Some connections between Abstract Elementary Classes and Accessible Categories

Marcos Mazari Armida

February 21, 2017

Abstract

We connect two generalizations of first order Model Theory. On one hand we have Abstract Elementary Classes which were introduced in the late seventies by Shelah in [4]. On the other we have Accessible Categories which were introduced in the late eighties by Makkai and Paré in [3]. In this paper we will see that on some natural framework both notions coincide. Most of what we present here was presented by Lieberman in [2].

1 Basic Model Theory

In this section we will introduce the basic concepts of first order Model Theory, for a detailed exposition you may wish to consult [1].

Definition 1.1. A signature Σ is a triple $\langle C, (F_n)_{n \in \mathbb{Z}^+}, (R_n)_{n \in \mathbb{Z}^+} \rangle$, where C is the set of constants, for each $n \in \mathbb{Z}^+$ F_n is the set of n-ary functions and for each $n \in \mathbb{Z}^+$ R_n is the set of n-ary relations. Furthermore we require that all the sets are pairwise disjoint.

Given a signature Σ the next definition that one should introduce is that of Σ -terms and Σ -formulas, the definition of these are analogous to that given for *first order logic* in class, so we'll omit them. Although the definition of Σ -structure given in class is close to what we mean by Σ -model, we'll present it below for sake of exposition.

Definition 1.2. Given a signature Σ , a Σ -model M is a pair $\langle |M|, I \rangle$ where |M| is the universe and I is the interpretation such that:

- |M| is a nonempty set.
- Given $c \in C$, $I(c) \in |M|$.
- Given $n \in \mathbb{Z}^+$ and $f \in \mathbf{F}_n$, $I(f) : |M|^n \to |M|$.
- Given $n \in \mathbb{Z}^+$ and $r \in \mathbf{R}_n$, $I(r) \subseteq |M|^n$.

We will usually denote the structure and universe by M and instead of writing I(c), I(f) and I(r) we will write c^M , f^M and r^M . Again the natural step would be to define

interpretation of terms and satisfaction of formulas given a structure M but this is similar to what we did in class and as we will see below one of the advantages of AECs is that we don't have to deal with them.

As before the notions of morphism between Σ -models is close to that of Σ -structures presented in class, but due to its importance we'll present it below.

Definition 1.3. Given $M, N \Sigma$ -models, a function f from the universe of M to the universe of N is a morphism, denoted by $f : M \to N$, if the following hold:

- Given $c \in C$, $f(c^M) = c^N$.
- Given $n \in \mathbb{Z}^+$, $g \in \mathbf{F}_n$ and $\bar{a} \in M^n$, $f(g^M(\bar{a})) = g^N(f(\bar{a}))$.
- Given $n \in \mathbb{Z}^+$, $r \in \mathbf{R}_n$ and $\bar{a} \in M^n$, $\bar{a} \in r^M$ if and only if $f(\bar{a}) \in r^N$.

Of primary importance is the case when $|M| \subseteq |N|$ and f = i (the inclusion), in that case we say that M is a substructure of N and we denote it by $M \subseteq N$.

We say that f is a monomorphism if f is injective and we say that f is an isomorphism if it is bijective.

Although there is something to say regarding morphisms in general, the notion is too weak to preserve formulas, i.e., given $f: M \to N$, $\bar{a} \in M$ and $\phi(\bar{x})$ it is not true that:

$$M \models \phi[\bar{a}]$$
 iff $N \models \phi[f(\bar{a})]$

Therefore the right notion for relating Σ -models is that of elementary embedding.

Definition 1.4. Given M, N two Σ -models, a function f from the universe of M to the universe of N is an elementary embedding, denoted by $f: M \to N$, if f is a morphism and for every $\bar{a} \in M$ and $\phi(\bar{x})$:

$$M \vDash \phi[\bar{a}] \quad iff \quad N \vDash \phi[f(\bar{a})].$$

Of primary importance is the case $|M| \subseteq |N|$ and f = i (the inclusion), in that case we say that M is an elementary substructure of N and we denote it by $M \preceq N$.

Definition 1.5. Let Σ be a signature. $\langle K, \leq_K \rangle$ is an elementary class if there is T a set of first order Σ -sentences (formulas without free variables) such that $K = Mod(T) = \{M | \forall \phi \in T(M \vDash \phi)\}$ and $\leq_K = \preceq$.

