

# THE ABSTRACTION MYSTIQUE

G. Aldo Antonelli

Department of Philosophy  
University of California, Davis

Workshop on Set Theory and Higher-Order Logic  
5 August 2011



- What are Abstraction Principles?
- Definitions and preliminary results
- Satisfiability of abstraction principles
- Logical invariance
- Notions of invariance for abstraction principles
- Abstraction and demarcation problems
- Assessment and Conclusion

# WHAT ARE ABSTRACTION PRINCIPLES?

The notion of a “classifier” is known from descriptive set theory:

## DEFINITION

If  $R$  is an equivalence relation over a set  $X$ , a *classifier* for  $R$  is a function  $f : X \rightarrow Y$  such that  $f(x) = f(y) \iff R(x, y)$ .

An *abstraction operator* is a classifier  $f$  for the specific case in which  $X = \mathcal{P}(Y)$ , i.e., an assignment of first-order objects to “predicates” (subsets of the first-order domain), which is governed by the given equivalence relation.

Abstraction operators are particular functional terms that take predicates (“concepts,” in Frege’s sense) as input. The statement that an operator  $f$  assigns objects to concepts according to an equivalence  $R$  is called an *abstraction principle*:

$$f(X) = f(Y) \iff R(X, Y).$$

Philosophers often view abstraction principles as the preferred vehicle for the delivery of a special kind of objects — *abstract entities* — whose somewhat mysterious nature includes such properties as non-spatio-temporal existence and causal inefficacy.

One particular abstraction principle, known as *Hume's Principle* (HP) plays a crucial role in the neo-Fregean program initiated by Crispin Wright and Bob Hale.

HP assigns objects to concepts on the basis of the *equinumerosity* relation  $\approx$  between concepts:

$$N(X) = N(Y) \iff X \approx Y,$$

where the object  $N(X)$  assigned to  $X$  is interpreted as “the number of  $X$ ,” and HP is variously advertised as being *logically true*, *analytic*, or *constitutive* of the notion of number.

Some abstraction principles *can* be inconsistent. The first such example is “Basic Law V” from Frege’s *Grundgesetze*, which assigns objects to concepts based on the equiextensionality relation:

$$\text{Ext}(X) = \text{Ext}(Y) \iff \forall z(z \in X \leftrightarrow z \in Y).$$

The operator  $\text{Ext}$  is then an injection of concepts into objects. If the domain of the concepts is identified with the true power-set of the first-order domain, this contradicts Cantor’s theorem.

In principle, consistency can be restored either by making the operator *partial* or the equivalence *coarser* (or both).

Inconsistent abstraction principles give rise to the *Demarcation problem*: the problem of identifying necessary and sufficient conditions for separating the “bad” abstraction principles from the “good” ones.

- We fix throughout a non-empty first-order domain  $D_1$  and a second-order domain  $D_2 \subseteq \mathcal{P}(D_1)$ .
- We assume that  $D_2$  always contains the first-order definable subsets of  $D_1$  (in some given background language  $\mathcal{L}$ ).
- The interpretation  $(D_1, D_2)$  is *standard* if  $D_2 = \mathcal{P}(D_1)$ .

## DEFINITION

An *abstraction operator* is a functional assignment  $f_R : D_2 \rightarrow D_1$  satisfying an abstraction principle  $\text{Ab}_R$  of the form

$$\text{Ab}_R \quad f_R(X) = f_R(Y) \iff R(X, Y),$$

where  $R$  is an equivalence relation over  $D_2$ .

## DEFINITION

An abstraction operator  $f$  is *choice-like* if  $f(X) \in X$  whenever  $X$  is non-empty.

## THEOREM (CANTOR)

If  $(D_1, D_2)$  is standard, then no abstraction operator is injective.

## PROPOSITION

There is no choice-like injective abstraction operator.

## PROOF.

Obvious if  $D_2$  is standard. Even if  $D_2$  is non-standard,  $|D_1| > 1$  (in fact infinite), and the result follows from the fact that there is no injective choice function on  $\{\{a\}, \{b\}, \{a, b\}\}$ .  $\square$

## COROLLARY

Hilbert's  $\varepsilon$ -calculus cannot be consistently augmented by adding the axiom  $\varepsilon x\varphi(x) = \varepsilon x\psi(x) \rightarrow \forall x(\varphi(x) \leftrightarrow \psi(x))$ .

## DEFINITION

Given equivalences  $R$  and  $S$ , define  $R \preceq S$  ( $R$  is *finer* than  $S$ ,  $S$  is *coarser* than  $R$ ), if and only if  $R(x,y)$  implies  $S(x,y)$ .

