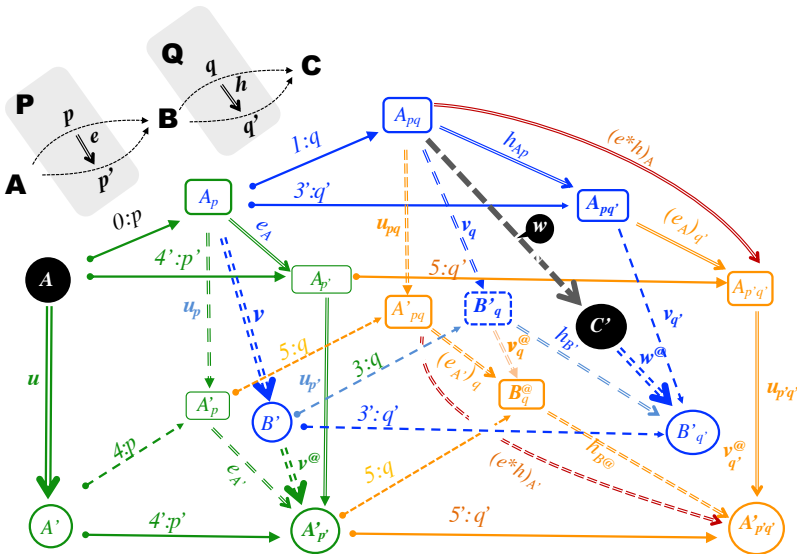


Fig. 5: Sequential composition of apa-lenses

Now we define $\kappa; \ell$'s propagation operations puts . Let (A, pq, A_{pq}) with $A \in \mathbf{A}_0$, $pq \in (\mathbf{P} \times \mathbf{Q})_0$, $A.\text{get}_p^\kappa.\text{get}_q^\ell = A_{pq} \in \mathbf{C}_0$ be a state of lens $\kappa; \ell$, and $w: A_{pq} \rightarrow C'$ is a target update as shown in Fig. 3. For the first propagation step, we run lens ℓ as shown in Fig. 3 with the blue colour for derived elements: this is just an

instantiation of the pattern of Fig. 2 with the source object being $A_p = A.\text{get}_p$ and parameter q . The results are deltas

$$(3) \quad h = \text{put}_{q,A_p}^{\ell.\text{upd}}(w): q \rightarrow q', v = \text{put}_{q,A_p}^{\ell.\text{req}}(w): A_p \rightarrow B', w^{\textcircled{a}} = \text{put}_{q,A_p}^{\ell.\text{self}}(w): C' \rightarrow B'_{q'}.$$

Next we run lens κ at state (p, A) and the target update v produced by lens ℓ ; it is yet another instantiation of pattern in Fig. 2 (this time with the green colour for derived elements), which produces three deltas

$$(4) \quad e = \text{put}_{p,A}^{\kappa.\text{upd}}(v): p \rightarrow p', u = \text{put}_{p,A}^{\kappa.\text{req}}(v): A \rightarrow A', v^{\textcircled{a}} = \text{put}_{p,A}^{\kappa.\text{self}}(v): B' \rightarrow A'_{p'}.$$

These data specify the green prism adjoint to the blue prism: the edge v of the latter is the “first half” of the right back face diagonal $A_p A'_{p'}$ of the former. In order to make an instance of the pattern in Fig. 2 for lens $\kappa; \ell$, we need to extend the blue-green diagram to a triangle prism by filling-in the corresponding “empty space”. These filling-in arrows are provided by functors get^ℓ and get^κ and shown in orange (where we have chosen one of the two equivalent ways of forming the Godeмент product – note two curve brown arrows). In this way we obtain yet another instantiation of the pattern in Fig. 2 denoted by $\kappa; \ell$:

$$(5) \quad \text{put}_{A,pq}^{(\kappa;\ell).\text{upd}}(w) = (e, h), \quad \text{put}_{A,pq}^{(\kappa;\ell).\text{req}}(w) = u, \quad \text{put}_{A,pq}^{(\kappa;\ell).\text{self}}(w) = w^{\textcircled{a}}; v_q^{\textcircled{a}}$$

where $v_q^{\textcircled{a}}$ denotes $v^{\textcircled{a}}.\text{get}_{q'}$. Thus, we built an ala-lens $\kappa; \ell$, which satisfies equation Putget_0 by construction.