Finally, let us define two categories.

Definition 1.6. Fix a Σ signature.

- Let Σ-Mon the category in which objects are Σ-models and arrows are monomorphisms.
- Let Σ Emb the category in which objects are Σ-models and arrows are elementary embeddings.

2 Abstract Elementary Classes

An Abstact Elementary Class (AEC) is a semantic generalization of an *elementary* class. In this section we will introduce some basic definitions and prove some basic lemmas, for a detailed exposition the reader may want to consult [1].

Definition 2.1. Given a signature Σ . An ordered pair $\langle K, \leq_K \rangle$ is an Abstract Elementary Class (AEC) if K is a class of Σ -structures such that:

- 1. (K, \leq_K) is a preorder and if $M \leq_K N$ then $M \subseteq N$.
 - (Closure under isomorphisms) If $N \in K$ and $N \cong M$, then $M \in K$.
 - Let $N_1, N_2 \in K$ and $M_1, M_2 \in K$, if the following diagram commutes:

$$\begin{array}{ccc} M_1 & \stackrel{\subseteq}{\longrightarrow} & M_2 \\ \downarrow \cong & & \downarrow \cong \\ N_1 & \stackrel{\leq_K}{\longrightarrow} & N_2 \end{array}$$

then $M_1 \leq_K M_2$.

- (Coherence) If $M_1, M_2, M_3 \in K$, $M_1 \leq_K M_3$, $M_2 \leq_K M_3$ and $M_1 \subseteq M_2$ then $M_1 \leq_K M_2$.
- 2. (Downward Löwenheim-Skolem [DLS]) There is a cardinal $\lambda \ge |\Sigma| + \aleph_0$ such that for every $M \in K$ and $A \subseteq M$, there is $N \in K$ such that $A \subseteq |N|$, $N \le_K M$ and $||N|| \le \lambda + |A|$. The minimal such λ is denoted by LS(K).
- 3. (Chain axiom) For every α limit ordinal.
 - If $\langle M_i | i < \alpha \rangle \subseteq K$ is a chain, i.e., $\forall i < j < \alpha(M_i \leq_K M_j)$, then $M = \bigcup_{i < \alpha} M_i \in K$ and morever $\forall i < \alpha(M_i \leq_K M)$.
 - If $N \in K$ and $\langle M_i | i < \alpha \rangle \subseteq K$ is a chain and $\forall i < \alpha(M_i \leq_K N)$, then $M = \bigcup_{i < \alpha} M_i \leq_K N$.

We will always assume that there aren't any models of size less than the LS(K), the reason we can do that is because if there were we could just define K' to be the class that didn't have them and this would still be an AEC. Also in the studying of AECs we are interested in big cardinalities. Before moving on let us give a series of examples.

Examples. • Every elementary class is an AEC with $LS(K) = |\Sigma(T)| + \aleph_0$.

- Let $LF = \langle LF, \leq_{LF} \rangle$, where $M \in LF$ iff M is a locally finite group, i.e., every finite subset of M generates a finite group, and $M \leq_{LF} N$ iff M is a subgroup of N.
- Let ϕ a sentence in $L_{\omega_1,\omega}$ then $\langle Mod(\phi), \subseteq \rangle$ is an AEC.

Definition 2.2. Given λ an infinite cardinal, let $K_{\lambda} = \{M \in K |||M|| = \lambda\}$ and let $I(\lambda, K) = |K_{\lambda}/\cong|$. We say that K is λ -categorical if $I(\lambda, K) = 1$.

Definition 2.3. • A poset $\langle I, \leq \rangle$ is said to be λ - directed if for any $X \subseteq I$ with $|X| < \lambda$, there is $i \in I$ such that $\forall x \in X(x \leq i)$. We say it is directed if $\lambda = \omega$.

 Given (I, ≤) a poset, we say that {M_i|i ∈ I} ⊂ K is a directed system if i ≤ j then M_i ≤_K M_j.

The following two lemmas will be very useful when connecting AECs to Accessible Categories.

Lemma 2.1. In axiom 3. of an AEC we can change "chain" for "directed system over $a \omega$ -directed poset".

