An Note on a Categorical Semantics for ER-Models

Ernst-Erich Doberkat
Chair for Software-Technology
University of Dortmund
doberkat@acm.org

November 18, 1999
Abstract

We have a look at the semantics of Entity-Relationship models, a popular device for modelling data, but lacking a stringent semantics. In an earlier paper we have shown how to generate an algebraic specification for an extended ER-model; in this paper we extend the algebraic view of a model through a categorial interpretation.

Inheritance induces a tree structure for an ER-model. This is decorated with objects from a suitable category, and we ask for a unifying view of this model. Our approach suggests using colimits as the semantics. It is shown that under very mild conditions the colimit exists, and that this colimit is an adjoint to the diagonal. Finally we show how to integrate binary relations into this approach by studying two general conditions on the morphism associated to a relation.
1 Introduction

Entity-Relationship modelling is a rather popular approach to data modelling, as witnessed by the literature on data base systems [Ull88] and on software engineering [GJM91]. This popularity is partly due to the visual representability of the models constructed, and to the relative ease with which these models may be implemented. Furthermore, ER models come in many different variants, hence different kinds of problems may be subjected to ER modeling, indicating the wide applicability of this method, nay, of this family of methods.  
This kind of data modeling has, however, quite a notable drawback from the conceptual point of view. Its modeling facilities are almost seductive, but the formal foundations for its semantics are not that readily described. In fact, a set theoretic semantics is easily conceived, but it is not difficult to see that set theory due to its implied semantics is not always appropriate. An alternative for giving some meaning to an ER model is provided through an algebraic way of life: [CLWW94], [Het93] and [GH91, Hoh93] formulate an algebraic view, and [Dob97] gives methods for generating an algebraic specification for an ER model.
These approaches are mainly characterized by providing some axiomatic description for the operations involved in a suitable variant of the model, and by establishing subsequently the models provided by these descriptions as the (class of) semantics for the ER model. This is a mathematically sound way providing a solid foundation for this powerful method. It enables constructing concise models when ER modeling is combined with other approaches that focus on modelling the functionality of an application. In [Dob97, Sec. 5] it is shown with an example how data modeling with ER may be combined with functional modeling using a Petri nets, essentially attributing the edges of the net with enabling terms from an algebraic specification.
This note carries the algebraic approach developed in [Dob97] a bit further. Given an algebraic specification, we may consider models for it. These models form a category, hence the specification determines objects of a category (which is usually unrelated to the category of sets). Since each entity and each relation in determined by a specification of its own, we see a collection of models, hence a collection of objects in our model category. In addition, some entities are related to each other (e.g., by ISA links). This translates to the edges of a graph, each node of which is decorated with an algebraic specification. Translating the edges to morphisms between the corresponding objects, hence one ends up with a functor from a diagram to the category under consideration.
It is this functor which is of interest here.

Carried away? It is easy to get carried away by the algebraic arguments, so the reason for why it is helpful to have such an algebraic discussion should be elaborated. First, categories abstract away all the implementation details and permit focussing on the structure of the objects under consideration and their relationships. In fact, the details of an object in a category is not accessible at all. An object may be internally as rich as Croesus (or Bill Gates) — this is of no concern, as long as this richness is not expressible on the outside. This, in turn, is determined by the set of morphisms associated with the object. Hence the approaches (ER, categories) are in fact rather similar by focussing on external relations. They may be oriented towards preserving structures (like homomorphisms familiar from group theory, or like signature morphisms from algebraic specifications), but they may also be considered as models for channels (like in the COMMUNITY approach, see [WF98]). This indicates that a categorical approach may encompass also many other views to modelling. The full generality
has, however, its price: further work is required for modelling these aspects which are not
determined by the structure alone, in particular questions on capturing attributes and their
dependence. This will require a judicious restriction to the kind of category that is used for
formulating the model.

**Organization** This paper is organized as follows: we first recall from [Dob97] the necessary
properties of ER models and focus on the forest induced by the model. Then we have a look
at the semantics of a diagram. It is formulated in terms of colimits, and we investigate
the question under which conditions such a colimit exists. It turns out that an answer for
the special case of trees is rather immediate, and that constructing an adjunction helps in
establishing the general model for forests (supporting once more MacLane’s *adjoints-are-
everywhere*-hypothesis put forward in [Mac98, p. 97]). This result implies that the semantics
is compositional.
The constructions address entities only, and we discuss in the final section modelling the
interplay between relations and entities.