Theorem 1 (Sequential composition and lens laws). *Given ala-lenses $\kappa: \mathbf{A} \rightarrow \mathbf{B}$ and $\ell: \mathbf{B} \rightarrow \mathbf{C}$, let lens $\kappa; \ell: \mathbf{A} \rightarrow \mathbf{C}$ be their sequential composition as defined above. Then the lens $\kappa; \ell$ is wb as soon as lenses κ and ℓ are such.*

See [9, Appendix A.3] for a proof.

4.3 Parallel composition of ala-lenses

Let $\ell_i: \mathbf{A}_i \rightarrow \mathbf{B}_i$, $i = 1, 2$ be two ala-lenses with parameter spaces \mathbf{P}_i . The lens $\ell_1 \parallel \ell_2: \mathbf{A}_1 \times \mathbf{A}_2 \rightarrow \mathbf{B}_1 \times \mathbf{B}_2$ is defined as follows. Parameter space $\ell_1 \parallel \ell_2. \mathbf{P} = \mathbf{P}_1 \times \mathbf{P}_2$. For any pair $p_1 \parallel p_2 \in (\mathbf{P}_1 \times \mathbf{P}_2)_0$, define $\text{get}_{p_1 \parallel p_2}^{\ell_1 \parallel \ell_2} = \text{get}_{p_1}^{\ell_1} \times \text{get}_{p_2}^{\ell_2}$ (we denote pairs of parameters by $p_1 \parallel p_2$ rather than $p_1 \otimes p_2$ to shorten long formulas going beyond the page width). Further, for any pair of models $\in (\mathbf{A}_1 \times \mathbf{A}_2)_0$ and deltas $v_1 \parallel v_2: () . \text{get}_{p_1 \parallel p_2}^{\ell_1 \parallel \ell_2} \rightarrow B'_1 \parallel B'_2$, we define componentwise

$$e = \text{put}_{p_1 \parallel p_2}^{(\ell_1 \parallel \ell_2).\text{upd}}(v_1 \parallel v_2): p_1 \parallel p_2 \rightarrow p'_1 \parallel p'_2$$

by setting $e = e_1 \parallel e_2$ where $e_i = \text{put}_{p_i, S_i}^{\ell_i}(v_i)$, $i = 1, 2$ and similarly for $\text{put}_{p_1 \parallel p_2}^{(\ell_1 \parallel \ell_2).\text{req}}$ and $\text{put}_{p_1 \parallel p_2}^{(\ell_1 \parallel \ell_2).\text{self}}$. The following result is obvious.

Theorem 2 (Parallel composition and lens laws). *Lens $\ell_1 \parallel \ell_2$ is wb as soon as lenses ℓ_1 and ℓ_2 are such.*

4.4 Symmetric monoidal structure over ala-lenses

Our goal is to organize ala-lenses into an sm-category. To make sequential composition of ala-lenses associative, we need to consider them up to some equivalence (indeed, Cartesian product is not strictly associative).

Definition 5 (Ala-lens Equivalence) Two parallel ala-lenses $\ell, \hat{\ell}: \mathbf{A} \rightarrow \mathbf{B}$ are called *equivalent* if their parameter spaces are isomorphic via a functor $\iota: \mathbf{P} \rightarrow \hat{\mathbf{P}}$ such that for any $A \in \mathbf{A}_0$, $e: p \rightarrow p' \in \mathbf{P}$ and $v: (A.\text{get}_p) \rightarrow T'$ the following holds (for $\mathbf{x} \in \{\text{req}, \text{self}\}$):

$$A.\text{get}_e = A.\widehat{\text{get}}_{\iota(e)}, \iota(\text{put}_{p,A}^{\text{upd}}(v)) = \widehat{\text{put}}_{\iota(p),A}(v), \text{ and } \text{put}_{p,A}^{\mathbf{x}}(v) = \widehat{\text{put}}_{\iota(p),A}^{\mathbf{x}}(v)$$