Proof. The proof is done by induction on the size of I and can be consulted in [1]. \Box

Lemma 2.2. Let K be an AEC and $M \in K$. There is $\langle I, \leq \rangle$ directed poset and $\{M_i | i \in I\} \subseteq K_{LS(K)}$ a directed system such that $M = \bigcup_{i \in I} M_i$.

Proof. Let $I = \{A \subseteq M | |A| < \aleph_0\}$ and we say that $A \leq_I B$ iff $A \subseteq B$. Clearly I is a directed poset. Now, let us build $\{M_A | A \in I\}$ by induction on the size of A.

- Base: If |A| = 0, then $A = \emptyset$. Let $B \subseteq M$ with |B| = LS(K). Then apply DLS axiom to get $M_A \leq M$ s.t. $||M_A|| = LS(K)$.
- Induction step: Let |A| = n + 1. Let $E = \{B|B \subsetneq A\}$, by induction hypothesis for each $B \in E$ there is $M_B \in K_{LS(K)}$ with $M_B \leq M$. Let $X = \bigcup_{B \in E} M_B \cup A$ and let M_A the structure obtained by applying DLS axiom to X in M. Clearly $A \subseteq M_A \in K_{LS(K)}, M_A \leq M$ and by coherence for each $B \in E(M_B \leq M_A)$.

It is easy to see that the construction above makes $\{M_A | A \in I\}$ into a directed system and clearly $M = \bigcup_{A \in I} M_A$, since given $m \in M$ we have that $m \in M_{\{m\}}$. \Box

We will refer to the poset and directed system constructed above as I_M .

Since we are interested in connecting AECs to Category Theory, let us define the *arrows* of an AECs.

Definition 2.4. Given $M, N \in K$ a function f from |M| to |N| is a K-embedding/morphism, denoted by $f: M \to N$, iff there is $M^* \in K$ such that $f: M \cong M^*$ and $M^* \leq_K N$.

Realize that in particular every K-embedding is injective.

To conclude with this section, we would like to point out that the study of the function I(-, -) is the main line of study in AECs, the main conjecture is the following.

Shelah's Categoricity conjecture. Let K be an AEC. If K is λ -categorical for $\lambda \geq \beth_{(2^{LS(K)})^+}$ then K is μ -categorical for all $\mu \geq \beth_{(2^{LS(K)})^+}$.

What we need for our discussion is the upper bound for I(-, -).

Lemma 2.3. Let K be an AEC and $\lambda \geq LS(K)$, then $I(\lambda, K) \leq 2^{\lambda}$.

Proof. (Idea) Given λ there is λ ways to interpret each constant, there is λ^{λ} ways to interpret each relation and function, therefore

$$I(\lambda, K) \le (\lambda + \lambda^{\lambda} + \lambda^{\lambda})^{|\Sigma|} = 2^{\lambda}.$$

ſ			

3 Accesible categories and other categorical notions

It is time to turn our attention to Category Theory, we will start with two basic definitions that will be used to define what we understand by an Accessible Category. For a detailed exposition the reader may want to consult [3].

Definition 3.1. A λ -directed colimit in a category \mathbb{C} is a colimit in which the indexing category is a λ -directed poset.

Definition 3.2. An object N in \mathbb{C} is λ -presentable if $Hom_{\mathbb{C}}(N, -) : \mathbb{C} \to Sets$ preserves λ -directed colimits.

Realize that in this case preserving λ -directed colimits means that given $(C, c_i : D_i \to C)$ a cocone and $f : N \to C$ there is a $i \in I$ and $g : N \to D_i$ such that $f = c_i \circ g$. Moreover, if $g' : N \to D_i$ is such that $f = c_i \circ g'$, then there is $j \ge i$ such that $D_{i\to j} \circ g = D_{i\to j} \circ g'$. This follows from the way directed colimits are constructed in Sets.

With this two definitions we are ready to introduce the concept of Accessible Category.

Definition 3.3. Let λ a regular infinite cardinal. A category \mathbb{C} is λ -accessible if:

- \mathbb{C} is closed under λ -directed colimits.
- \mathbb{C} contains only a set of λ -presentable objects up to isomorphisms.
- Every object in \mathbb{C} is a λ -directed colimit of λ -presentables.