## THEOREM

If  $R$  and  $S$  are equivalences over  $D_2$  and  $R \preceq S$ , then if  $\text{Ab}_R$  is satisfiable over  $(D_1, D_2)$ , so is  $\text{Ab}_S$ .

## PROOF.

Assume  $f_R$  satisfies  $\text{Ab}_R$  and let  $f_R[X]_S$  be the point-wise image of  $[X]_S$  under  $f_R$ . Let  $g$  be a choice function for the set  $\{f_R[X]_S : X \in D_2\}$ , and define  $f_S(X) = g(f_R[X]_S)$ . Then  $f_S$  satisfies  $\text{Ab}_S$ . □

## THEOREM

For every equivalence  $R$ ,  $Ab_R$  is satisfiable.

## PROOF.

There is a (perforce non-standard) domain  $(D_1, D_2)$  over which  $Ab_=$  (i.e., Frege's "Basic Law V") is satisfiable (see Parsons, Heck, Wehmeier, etc.). The relation  $=$  is, of course, the finest equivalence relation over  $D_2$ . □

## REMARK

In *Grundgesetze* Frege went to great length to derive Hume's Principle from Basic Law V. The semantic proof just given shows that not just HP, but every abstraction principle is reducible to BLV.

## DEFINITION

A permutation  $\pi$  is a one-to-one function from  $D_1$  onto itself;  $\pi$  can be “lifted” to subsets of  $D_1$ :  $\pi[X] = \{\pi(x) : x \in X\}$ .

The domain  $D_2$  is  $p$ -closed if  $X \in D_2$  implies  $\pi[X] \in D_2$ .

## DEFINITION

A permutation  $\pi$  is *finitely based* if  $\pi(x) = x$  for all but finitely many  $x \in D_1$ .

For subsets  $X$  and  $Y$  of  $D_1$ , the equivalence  $X \sim Y$  holds if and only if  $\pi[X] = Y$  for some finitely based  $\pi$ .

## DEFINITION

The *Nuisance Principle*, NP, governs the assignment of objects to concepts according to  $\sim$ :

$$\text{Ab}_{\sim} \quad \nu(X) = \nu(Y) \iff X \sim Y.$$

## THEOREM

If  $|D_1| \geq \aleph_0$  then NP is satisfiable in  $(D_1, D_2)$  if and only if  $|D_2| \leq |D_1|$ .

## PROOF.

If  $|D_1| = \kappa \geq \aleph_0$ , for each  $X \in D_2$ , the cardinality of  $[X]_\sim$  is bounded by  $\kappa$ , i.e., the number of finite sets of pairs in some finitely based  $\pi$  that are of the form  $\langle x, y \rangle$  with  $x \neq y$ .

In one direction, using  $\nu$ , the equivalence classes themselves can be *indexed* by a set of size  $\kappa$ , i.e.,  $D_1$ , so that  $|D_2| \leq \kappa^2 = \kappa$ .

Conversely, if  $f : D_2 \rightarrow D_1$  is an injection, let  $\nu$  be a choice function for

$$\{f[X]_\sim : X \in D_2\},$$

where  $f[X]_\sim$  is the pointwise image of  $[X]_\sim$  under  $f$ . Then  $\nu$  satisfies NP.



## COROLLARY 1 (BOOLOS)

If  $D_2$  standard, then NP is satisfiable in a domain  $(D_1, D_2)$  only if  $D_1$  is finite.

## COROLLARY 2

Let  $(D_1, D_2)$  be a domain where  $|D_1| = \aleph_1$  and  $D_2 = \mathcal{P}(C)$ , where  $C \subseteq D_1$  has size  $\aleph_0$ . Then the following are equivalent:

- 1 NP is satisfiable in  $(D_1, D_2)$ ;
- 2 CH is true.

## REMARK

Note the role played by  $\nu$  in both proofs as an *indexing* device.

*Invariance under permutation* was first identified by Tarski as a criterion demarcating logical notions, on the idea that such notions are independent of the subject matter.

If  $\pi$  is a permutation of  $D_1$ , it can be lifted to properties and relations:

$$\pi[R] = \{ \langle \pi(x_1), \dots, \pi(x_n) \rangle : R(x_1, \dots, x_n) \}.$$

A predicate  $R$  is invariant iff  $\pi[R] = R$ .

The following are all invariant:

- One-place predicates:  $\emptyset, D$ ;
- Two place predicates:  $\emptyset, D^2, =, \neq$ ;
- Predicates *definable* (in FOL, infinitary logic, etc.) from invariant predicates.