Remark 3. It would be more categorical to require delta isomorphisms (i.e., commutative squares whose horizontal edges are isomorphisms) rather than equalities as above. However, model spaces appearing in Bx-practice are skeletal categories (and even stronger than skeletal in the sense that all isos, including iso loops, are identities), for which isos become equalities so that the generality would degenerate into equality anyway.

It is easy to see that operations of lens' sequential and parallel composition are compatible with lens' equivalence and hence are well-defined for equivalence classes. Below we identify lenses with their equivalence classes by default.

Theorem 3 (Ala-lenses form an sm-category). *Operations of sequential and parallel composition of ala-lenses defined above give rise to an sm-category **aLaLens**, whose objects are model spaces (= categories) and arrows are (equivalence classes of) ala-lenses. See [9, p.17 and Appendix A.4] for a proof.*

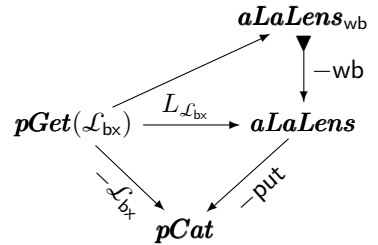
4.5 Functoriality of learning in the delta lens setting

As example in Sect. 4.1 shows, the notion of update policy transcends individual lenses. Hence, its proper formalization needs considering the entire category of ala-lenses and functoriality of a suitable mapping.

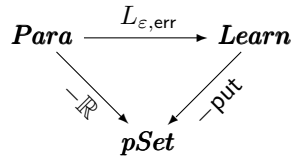
Definition 6 (Bx-transformation language)

A *compositional bidirectional model transformation language* \mathcal{L}_{bx} is given by (i) an sm-category $\mathbf{pGet}(\mathcal{L}_{\text{bx}})$ whose objects are $(\mathcal{L}_{\text{bx}})$ -model spaces and arrows are $(\mathcal{L}_{\text{bx}})$ -transformations which is supplied with forgetful functor into \mathbf{pCat} , and (ii) an sm-functor $L_{\mathcal{L}_{\text{bx}}}: \mathbf{pGet}(\mathcal{L}_{\text{bx}}) \rightarrow \mathbf{aLaLens}$ such that the lower triangle in the inset diagram commutes. (Forgetful functors in this diagram are named “ $-X$ ” with X referring to the structure to be forgotten.)

An \mathcal{L}_{bx} -language is *well-behaved (wb)* if functor $L_{\mathcal{L}_{\text{bx}}}$ factorizes as shown by the upper triangle of the diagram.



Example. A major compositionality result of Fong *et al* [17] states the existence of an sm-functor from the category of Euclidean spaces and parameterized differentiable functions (pd-functions) **Para** into the category **Learn** of learning algorithms (*learners*) as shown by the inset commutative diagram. (The functor is itself parameterized by a *step size* $0 < \varepsilon \in \mathbb{R}$ and an *error function* $\text{err}: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ needed to specify the gradient descent procedure.) However, learners are nothing but codiscrete ala-lenses (see Sect. A.2), and thus the inset diagram is a codiscrete specialization of the diagram in Def. 6 above. That is, the category of Euclidean spaces and pd-functions, and the gradient descent method for back propagation, give rise to a (codiscrete) compositional bx-transformation language (over **pSet** rather than **pCat**).



Finding a specifically Bx instance of Def. 6 (e.g., checking whether it holds for concrete languages and tools such as EMOFLON [23] or GROUNDTRAM [22]) is laborious and left for future work.