We say that \mathbb{C} is accessible if there is a regular λ such that \mathbb{C} is λ -accessible.

There are three other categorical notions that will play a central role in our main theorems.

Definition 3.4. Let \mathbb{C} a category and \mathbb{D} a subcategory of \mathbb{C} .

- D is a replete subcategory of C if given A ∈ obj(D) and f : A → B isomorphism in C we have that f and B are in D.
- \mathbb{D} is a coherent subcategory of \mathbb{C} if for every $A_1, A_2, A_3 \in obj(\mathbb{D})$ and $h, g \in Arrow(\mathbb{D})$ such that $g: A_1 \to A_3$ and $h: A_2 \to A_3$ if there is $f: A_1 \to A_2 \in Arrow(\mathbb{C})$ such that $h \circ f = g$, then f is already in \mathbb{D} .
- \mathbb{D} is a strong subcategory of \mathbb{C} if for every $g: A_1 \to A_2 \in Arrow(\mathbb{D})$, if there are $k: A_1 \to A_3 \in Arrow(\mathbb{D})$ isomorphism and $l: A_3 \to A_1 \in Arrow(\mathbb{C})$ such that $g = l \circ k$, then $l \in Arrow(\mathbb{D})$.

4 Some connections

Now that we have all the definitions out of the way it is time to connect the concepts introduced in the previous two sections.

Construction 4.1. Realize that given any $AEC \langle K, \leq_K \rangle$, this can be seen as a category \mathbb{K} where $Obj(\mathbb{K}) = K$ and given $M, N \in K$, $f \in Arrow_{\mathbb{K}}(M, N)$ if and only if f is a K-embedding from M to N. We will identify \mathbb{K} and $\langle K, \leq_K \rangle$ with simply K.

Now that we have a new category K we are ready to present the first half of our main theorem.

Theorem 4.1. Let K be an AEC. Then K (or more precisely \mathbb{K}) has all directed colimits and is λ -accessible for all regular $\lambda > LS(K)$. Moreover \mathbb{K} is a replete, coherent and strong subcategory of the category $\Sigma - Mon$.

Proof. Follows from lemmas 4.1-4.5.

Let us prove the series of lemmas referred above. In what follows K and λ regular are fix. Let us start by calculating directed colimits in K.

Lemma 4.1. K is closed under directed colimits.

Proof. Let μ a cardinal and $I = (I, \leq)$ be a μ -directed poset. We have $\langle \{M_i\}_{i \in I}, \{f_{i,j} : M_i \to M_j\}_{i \leq j} \rangle$. We may assume with out lost of generality that if $i \neq j$ then $M_i \cap M_j = \emptyset$.

Let $A = \bigcup_{i \in I} M_i$ and given $a, b \in A$ such that $a \in M_{i_0}$ and $b \in M_{i_1}$ we say that $a \sim b$ iff $\exists j \geq i_0, i_1(f_{i_0,j}(a) = f_{i_1,j}(b))$. Using that I is μ -directed it follows that it is an equivalence relation.

Let $M = \langle |M|, I \rangle$ where $|M| = A / \sim$ and I is defined as follows:

- Given $c \in \mathbf{C}$ we define $c^M = [c^{M_i}]$ for some $i \in I$.
- Given $f \in \mathbf{F}_n$ and $[a_1], ..., [a_n] \in A$ s.t. $a_1 \in M_{i_1}, ..., a_n \in M_{i_n}$, let $j \ge i_m$ for all $m \in \{1, ..., n\}$. We define:

$$f^{M}([a_{1}],...,[a_{n}]) = [f^{M_{j}}(f_{i_{1},j}(a_{1}),...,f_{i_{n},j}(a_{n})).]$$

• Given $r \in \mathbf{R}_n$ and $[a_1], ..., [a_n] \in M$ s.t. $a_1 \in M_{i_1}, ..., a_n \in M_{i_n}$ we define:

$$\langle [a_1], ..., [a_n] \rangle \in r^M$$
 iff $\exists j \in I(\langle f_{i_1,j}(a_1), ..., f_{i_n,j}(a_n) \rangle \in r^{M_j}).$