**Further work** Attributes are not considered in this note, they are rather a subject of
further work. The solution we present is completely general as far as entities are concerned,
but could probably be fine tuned when it comes to different kinds of relations. This is also
delegated to the drawer labelled *Further Work*.

## 2 ER-Models

An entity-relationship model [ULL88, 2.4] consists of entities, relationships on these entities
and attributes both on entities and relations. Only binary relations will be considered for
the sake of simplicity. Entities may be related by the lSA relation: $E_1 \isa E_2$ indicates that
each instance of $E_1$ is also an instance of $E_2$, hence shares all the attributes defined on the
latter entity. Multiple inheritance is not permitted here (i.e., no entity may be related to
more than one other entity via an lSA-relation). If entities are represented by their *extension*,
then relations are subsets of the Cartesian product. If $R$ relates the entities $E_1$ and $E_2$, the
entities in $E_1$ (in $E_2$) are said to be in the domain (in the co-domain) of $R$. In the graphical
representation the order of the factors for the product is not immediate, hence we number the
corners of the diamond counterclockwise starting in the northern corner, identifying domain
and co-domain uniquely. Attributes are mentioned for the sake of completeness; they are
usually represented as maps; as usual, an attribute is a *key* for an entity iff it uniquely
determines each instance. A relation $R$ is $N:1$ iff $b_1 = b_2$ is true whenever both $aRb_1$ and
$aRb_2$ hold (i.e., whenever $R$ is a partial map), i.e. iff for each instance $a$ in the domain of $R$ the
set $\{b : aRb\}$ contains at most one element. In a similar way $1:N$ relations are characterized:
$R$ is an $1:N$ relation iff its inverse $R^{-1}$ is $N:1$. A relation is said to be $N:M$ iff there
are no restrictions concerning the domain or the co-domain of the pairs participating in the
relation. That a relation is $N:1$ is indicated in the graphical representation by labelling the
dge leading to the domain with an *, and a 1 as a label for the co-domain.

Fig. 1 displays an example for modelling a simple graphical user interface.
The entities are *window*, *button*, *textfield*, *menu entry*, moreover *trigger* and *text (fixed)*,
both of which are related to *menu entry* via the lSA relation, and output *window* and *icon*,

---

*An Note on a Categorical Semantics for ER-Models*
for which isA window holds. The relations are sequence, which is an $N : M$ relation between windows, residesIn, a $1 : N$ relation between window and button, contains relates textfield and window as an $N : 1$ relation, inMenu is an $N : 1$ relation between trigger and window, and finally invocation relates trigger and menu entry $1 : N$. Attributes are e.g. window layout defined on entity window or button position defined on relation residesIn. As usual, key attributes are underlined, and total relations or attributes carry a dot where they are total. We will, however, not deal with attributes here.

3 Preparations

This section will relate the construction of a graph from an ER-model (see [Dob97, Sec. 3]). Furthermore it will collect some notations and results from category theory following [Mac98] for easier reference and the reader’s convenience.

3.1 The Graph

Given an ER-model $\mathcal{M}$, denote by $\mathcal{E}$ and $\mathcal{R}$ the respective sets of entities and of relations. Let $N_\mathcal{E}$ and $N_\mathcal{R}$ be fresh and disjoint sets of nodes representing $\mathcal{E}$ and the domains and co-domains for the relations in $\mathcal{M}$, so that each $E \in \mathcal{E}$ is associated with a unique node $n_E \in N_\mathcal{E}$, similarly for $\mathcal{R}$. Construct a directed edge $n_{E_1} \to n_{E_2}$ iff $E_1 \text{ isA } E_2$ holds in $\mathcal{M}$.

If $r \in \mathcal{R}$ is a relation with $E_1$ as domain and $E_2$ as co-domain, generate two fresh nodes $j^{\delta(r)}$ and $j^{\gamma(r)}$ in $N_\mathcal{E}$ which are linked through the directed edges $j^{\delta(r)} \to n_{E_1}$ and $j^{\gamma(r)} \to n_{E_2}$ to their domain and co-domain, resp. (this reflects the fact that the domain and the co-domain of $r$ have to be taken care of when it comes to manipulate the relation). The construction from [Dob97] additionally constructs non-directed edges $n_r \leftrightarrow n_{E_1}$ and $n_r \leftrightarrow n_{E_2}$; it also takes care of the attributes. But the simpler construction from above will suffice for the purposes of the present paper.