Conversely, invariant notions are all definable in a possibly higher-order or infinitary language (McGee).

In the modern generalized conception, a *quantifier*  $Q$  is a collection of subsets of the first-order domain, for instance:

$$\begin{aligned}\exists &= \{X \in D_2 : X \neq \emptyset\}; \\ Q_1 &= \{X \in D_2 : |X| \geq \aleph_1\}.\end{aligned}$$

A quantifier  $Q$  is *invariant* under permutations iff for every  $\pi$ :

$$X \in Q \iff \pi[X] \in Q.$$

$\exists$ ,  $\forall$ , as well cardinality quantifiers such as “there are exactly  $k$ ,” “there are infinitely many,” and so on, are all invariant in this sense.

Although it is not clear whether permutation invariance is a sufficient demarcating criterion for logical notions (it appears to be too liberal), it is generally agreed that it provides a necessary condition.

In spite of the accompanying mystique, logical invariance has only been applied to abstraction in a limited way.

Kit Fine uses invariance to constrain the equivalences mentioned in abstraction principles to be *coarse* enough to ensure satisfiability of the principle in which they occur.

There are, *prima facie*, three different way in which invariance can be brought to bear on abstraction:

- Invariance of the equivalence *relation*  $R$ ;
- Invariance of the *operator*  $f_R$ ;
- Invariance of the abstraction *principle*.

We take up each one of them in turn.

## DEFINITION

Let  $R$  be an equivalence relation on  $D_2$ , and assume  $D_2$  is  $p$ -closed.

- $R$  is *internally invariant* if, for any permutation  $\pi$ ,  $R(X, Y)$  if and only if  $R(\pi[X], \pi[Y])$ .
- $R$  is *doubly invariant* if, for any pair  $\pi_1$  and  $\pi_2$  of permutations,  $R(X, Y)$  if and only if  $R(\pi_1[X], \pi_2[Y])$ .
- $R$  is (simply) *invariant* if and only if  $R(X, \pi[X])$  holds for any permutation  $\pi$ .

## PROPOSITION

Let  $R$  be an equivalence relation on  $D_2$ , which is  $p$ -closed.

- If  $R$  is simply invariant then it is internally invariant.
- $R$  is doubly invariant if and only if it is simply invariant.

*Simple invariance* is the strongest notion of invariance for  $R$ , and a very plausible necessary condition on  $\text{Ab}_R$ . The equinumerosity relation  $\approx$  is simply invariant.

## DEFINITION

Let  $f_R$  be an abstraction operator. Then:

- For any permutation  $\pi$  of  $D_1$ , the function  $f_R^\pi$ , is the set-theoretic image of  $f_R$  under  $\pi$ :

$$f_R^\pi = \{ \langle \pi[X], \pi(x) \rangle : \langle X, x \rangle \in f_R \}.$$

- $f_R$  is *objectually invariant* if  $f_R$  is invariant as a set-theoretic entity: i.e., if and only if  $f_R^\pi = f_R$  for any  $\pi$ .

## PROPOSITION

No function  $f_\approx$  satisfying HP is objectually invariant (on any  $p$ -closed domain).

## REMARK

Objectual invariance is the notion that speaks to the character of abstraction as a *logical operation*.

# SIMPLE INVARIANCE AND OBJECTUAL INVARIANCE

## THEOREM

Let  $f_R$  be an abstraction operator and suppose  $D_2$  is  $p$ -closed,  $|D_1| > 1$ , and  $R$  is simply invariant. Then  $f_R$  is *not* objectually invariant.

## PROOF.

Let  $X \in D_2$  where  $f_R(X) = a$ . Since  $|D_1| > 1$ , pick any  $b \neq a$  and let  $\pi$  be a permutation such that  $\pi(a) = b$ . Since  $R$  is simply invariant,  $R(X, \pi[X])$ , so that  $f_R(X) = f_R(\pi[X]) = a$ . Therefore

$$\langle \pi[X], a \rangle \in f_R,$$

whence  $\langle \pi[X], \pi(a) \rangle \notin f_R$  (because  $f_R$  is a function), so that  $f_R \neq f_R^\pi$  and  $f_R$  is not objectually invariant.  $\square$

## REMARK

Objectual invariance is quite rare and mostly incompatible with simple invariance.

# THE INVARIANCE OF $Ab_R$

## DEFINITION

An abstraction principle  $Ab_R$  is *contextually invariant* if and only if, for any function  $f_R : D_2 \rightarrow D_1$  and permutation  $\pi, f_R^\pi$  satisfies  $Ab_R$  whenever  $f_R$  does.