5 Related work

Figure 6 on the right is a simplified version of Fig. 8 on p. 194 convenient for our discussion here: immediate related work should be found in areas located at points (0,1) (codiscrete learning lenses) and (1,0) (delta lenses) of the plane. For the point (0,1), the paper [17] by Fong, Spivak and Tuyéras is fundamental: they defined the notion of a codiscrete learning lens (called a learner), proved a fundamental results about sm-functoriality of the gradient descent approach to ML, and thus laid a foundation for the compositional approach to change propagation with learning. One follow-up of that work is paper [16] by Fong and Johnson, in which they build an sm-functor **Learn** \rightarrow **sLens** which maps learners to so called symmetric lenses. That paper is probably the first one where the terms 'lens' and 'learner' are met, but the initial observation that a learner whose parameter set is a singleton is actually a lens is due to Jules Hedges, see [16].

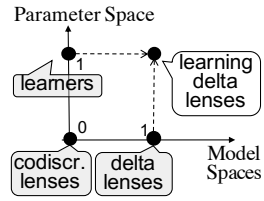


Fig. 6

There are conceptual and technical distinctions between [16] and the present paper. On the conceptual level, by encoding learners as symmetric lenses, they “hide” learning inside the lens framework and make it a technical rather than conceptual idea. In contrast, we consider parameterization and supervised learning as a fundamental idea and a first-class citizen for the lens framework, which grants creation of a new species of lenses. Moreover, while an ordinary lens is a way to invert a functor, a learning lens is a way to invert a parameterized functor so that learning lenses appear as an extension of the parameterization idea from functors to lenses. (This approach can probably be specified formally by treating parameterization as a suitably defined functorial construction.) Besides

technical advantages (working with asymmetric lenses is simpler), our asymmetric model seems more adequate to the problem of learning functions rather than relations. On the technical level, the lens framework we develop in the paper is much more general than in [16]: we categorificated both the parameter space and model spaces, and we work with lenses with amendment (which allows us to relax the Putget law if needed).

As for the delta lens roots (the point (1,0) in the figure), delta lenses were motivated and formally defined in [12] (the asymmetric case) and [13] (the symmetric one). Categorical foundations for the delta lens theory were developed by Johnson and Rosebrugh in a series of papers (see [20] for references); this line is continued in Clarke’s work [6]. The notion of a delta lens with amendments (in both asymmetric and symmetric variants) was defined in [11], and several composition results were proved. Another extensive body of work within the delta-based area is modelling and implementing model transformations with triple-graph grammars (TGG) [4,23]. TGG provide an implementation framework for delta lenses as is shown and discussed in [5,19,2], and thus inevitably consider change propagation on a much more concrete level than lenses. The author is not aware of any work considering functoriality of update policies developed within the TGG framework.

The present paper is probably the first one at the intersection (1,1) of the plane. The preliminary results have recently been reported at ACT’19 in Oxford to a representative lens community, and no references besides [17], [16] mentioned above were provided.

6 Conclusion

The perspective on Bx presented in the paper is an example of a fruitful interaction between two domains—ML and Bx. In order to be ported to Bx, the compositional approach to ML developed in [17] is to be categorificated as shown in Fig. 8 on p. 194. This opens a whole new program for Bx: checking that currently existing Bx languages and tools are compositional (and well-behaved) in the sense of Def. 6 p. 190. The wb compositionality is an important practical requirement as it allows for modular design and testing of bidirectional transformations. Surprisingly, but this important requirement has been missing from the agenda of the Bx community, e.g., the recent endeavour of developing an effective benchmark for Bx-tools [3] does not discuss it.