To illustrate things, we borrow from [Dob97] a simple ER-model for constructing a graphical user interface. This is displayed in Fig. 1. The directed graph generated from it is shown in Fig. 2.

3.2 Morphisms and all that

If $\mathcal{C}$ is a category, $\mathcal{C}(a,b)$ is the set of all morphisms $a \to b$ in $\mathcal{C}$. Suppose $\mathcal{D}$ is another category, and $S : \mathcal{C} \to \mathcal{D}$ is a functor, then $\langle v, r \rangle$ is called an arrow from $x$ to $S$ iff $r : x \to Sv$ is a morphism in $\mathcal{D}$, or, equivalently, if $\langle v, r \rangle$ is a member of the comma category $x \downarrow S$. An initial object in this category is called universal; thus universal objects are unique up to isomorphisms. Thus $\langle v, r \rangle$ is a universal arrow from $x$ to $S$ iff for each arrow $\langle v', r' \rangle$ from $x$ to $S$ there exists a unique morphism $f : v \to v'$ such that $r' = Sf \circ r$ holds, thus the following diagram is commutative:

$$
\begin{aligned}
\text{Diagram:}
\end{aligned}
$$
Recall that a natural transformation \( \tau : R \Rightarrow S \) for the functors \( R, S : \mathcal{C} \to \mathcal{D} \) assigns to each object \( c \) in \( \mathcal{C} \) an arrow \( \tau_c : Rc \to Sc \) in \( \mathcal{D} \) such that for each morphism \( f : c \to c' \) the diagram

\[
\begin{array}{ccc}
Rc & \xrightarrow{\tau_c} & Sc \\
\downarrow Rf & & \downarrow Sf \\
Rc' & \xrightarrow{\tau_{c'}} & Sc'
\end{array}
\]

is commutative. Denote for two natural transformations \( \tau : R \Rightarrow S \) and \( \sigma : S \Rightarrow T \) their \textit{vertical composition} by \( \sigma \circ \tau \), hence

\[
(\sigma \circ \tau)_c = \sigma_c \circ \tau_c
\]

holds for each object \( c \) in \( \mathcal{C} \). The vertical composition is again a natural transformation.

Trees are directed towards their root, a \textit{forest} is a finite collection of finite trees. Each forest \( \mathcal{B} \) may be considered as a category and each map \( S : \mathcal{B} \to \mathcal{C} \) that assigns nodes to objects and edges to morphisms (so that each edge \( e : i \to j \) in \( \mathcal{B} \) yields a morphism \( Se \in \mathcal{C}(c_i, c_j) \)) may be considered as a functor; we can define a unique morphism \( Sp \in \mathcal{C}(c_{i_0}, c_{i_k}) \) for each path \( p \) from \( i_0 \) to \( i_k \) in \( \mathcal{B} \) upon piecing \( Sp \) together from these edges. This is an old trick discussed at length e.g. in [Mil65, II.1], see also [Mac98, II.7].

The category \( \mathcal{C}^{\mathcal{B}} \) is the category of all functors from \( \mathcal{B} \) to \( \mathcal{C} \), natural transformations serving as usual as morphisms between functors. These functors are called \( \mathcal{B} \)-\textit{diagrams over} \( \mathcal{C} \), or
Figure 2: The graph derived from the ER-model

simply diagrams, if the context is clear. Denote by $\Delta_B$ the diagonal functor from $\mathcal{C}$ to $\mathcal{CB}$, hence $\Delta_B(v)$ is a constant functor on $\mathcal{B}$, mapping each object to $v$, and each morphism to the identity $1_v$. The pair $\langle v, r \rangle$ is called a colimit for the diagram $S$ iff it is a universal arrow from $S$ to $\Delta_B$. Thus for each node $j$ in $\mathcal{B}$, $r_j : S_j \rightarrow v$ is a morphism in $\mathcal{C}$ such that for each edge $e : i \rightarrow j$ the diagram

\[
\begin{array}{ccc}
S_i & \xrightarrow{Se} & S_j \\
\downarrow{r_i} & & \downarrow{r_j} \\
v & & v
\end{array}
\]

commutes, and if we have another arrow $\langle v', r' \rangle$ from $S$ to $\Delta_B$, hence a commutative diagram for $\langle r'_j : S_j \rightarrow v' \rangle$, then

$\text{r}'_j = f \circ r_j$

holds for each node $j$, where $f : v \rightarrow v'$ is a uniquely determined morphism in $\mathcal{C}$.