It is easy to show that M is a well-defined Σ -model. Moreover, given $i \in I$, let $g_i : M_i \to M$ given by $g_i(a) = [a]$; it follows that $M_i \cong g_i[M_i]$, therefore by closure under isomorphisms $g_i[M_i] \in K$ for all $i \in I$. Observe that $\{g_i[M_i]|i \in I\}$ is a directed system over a ω -directed poset and $M = \bigcup_{i \in I} g_i[M_i]$, then by lemma 2.1 we have that $M \in K$ and that $\forall i \in I(g_i[M_i] \leq M)$, hence $\forall i \in I(g_i$ is a K-embeding). Since given $i \leq j$ we have that $g_i = g_j \circ f_{i,j}$, we can conclude that $(M, \{g_i\}_{i \in I})$ cocone in K.

To show it is initial, suppose that $(C, \{c_i : M_i \to C\}_{i \in I})$ is another cocone, then define $h : M \to C$ by $h([a]) = c_i(a)$ where $a \in M_i$. It follows that h is K-embedding by the second part of the chain axiom. Hence $(M, \{g_i\}_{i \in I})$ is a μ -directed colimit.

An important thing to observe is that the colimits are computed the same way as they are computed in $\Sigma - Mon$.

Lemma 4.2. Le $M \in K$. M is λ -presentable if and only if $||M|| < \lambda$.

Proof. \longrightarrow Suppose M is λ -presentable. Let $I = \{A \subseteq M | |A| < \lambda\}$ and given $A, B \in I$ we say that $A \leq B$ iff $A \subseteq B$. Using that λ is regular it follows that (I, \leq) is a λ -directed poset.Let I_M be the poset described in lemma 2.2 and $\{M_A | A \in I_M\}$ its directed system. Let us define $D: I \to K$ as follows:

$$D(A) = \begin{cases} M_A & \text{if } A \text{ is finite} \\ \bigcup_{B \subseteq_{fin} A} M_B & \text{if } A \text{ is infinite} \end{cases}$$

By the second part of the chain axiom D is a functor and observe that since λ is regular for each $A \in I$ we have that $||M_A|| < \lambda$. Let $C = (\bigcup_{A \in I} D(A), g_A : D(A) \hookrightarrow M)$, it is easy to see that this is a λ -directed colimit and that $\bigcup_{A \in I} D(A) = M$.

Consider $1_M : M \to M$, since M is λ -presentable there is $A \in I$ and $g : M \to D(A)$ s.t. $f = g_A \circ g$. In particular g is an injective function, hence $||M|| \leq |D(A)| < \lambda$.

 $\begin{array}{l} \overleftarrow{\leftarrow} \text{Suppose } ||M|| < \lambda. \text{ Let } I = (I, \leq) \text{ be a } \lambda-\text{directed poset. We have } \langle \{N_i\}_{i \in I}, \{f_{i,j} : N_i \to N_j\}_{i \leq j} \rangle \text{ and } (N, g_i) \text{ its colimit as in previos lemma (just change M for N in it).} \\ \text{Let } f: M \to N \text{ a } K-\text{embedding. Clearly}|f[M]| < \lambda, \text{ then from the fact that } I = (I, \leq) \\ \text{ is a } \lambda-\text{directed, there is } A \in I \text{ such that } f[M] \leq g_A[N_A]. \text{ Let } g = g_A^{-1} \circ f : M \to N_A \\ \text{where } g_A^{-1} \text{ is the inverse on the image. It is easy to see that } g \text{ satisfies the definition} \\ \text{of } \lambda-\text{presentable. So } M \text{ is } \lambda-\text{presentable.} \end{array}$

Lemma 4.3. K contains only a set of λ -presentable objects up to isomorphisms.

Proof. By previous lemma $\{M \in K | M\lambda - \text{presentable}\} = \bigcup_{LS(K) \le \mu < \lambda} K_{\mu}$. By lemma 2.3 it follows that for a fix μ there are at most 2^{μ} nonisomorphic models, hence:

$$|\{M \in K | M\lambda - \text{presentable}\}| \cong | \leq \sum_{LS(K) \leq \mu < \lambda} 2^{\mu} \leq 2^{\lambda}$$

So there are at most 2^λ $\lambda\text{-presentable objects up to isomorphisms, so set many of them.$

Lemma 4.4. Every object in K is a λ -directed colimit of λ -presentables. Moreover K is λ -accesible

Proof. The first part follows directly from the construction done in lemma 4.2.. As for the second, it follows from lemmas 4.1, 4.2 and 4.3 and from what we just said.