In a wider context, the main message of the paper is that the learning idea transcends its applications in ML: it is applicable and usable in many domains in which lenses are applicable such as model transformations, data migration, and open games [18]. Moreover, the categorificated learning may perhaps find useful applications in ML itself. In the current ML setting, the object to be learnt is a function $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ that, in the OO class modelling perspective, is a very simple structure: it can be seen as one object with a (huge) amount of attributes, or, perhaps, a predefined set of objects, which is not allowed to be changed during the search — only attribute values may be changed. In the delta lens view,

such changes constitute a rather narrow class of updates and thus unjustifiably narrow the search space. Learning with the possibility to change dimensions m, n may be an appropriate option in several contexts. On the other hand, while categorification of model spaces extends the search space, categorification of the parameter space would narrow the search space as we are allowed to replace a parameter p by parameter p' only if there is a suitable arrow $e: p \rightarrow p'$ in category \mathbf{P} . This narrowing may, perhaps, improve performance. All in all, the interaction between ML and Bx could be bidirectional!

A Appendices

A.1 Category of parameterized functors \mathbf{pCat}

Category \mathbf{pCat} has all small categories as objects. \mathbf{pCat} -arrows $\mathbf{A} \rightarrow \mathbf{B}$ are *parameterized functors* (p -functors) i.e., functors $f: \mathbf{P} \rightarrow [\mathbf{A}, \mathbf{B}]$ with \mathbf{P} a small category of *parameters* and $[\mathbf{A}, \mathbf{B}]$ the category of functors from \mathbf{A} to \mathbf{B} and their natural transformations. For an object p and an arrow $e: p \rightarrow p'$ in \mathbf{P} , we write f_p for the functor $f(p): \mathbf{A} \rightarrow \mathbf{B}$ and f_e for the natural transformation $f(e): f_p \Rightarrow f_{p'}$. We will write p-functors as labelled arrows $f: \mathbf{A} \xrightarrow{\mathbf{P}} \mathbf{B}$. As \mathbf{Cat} is Cartesian closed, we have a natural isomorphism between $\mathbf{Cat}(\mathbf{P}, [\mathbf{A}, \mathbf{B}])$ and $\mathbf{Cat}(\mathbf{P} \times \mathbf{A}, \mathbf{B})$ and can reformulate the above definition in an equivalent way with functors $\mathbf{P} \times \mathbf{A} \rightarrow \mathbf{B}$. We prefer the former formulation as it corresponds to the notation $f: \mathbf{A} \xrightarrow{\mathbf{P}} \mathbf{B}$ visualizing \mathbf{P} as a hidden state of the transformation, which seems adequate to the intuition of parameterized in our context. (If some technicalities may perhaps be easier to see with the product formulation, we will switch to the product view thus doing currying and uncurrying without special mentioning.) Sequential composition of $f: \mathbf{A} \xrightarrow{\mathbf{P}} \mathbf{B}$ and $g: \mathbf{B} \xrightarrow{\mathbf{Q}} \mathbf{C}$ is $f.g: \mathbf{A} \xrightarrow{\mathbf{P} \times \mathbf{Q}} \mathbf{C}$ given by $(f.g)_{pq} \stackrel{\text{def}}{=} f_p.g_q$ for objects, i.e., pairs $p \in \mathbf{P}, q \in \mathbf{Q}$, and by the Godement product of natural transformations for arrows in $\mathbf{P} \times \mathbf{Q}$. That is, given a pair $e: p \rightarrow p'$ in \mathbf{P} and $h: q \rightarrow q'$ in \mathbf{Q} , we define the transformation $(f.g)_{eh}: f_p.g_q \Rightarrow f_{p'}.g_{q'}$ to be the Godement product $f_e * g_h$.