The coproduct $\langle v, r \rangle$ of an object $a$ in $\mathcal{C}^n$ is a universal arrow from $a$ to $\Delta_n$, $v$ being denoted...
by $\prod_{i=1,\ldots,n} a_i$, and $r$ being identified by a collection

$$r_j : a_j \rightarrow \prod_{i=1,\ldots,n} a_i$$

of injections. The coproduct is sometimes only identified by its object, similarly for the colimit.

Suppose that $R : \mathcal{C} \rightarrow \mathcal{D}$ and $S : \mathcal{D} \rightarrow \mathcal{C}$ are functors, and that for each pair of objects $c$ in $\mathcal{C}$ and $d$ in $\mathcal{D}$ there exists a bijection

$$\varphi_{cd} : \mathcal{D}(Rc, d) \rightarrow \mathcal{C}(c, Sd)$$

that is natural in $c$ and $d$. Then $\langle R, S, \varphi \rangle : \mathcal{C} \rightarrow \mathcal{D}$ is called an adjunction, and $R$ is called the left adjoint for $S$.

4 Basic Constructions

Consider the graph $\mathcal{B}$ associated with the entities of an ER-model. This graph is actually a forest of trees, since only the nodes coming from entities are considered, reflecting the fact that we do not permit multiple inheritance, hence each entity inherits from at most one other entity. Suppose that we have constructed for each node an algebraic specification, as outlined at length in [Dob97, 4.2]. Interpreting the specification, and assuming that the specification is valid, this yields for each node a model living in that node.

Formally, we map each node $j$ to a model $S_j$. These models are linked by an edge, whenever the corresponding nodes are related by the IsA-relation, hence there is a directed edge

$$e : j_1 \rightarrow j_2$$

between two nodes $j_1$ and $j_2$ iff $j_2$ inherits from $j_1$. Reflecting this in the world of models (no, not Claudia Schiffer’s), we establish a homomorphism

$$Se : Sj_1 \rightarrow Sj_2.$$ 

The models for a given specification form a category $\mathcal{C}$, so we end up with a map $S$ associating each node in $\mathcal{B}$ an object from $\mathcal{C}$, cf. 3.2.

Now the project investigating the semantics of an ER-model may be formulated more specific: Given the functor above, associate an object — provisionally called $s$ — and morphisms $\alpha$ with the diagram such that

- there is an arrow $\alpha_j : Sj \rightarrow s$ for each node in the specification tree (indicating that each model for a specification may be associated with the object through a morphism),

- the arrows are compatible, hence if $e : j_1 \rightarrow j_2$ is an edge in $\mathcal{B}$, then $\alpha_{j_1} = \alpha_{j_2} \circ Se$, hence making the diagram

$$\begin{array}{ccc}
Sj_1 & \xrightarrow{Se} & Sj_2 \\
\alpha_{j_1} \downarrow & & \downarrow \alpha_{j_2} \\
s & & s
\end{array}$$

commutative,
• the object $s$ should be as close as possible to the objects described by $S$: thus if there is another object $s'$ together with morphisms $\beta_j : S_j \to s'$ having the two properties above, then $\beta$ should factor uniquely through $\alpha$, hence we postulate that

$$\exists! \gamma : s \to s' \forall j : \beta_j = \gamma \circ \alpha_j$$

holds (hence requesting information which is as precise as possible).

In other words: we are looking for a colimit for the functor $S$.

## 5 ER-Completeness

Fix a category $\mathcal{C}$ and a forest $\mathcal{B}$. The following definition is just an abbreviation.

**Definition 1** $\mathcal{C}$ is called **ER-complete** iff each $\mathcal{B}$-diagram over $\mathcal{C}$ has a colimit.

We want to investigate these categories first.

Suppose $S$ is a $\mathcal{B}$-diagram over $\mathcal{C}$, and let $\{w_1, \ldots, w_k\}$ be the roots of the trees in $\mathcal{B}, w_i$ being the root of the tree $T_i$. Restrict $S$ to this tree, obtaining a $T_i$-diagram over $\mathcal{C}$. If $v_i$ is its colimit, then $\prod_{i=1}^{n} v_i$ is the colimit for $S$. On the other hand, suppose that the colimit for $S$ exists, then all colimits for the trees exist, and their coproduct form the colimit. By the way, this colimit is the least upper bound of all node labels if $\mathcal{C}$ is a partially ordered set with an antisymmetric order relation.