Lemma 4.5. \mathbb{K} is a replete, coherent and strong subcategory of the category $\Sigma - Mon$.

Proof. K is replete and coherent because K is closed under isomorphisms and coherent. To show that it is strong let $g: M \to N$ K-embedding and suppose there are $k: M \to M'$ isomorphism and $l: M' \to N$ monomorphism with $g = l \circ k$, then realize that we have the following commutative diagram:

$$\begin{array}{ccc} l[M'] & \stackrel{\subseteq}{\longrightarrow} & M \\ & & \downarrow \cong_{Id} & & \downarrow \cong_{Id} \\ g \circ k^{-1}[M'] & \stackrel{\leq \kappa}{\longrightarrow} & M \end{array}$$

Then from axiom 1. of AECs it follows that $l[M'] \leq M$. So l is a K-embedding.

This finishes the proof of the first half of our main theorem, before stating the second half let us present a second construction.

Construction 4.2. Let \mathbb{C} a subcategory of $\Sigma - Mon$. Then consider $C = (C, \leq_C)$ where $C = obj(\mathbb{C})$ and given $M, N \in C$ we define $M \leq_K N$ if and only if $M \subseteq N$ and $i : M \hookrightarrow N \in Arrow(\mathbb{C})$. Realize that C might not be an AEC, to achive that we will need further conditions.

Let us present the second half of our main theorem.

Theorem 4.2. Let $\mu \geq |\Sigma| + \aleph_0$ infinite cardinal. If \mathbb{C} is a replete, coherent and strong subcategory of $\Sigma - Mon$, \mathbb{C} is closed under directed colimits (computed as in $\Sigma - Mon$) and is λ -accessible for all regular $\lambda > \mu$. Then C (as defined in construction 4.2) is an AEC. Moreover $LS(C) = \mu$.

Proof. (Sketch) Axiom 1. follows from the fact that \mathbb{C} is replete, coherent and strong subcategory of $\Sigma - \mathbb{M}on$. Axiom 3. follows from the fact that \mathbb{C} is closed under directed colimits, so let us show that axiom 2. holds. Let $A \subseteq M \in C$. Let $\lambda = |A| + \mu$, observe that $\lambda^+ > \mu$ and regular. By hypothesis C is λ^+ -accessible, so M is a λ^+ -directed colimit of λ^+ -presentable objects.

As colimits in C are computed as in $\Sigma - Mon$ we have that $M = \bigcup_{i \in I} g_i[N_i]$ where (I, \leq) is a λ^+ -directed, $\langle \{N_i\}_{i \in I}, \{f_{i,j} : N_i \to N_j\}_{i \leq j} \rangle$ is the image under the diagram, each N_i is λ^+ - presentable and $(N, \{g_i\}_{i \in I})$ is the colimit (as in lemma 4.1). Since $|A| < \lambda^+$ and I is λ^+ -directed, there is $i \in I$ such that $A \subseteq g_i[N_i]$.

Since C is a strong subcategory, we have that $g_i[N_i] \leq M$. Moreover, a similar proof to lemma 4.2 but in $\Sigma - \mathbb{M}on$ shows that $|g_i[N_i]| < \lambda^+$. Hence $A \subseteq g_i[N_i]$, $g_i[N_i] \leq M$ and $|g_i[N_i]| \leq \lambda = |A| + \mu$.

References

- [1] Rami Grossberg, A course in Model Theory I, In preparation.
- [2] Michael J. Lieberman, Topological and Category-Theoretic Aspects of Abstract Elementary Classes, PhD Thesis: University of Michigan, 2009.
- [3] M. Makkai and R. Paré, Accessible Categories: The Foundations of Categorical Model Theory, AMS 1989.
- [4] Saharon Shelah, Classification of nonelementary classes II, Abstract elementary classes, In Classification theory (Chicago, IL, 1985), volume 1292 of Lecture Notes in Mathematics, pages 419497. Springer, Berlin, 1987. Proceedings of the USAIsrael Conference on Classification Theory, Chicago, December 1985; ed. Baldwin, J.T.