Any category \mathbf{A} gives rise to a p-functor $\text{Id}_{\mathbf{A}}: \mathbf{A} \xrightarrow{\mathbf{1}} \mathbf{A}$, whose parameter space is a singleton category $\mathbf{1}$ with the only object $*$, $\text{Id}_{\mathbf{A}}(*) = \text{id}_{\mathbf{A}}$ and $\text{Id}_{\mathbf{A}}(\text{id}_*): \text{Id}_{\mathbf{A}} \Rightarrow \text{id}_{\mathbf{A}}$ is the identity transformation. It's easy to see that p-functors $\text{Id}_{\mathbf{A}}$ are units of the sequential composition. To ensure associativity we need to consider p-functors up to an equivalence of their parameter spaces. Two parallel p-functors $f: \mathbf{A} \xrightarrow{\mathbf{P}} \mathbf{B}$ and $\hat{f}: \mathbf{A} \xrightarrow{\hat{\mathbf{P}}} \mathbf{B}$, are *equivalent* if there is an isomorphism $\alpha: \mathbf{P} \rightarrow \hat{\mathbf{P}}$ such that two parallel functors $f: \mathbf{P} \rightarrow [\mathbf{A}, \mathbf{B}]$ and $\alpha; \hat{f}: \mathbf{P} \rightarrow [\mathbf{A}, \mathbf{B}]$ are naturally isomorphic; then we write $f \approx_{\alpha} \hat{f}$. It's easy to see that if $f \approx_{\alpha} \hat{f}: \mathbf{A} \rightarrow \mathbf{B}$ and $g \approx_{\beta} \hat{g}: \mathbf{B} \rightarrow \mathbf{C}$, then $f; g \approx_{\alpha \times \beta} \hat{f}; \hat{g}: \mathbf{A} \rightarrow \mathbf{C}$, i.e., sequential composition is stable under equivalence. Below we will identify p-functors and their equivalence classes. Using a natural isomorphism $(\mathbf{P} \times \mathbf{Q}) \times \mathbf{R} \cong \mathbf{P} \times (\mathbf{Q} \times \mathbf{R})$, strict associativity of the functor composition and strict associativity of the Godement product, we conclude that

sequential composition of (equivalence classes of) p-functors is strictly associative. Hence, \mathbf{pCat} is a category.

Our next goal is to supply it with a monoidal structure. We borrow the latter from the sm-category (\mathbf{Cat}, \times) , whose tensor is given by the product. There is an identical on objects embedding $(\mathbf{Cat}, \times) \hookrightarrow \mathbf{pCat}$ that maps a functor $f: \mathbf{A} \rightarrow \mathbf{B}$ to a p-functor $\bar{f}: \mathbf{A} \xrightarrow{1} \mathbf{B}$ whose parameter space is the singleton category $\mathbf{1}$. Moreover, as this embedding is a functor, the coherence equations for the associators and unitors that hold in (\mathbf{Cat}, \times) hold in \mathbf{pCat} as well (this proof idea is borrowed from [17]). In this way, \mathbf{pCat} becomes an sm-category. In a similar way, we define the sm-category \mathbf{pSet} of small sets and parametrized functions between them — the codiscrete version of \mathbf{pCat} . The diagram in Fig. 7 shows how these categories are related.

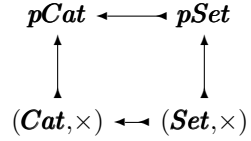


Fig. 7

A.2 Ala-lenses as categorification of ML-learners

Figure 8 shows a discrete two-dimensional plane with each axis having three points: a space is a singleton, a set, a category encoded by coordinates 0,1,2 resp. Each of the points x_{ij} is then the location of a corresponding sm-category of

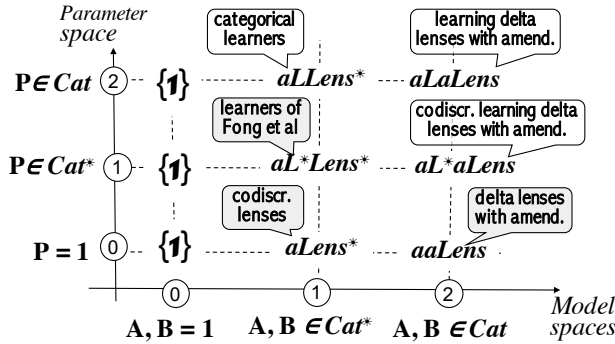


Fig. 8: The universe of categories of learning delta lenses

(asymmetric) learning (delta) lenses. Category $\{\mathbf{1}\}$ is a terminal category whose only arrow is the identity lens $\mathbf{1} = (\text{id}_1, \text{id}_1): \mathbf{1} \rightarrow \mathbf{1}$ propagating from a terminal category $\mathbf{1}$ to itself. Label \ast refers to the codiscrete specialization of the construct being labelled: L^\ast means codiscrete learning (i.e., the parameter space \mathbf{P} is a set considered as a codiscrete category) and $aLens^\ast$ refers to codiscrete model spaces. The category of learners defined in [17] is located at point (1,1), and the category of learning delta lenses with amendments defined in the present paper is located at (2,2). There are also two semi-categorificated species of learning lenses: categorical learners at point (1,2) and codiscretely learning delta lenses at (2,1), which are special cases of ala-lenses.