We have demonstrated:

**Observation 1** An ER-complete category has coproducts.

We will show now that the converse also holds. Let $T$ be a tree with root $w$, and let $S$ be a $T$-diagram over $\mathcal{C}$. Let for each node $j$ in $T$ be $\pi_j$ the unique path from $j$ to $w$, and put $\tau_j^S := S\pi_j$, hence $\tau_j^S : S_j \to S_w$. Then $\langle S_w, \tau^S \rangle$ is a universal arrow from $S$ to $\Delta_T$. First, we have to show that $\langle S_w, \tau^S \rangle$ is in fact an arrow from $S$ to $\Delta_T$; this follows from the construction. Next, suppose that we have another arrow $\langle v, r \rangle$ from $S$ to $\Delta_T$, then $r_w : S_w \to v$ is a morphism in $\mathcal{C}$, and

$$r_j = r_w \circ \tau_j^S$$

holds for each node $j$. The latter property determines the factor for $\tau^S$ uniquely.

Thus we have established:

**Proposition 1** The category $\mathcal{C}$ is ER-complete iff it has all finite coproducts.

In fact, we have shown more: let

$$R^0_T : \mathcal{C}^T \to \mathcal{C}$$

be the map that assigns to each diagram $S$ the value $S_w$ at the root $w$ of tree $T$. The argumentation above indicates that

$$\eta_S : S \mapsto \Delta_T(R^0_T S)$$

constitutes a universal arrow from $S$ to $\Delta_T \circ R^0_T$, and from MacLane’s Portmanteau Theorem [Mac98, IV.1.2(ii)] we may conclude that $R^0_T$ is the object function of a functor $R_T$ which is the left adjoint to $\Delta_T$, specifically:

*An Note on a Categorical Semantics for ER-Models*
Proposition 2 \( (R_T, \Delta_T, \varphi) : C^T \rightarrow C \) is an adjunction, where \( \varphi \) is defined through

\[
\varphi_{S,c} : \begin{cases}
C(R_T S, c) & \rightarrow & C^T(S, \Delta_T(c)) \\
f & \rightarrow & \Delta_T(f) \circ \eta_S
\end{cases}
\]

Recall that the maps \( \varphi \) for an adjunction are bijections. Thus the result above may be restated as follows: the meaning of each diagram is uniquely determined by the meaning which is assigned to the root.

This reflects the usual emphasis and care on modeling the root class in an inheritance hierarchy, because this class essentially determines the properties of all its descendants. Recall, moreover, that these maps are natural transformations. Consequently, homomorphic changes to the modeling of the root induce homomorphic changes to the meaning of the entire hierarchy. The latter may be considered to be a sort of continuity property: smooth changes at the root do not induce drastic changes in the object hierarchy.

Let us turn to forests. The colimit of a forest was shown to be the coproduct, the factors being determined by the constituting trees. Denote for the diagram \( S \) and the forest \( B \) this colimit again by \( R_B S \), then we will show that there exists a natural bijection between \( C(R_B S, c) \) and \( C^B(S, \Delta_B(c)) \) for each diagram \( S \) and each object \( c \) in \( C \), hence the adjunction generalizes to forests.

Suppose again that \( \{w_1, \ldots, w_l\} \) is the collection of roots in \( B \), so that each node \( j \) lies in a uniquely determined tree with root \( w(j) \). Denote for the diagram \( S \) the unique morphism \( S_j \rightarrow S_{w(j)} \) by \( s_j \). Then

\[
R_B S = \coprod_{i=1}^{l} S_{w_i}.
\]

Denote the canonical injection \( S_{w_i} \rightarrow R_B S \) by \( r_{w_i} \). Fix for the moment \( S \) and \( c \), then a morphism \( f : R_B S \rightarrow c \) is uniquely determined by the morphisms

\[
\psi_{S,c}(f)_j := f \circ r_{w(j)} \circ s_j
\]

yielding for fixed \( S, c \) and \( f \) a natural transformation

\[
\psi_{S,c}(f) : S \rightarrow \Delta_B(c).
\]