References

1. Abiteboul, S., McHugh, J., Rys, M., Vassalos, V., J.Wiener: Incremental Maintenance for Materialized Views over Semistructured Data. In: Gupta, A., Shmueli, O., Widom, J. (eds.) VLDB. Morgan Kaufmann (1998)
2. Anjorin, A.: An introduction to triple graph grammars as an implementation of the delta-lens framework. In: Gibbons, J., Stevens, P. (eds.) Bidirectional Transformations - International Summer School, Oxford, UK, July 25-29, 2016, Tutorial Lectures. Lecture Notes in Computer Science, vol. 9715, pp. 29–72. Springer (2016). <https://doi.org/10.1007/978-3-319-79108-1>
3. Anjorin, A., Diskin, Z., Jouault, F., Ko, H., Leblebici, E., Westfechtel, B.: Benchmarkx reloaded: A practical benchmark framework for bidirectional transformations. In: Eramo and Johnson [15], pp. 15–30, <http://ceur-ws.org/Vol-1827/paper6.pdf>
4. Anjorin, A., Leblebici, E., Schürr, A.: 20 years of triple graph grammars: A roadmap for future research. ECEASST **73** (2015). <https://doi.org/10.14279/tuj.eceasst.73.1031>
5. Anjorin, A., Rose, S., Deckwerth, F., Schürr, A.: Efficient model synchronization with view triple graph grammars. In: Modelling Foundations and Applications - 10th European Conference, ECMFA 2014, York, UK, July 21-25, 2014. Proceedings. Lecture Notes in Computer Science, vol. 8569, pp. 1–17. Springer (2014). https://doi.org/10.1007/978-3-319-09195-2_1
6. Clarke, B.: Internal lenses as functors and cofunctors. In: Pre-proceedings of ACT'19, Oxford, 2019. <http://www.cs.ox.ac.uk/ACT2019/preproceedings/BryceClarke.pdf>
7. Czarnecki, K., Foster, J.N., Hu, Z., Lämmel, R., Schürr, A., Terwilliger, J.F.: Bidirectional transformations: A cross-discipline perspective. In: Theory and Practice of Model Transformations, pp. 260–283. Springer (2009)
8. Diskin, Z.: Compositionality of update propagation: Lax putput. In: Eramo and Johnson [15], pp. 74–89, <http://ceur-ws.org/Vol-1827/paper12.pdf>
9. Diskin, Z.: General supervised learning as change propagation with delta lenses. CoRR **abs/1911.12904** (2019), <http://arxiv.org/abs/1911.12904>
10. Diskin, Z., Gholizadeh, H., Wider, A., Czarnecki, K.: A three-dimensional taxonomy for bidirectional model synchronization. Journal of System and Software **111**, 298–322 (2016). <https://doi.org/10.1016/j.jss.2015.06.003>
11. Diskin, Z., König, H., Lawford, M.: Multiple model synchronization with multiary delta lenses with amendment and K-Putput. Formal Asp. Comput. **31**(5), 611–640 (2019). <https://doi.org/10.1007/s00165-019-00493-0>, (Sect.7.1 of the paper is unreadable and can be found in <http://arxiv.org/abs/1911.11302>)
12. Diskin, Z., Xiong, Y., Czarnecki, K.: From State- to Delta-Based Bidirectional Model Transformations: the Asymmetric Case. Journal of Object Technology **10**, 6: 1–25 (2011)
13. Diskin, Z., Xiong, Y., Czarnecki, K., Ehrig, H., Hermann, F., Orejas, F.: From state-to delta-based bidirectional model transformations: the symmetric case. In: MODELS, pp. 304–318. Springer (2011)
14. El-Sayed, M., Rundensteiner, E.A., Mani, M.: Incremental Maintenance of Materialized XQuery Views. In: Liu, L., Reuter, A., Whang, K.Y., Zhang, J. (eds.) ICDE. p. 129. IEEE Computer Society (2006). <https://doi.org/10.1109/ICDE.2006.80>
15. Eramo, R., Johnson, M. (eds.): Proceedings of the 6th International Workshop on Bidirectional Transformations co-located with The European Joint Conferences on Theory and Practice of Software, Bx@ETAPS 2017, Uppsala, Sweden, April