It is not difficult to see that \( \psi_{S,c} \) is a bijection between \( C(R_B S, c) \) and \( C^B(S, \Delta_B(c)) \), and it now will be shown to be natural in \( S \) and in \( c \). Keeping \( c \) fixed, let \( \tau : S' \rightarrow S \) be a natural transformation. \( \tau \) induces a unique morphism

\[
\tau^B : R_B S' \rightarrow R_B S
\]

which makes the diagram

\[
\begin{array}{ccc}
S'_{w_i} & \xrightarrow{\tau_{w_i}} & \coprod_{i=1}^{l} S'_{w_i} \\
\downarrow \tau_{w_i} & & \downarrow \tau^B \\
S_{w_i} & \xrightarrow{r_{w_i}} & \coprod_{i=1}^{l} S_{w_i}
\end{array}
\]

An Note on a Categorical Semantics for ER-Models
commutative for each root \( v_i \) (the primed morphism refers to \( S' \)). Define
\[
\tau^B: \begin{cases} 
\mathcal{C}(R_S S, c) & \to \mathcal{C}(R_S S', c) \\
g & \mapsto g \circ \tau^B
\end{cases}
\]
and
\[
\tau_*: \begin{cases} 
\mathcal{C}(S, \Delta_{S}(c)) & \to \mathcal{C}(S', \Delta_{S}(c)) \\
\sigma & \mapsto \sigma \cdot \tau
\end{cases}
\]
and chase a morphism \( g : R_S S \to c \) around in the diagram (note the contravariance)
\[
\begin{array}{ccc}
\mathcal{C}(R_S S, c) & \xrightarrow{\psi_{S,c}} & \mathcal{C}(S, \Delta_{S}(c)) \\
\tau^B \downarrow & & \downarrow \tau_* \\
\mathcal{C}(R_S S', c) & \xrightarrow{\psi_{S',c}} & \mathcal{C}(S', \Delta_{S}(c))
\end{array}
\]
to obtain
\[
(g \circ \tau^B) \circ r'_{w(j)} \circ s'_{j} = g \circ r_{w(j)} \circ s_{j} \circ \tau
\]
This implies
\[
\psi_{S',c} \circ \tau^B \circ \tau_* = \tau_* \circ \psi_{S,c},
\]
hence the diagram above is commutative. Thus \( \psi \) is natural in \( S \), and it is immediate that it is natural in \( c \). Thus we have established another adjunction:

**Proposition 3** \(<R_S, \Delta_{S}, \psi>: \mathcal{C}^B \to \mathcal{C} \) is an adjunction.

Summarizing, we have proved

**Proposition 4** In a category with finite coproducts, the functor yielding the semantics of an ER-model is left adjoint to the diagonal.

Although we do not need it here, it might be interesting to note that the semantic functor is colimit-preserving [Mac98, p. 119]. This implies in particular that our semantics is compositional: if the ER-model is composed as, say, the coproduct of smaller models, then the semantics behaves civilized in the sense that it composes from the semantics of the coproduct’s factors, and similarly for other colimits. These aspects should be investigated further.

### 6 Relations

\( \mathcal{R} \) is the finite set of binary relations for the ER model with diagram \( \mathcal{B} \). We assume that the category \( \mathcal{C} \) has finite products as well as finite coproducts, so that \( S_B \) is defined for the schema \( S \). For each relation \( r \in \mathcal{R} \) with \( \delta(r) \) and \( \gamma(r) \) as the domain and the codomain entities with associated nodes \( n_{\delta(r)} \) and \( n_{\gamma(r)} \), resp., let \( j^\delta(r) \) and \( j^\gamma(r) \) be the corresponding new nodes in \( \mathcal{B} \). Since \( \mathcal{C} \) has finite products, \( (S n_{\delta(r)}) \times (S n_{\gamma(r)}) \) exists in \( \mathcal{C} \) with respective projections \( \pi_{\delta(r)} \) and \( \pi_{\gamma(r)} \). Because relation \( r \) may always be represented as a subobject of the Cartesian product of the domain and the codomain, we assume that there exists a monic
\[
\rho^r : c_r \to (S n_{\delta(r)}) \times (S n_{\gamma(r)})
\]

An Note on a Categorical Semantics for ER-Models
for some object $c_r$ in $\mathcal{C}$. Hence we label the edges $(j^{\delta(r)}, n_{\delta(r)})$ and $(j^{\gamma(r)}, n_{\gamma(r)})$ with $\pi_{\delta(r)} \circ \rho^r$ and $\pi_{\gamma(r)} \circ \rho^r$, resp. The nodes are added together with the edges to the graph. This transmogrifies the graph construction outlined in [Dob97], cp. Section 4. This process yields a new diagram $\mathcal{B}^*$ and a new schema $S^*$ with a colimit $\Sigma := S^*_r$.

We want to investigate the relationship between an arbitrary relation $r$ and $\Sigma$. By construction, $r$ helps to define the colimit.

**Just a brief aside.** Recall in this stage of the development that each object $\ell \in \mathcal{C}$ gives rise to a set-valued functor

$$\Lambda_\ell : \mathcal{C} \rightarrow \mathcal{S},$$

the latter denoting the category of all small sets, upon putting

$$\Lambda_\ell := \mathcal{C}(\ell, -),$$

mapping a morphism $f : a \rightarrow b$ to

$$\Lambda_\ell f : \left\{ \begin{array}{l} \mathcal{C}(\ell, a) \rightarrow \mathcal{C}(\ell, b) \\ g \mapsto f \circ g \end{array} \right.$$  

The famous Yoneda Lemma [Mac98, p. 61] identifies the natural transformations $\Lambda_\ell \bullet \rightarrow \Lambda_\ell$ with the morphisms in $\mathcal{C}(\ell, t)$. Similarly,

$$\Gamma_\ell := \mathcal{C}(-, \ell)$$

defines a (contravariant) functor $\Gamma_\ell : \mathcal{C} \rightarrow \mathcal{S}$ mapping $f : a \rightarrow b$ to

$$\Gamma_\ell f : \left\{ \begin{array}{l} \mathcal{C}(b, \ell) \rightarrow \mathcal{C}(a, \ell) \\ g \mapsto g \circ f \end{array} \right.$$  

**Let us return to relations.** The first scenario capitalizes on the assumption that $\rho^r$ is a monic. Thus the natural transformation

$$\Gamma_{c_r} \rightarrow \Gamma_{(Sn_{\delta(r)} \times (Sn_{\gamma(r)})}$$

implied by this morphism, again denoted by $\rho^r$, is a monic, too, when considered as a morphism in the category $\mathcal{S}^\mathcal{B}$ of all functors. In particular, $\rho^r$ induces an injective map

$$\mathcal{C}(\Sigma, c_r) \rightarrow \mathcal{C} \left( (Sn_{\delta(r)} \times (Sn_{\gamma(r)}), c_r \right).$$

This means that each operation from $\Sigma$ to the object $c_r$ corresponds uniquely to an operation from $\Sigma$ to $(Sn_{\delta(r)} \times (Sn_{\gamma(r)}$. Since there exists the embedding as a morphism from each node to $\Sigma$, each morphism from a node to $c_r$ gives rise to a morphism from $\Sigma$ to $(Sn_{\delta(r)} \times (Sn_{\gamma(r)}$. This models the flow of information from the colimit to the relation (and does probably not constitute an entirely surprising observation).

Next, suppose that $\rho^r$ is a split mono, thus it has a left inverse. This is e.g. the case whenever $\mathcal{C}$ is a subcategory of $\mathcal{S}$. The Yoneda Lemma implies that the induced natural transformation

$$\Lambda_{(Sn_{\delta(r)} \times (Sn_{\gamma(r)}) \rightarrow \Lambda_{c_r}$$

*An Note on a Categorical Semantics for ER-Models*
is an epi (cp. [Mac98, Lemma IV.3]), again as a morphism in the functor category $S^B$. In particular, the induced map

$$\mathcal{C}((S_{\delta(r)} \times (S_{\gamma(r)}), \Sigma) \rightarrow \mathcal{C}(c_r, \Sigma)$$

is an epi between sets. Epis in $S$ are exactly the onto maps. Reformulating, each operation $c_r \rightarrow \Sigma$ is induced by an operation $(S_{\delta(r)} \times (S_{\gamma(r)}) \rightarrow \Sigma$. This applies in particular to operations between $c_r$ and $Sp$, where $n$ is an arbitrary node in the forest.

Alas, this is about how far the general considerations on relations can go. More specific results require more specific assumptions, and this is — as usual — indicated as subject to further work.

References