- 29, 2017, CEUR Workshop Proceedings, vol. 1827. CEUR-WS.org (2017), <http://ceur-ws.org/Vol-1827>
16. Fong, B., Johnson, M.: Lenses and learners. In: Cheney, J., Ko, H. (eds.) Proceedings of the 8th International Workshop on Bidirectional Transformations co-located with the Philadelphia Logic Week, Bx@PLW 2019, Philadelphia, PA, USA, June 4, 2019. CEUR Workshop Proceedings, vol. 2355, pp. 16–29. CEUR-WS.org (2019), <http://ceur-ws.org/Vol-2355/paper2.pdf>
 17. Fong, B., Spivak, D.I., Tuyéras, R.: Backprop as functor: A compositional perspective on supervised learning. In: The 34th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2019, Vancouver, BC, Canada, June 24–27, 2019. pp. 1–13. IEEE (2019). <https://doi.org/10.1109/LICS.2019.8785665>
 18. Hedges, J.: From open learners to open games. CoRR **abs/1902.08666** (2019), <http://arxiv.org/abs/1902.08666>
 19. Hermann, F., Ehrig, H., Orejas, F., Czarnecki, K., Diskin, Z., Xiong, Y., Gottmann, S., Engel, T.: Model synchronization based on triple graph grammars: correctness, completeness and invertibility. *Software and System Modeling* **14**(1), 241–269 (2015). <https://doi.org/10.1007/s10270-012-0309-1>
 20. Johnson, M., Rosebrugh, R.D.: Unifying set-based, delta-based and edit-based lenses. In: The 5th International Workshop on Bidirectional Transformations, Bx 2016. pp. 1–13 (2016), http://ceur-ws.org/Vol-1571/paper_13.pdf
 21. Kappel, G., Langer, P., Retschitzegger, W., Schwinger, W., Wimmer, M.: Model transformation by-example: A survey of the first wave. In: *Conceptual Modelling and Its Theoretical Foundations - Essays Dedicated to Bernhard Thalheim on the Occasion of His 60th Birthday*. pp. 197–215 (2012). https://doi.org/10.1007/978-3-642-28279-9_15
 22. Sasano, I., Hu, Z., Hidaka, S., Inaba, K., Kato, H., Nakano, K.: Toward bidirectionalization of ATL with GRoundTram. In: *Theory and Practice of Model Transformations - 4th International Conference, ICMT 2011, Zurich, Switzerland, June 27–28, 2011*. Proceedings. Lecture Notes in Computer Science, vol. 6707, pp. 138–151. Springer (2011). https://doi.org/10.1007/978-3-642-21732-6_10
 23. Weidmann, N., Anjorin, A., Fritsche, L., Varró, G., Schürr, A., Leblebici, E.: Incremental bidirectional model transformation with emoflon: Ibex. In: The 8th International Workshop on Bidirectional Transformations co-located with the Philadelphia Logic Week, Bx@PLW 2019, Philadelphia, PA, USA, June 4, 2019. CEUR Workshop Proceedings, vol. 2355, pp. 45–55. CEUR-WS.org (2019), <http://ceur-ws.org/Vol-2355/paper4.pdf>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