## REMARK

If every  $f_R$  satisfying  $Ab_R$  is objectually invariant, then  $Ab_R$  is contextually invariant.

## PROPOSITION

If  $R$  is internally invariant (and  $D_2$  is  $p$ -closed) then  $Ab_R$  is contextually invariant.

## COROLLARY

Both HP and NP are contextually invariant (over  $p$ -closed domains).

## PROOF.

Both  $\approx$  and  $\sim$  are internally invariant. □

- In spite of the surrounding mystique, abstraction operators are nothing more than assignments of “representatives” in  $D_1$  to equivalence classes over  $D_2$ .
- These are not representatives in the usual sense, because they are not themselves members of the corresponding equivalence classes.
- As a result, a crucial aspect of their function is their *indexing* role, which allows us to “count” equivalence classes by elements of  $D_1$ .
- Thus, abstraction principles enforce a correlation between the size of  $D_1$  and that of  $D_2$ .
- The more fine-grained the relation  $R$ , the greater the *inflationary* thrust exerted by  $\text{Ab}_R$  on the size of  $D_1$  (as in the case of HP), and the *deflationary* thrust on the size of  $D_2$  (as in the case of NP).

If we accept the instrumentalist view of abstraction, then it comes as no surprise that abstraction principles mostly fail to meet the test of logical invariance.

We have three competing notions of invariance for abstraction principles:

- 1 Simple invariance of  $R$ ;
- 2 Objectual invariance of  $f_R$ ;
- 3 Contextual invariance of  $\text{Ab}_R$ .

Simple invariance is a necessary condition for abstraction principles, providing an answer to the “proliferation problem” of ruling out spurious principles, but by itself it is simply *not germane* to the issue of abstraction as a logical operation.

Objectual invariance *does* speak to the issue of the nature of abstraction, but is mostly incompatible with simple invariance.

Contextual invariance turns out to be too weak a notion, weaker than the internal invariance of  $R$ .

# THREE KINDS OF DEMARCATION PROBLEMS

Proponents of abstraction principles as vehicles for the delivery of abstract entities have historically faced three sorts of objections:

**The Bad Company Objection.** While some principles such as HP are essential for the development of arithmetic, it's not clear by which criterion they can be demarcated from their inconsistent brethren.

**The Embarrassment of Riches Objection.** Even if we rule out inconsistent principles by *fiat*, we still have consistent but incompatible principles, such as HP and NP.

**The Caesar Problem.** Abstraction principles do not fix the identity conditions of the objects delivered as values of the abstraction operator except relatively to objects that are themselves *abstracta*.

# THE BAD COMPANY OBJECTION

The attention given to the Bad Company Objection results from the failure to pay due attention to the possibility of *non-standard models* for abstraction principles. This possibility corresponds with the first option for restoring consistency to BLV, i.e., making the assignment *partial*.

Investigations into consistent fragments of *Grundgesetze* (e.g.,  $\Delta_1^1$ -CA) can be viewed as concerning special kinds of non-standard models (satisfying certain closure conditions).

When non-standard models enter the picture, we can see that all abstraction are satisfiable, since they are all *coarsenings* of  $Ab_=$ , i.e., Frege's Basic Law V. We turn the quest for consistency on its head: instead of looking for consistent abstraction principles, we consider what constraints they impose on the models that satisfy them.

Inconsistencies derive from abstraction principles only in conjunction with particular instances of second order comprehension (and neither one is inconsistent by itself). For instance, closure of conceptual space under (impredicative) definability, is incompatible with  $Ab_=$ .

Even when the Bad Company Objection is discarded, a problem arises because of individually consistent but mutually incompatible abstraction principles. The main example is given by HP and NP.

There is a delicate balance between the sizes of the first- and second-order domain when they are tied together by an abstraction principle. In some cases the balance tilts towards the *inflationary* thrust on the first-order domain and sometimes towards the *deflationary* pressure on the second-order domain. Different principles can tilt different ways and thus be incompatible.

But this is just a mathematical fact, not a problem — as it can be easily seen once Abstraction Principles are correctly regarded as *extra-logical* principles (witness their failure to be logically invariant) introduced for very specific mathematical purposes.

The introduction of natural numbers via HP gives rise to the “Caesar Problem,” i.e., the fact that HP only gives us enough information to settle the truth-value of identities in which both terms are abstracts, but says nothing about identities involving an abstract and a term of a different kind:

The number of the planets = Julius Caesar

This account of the natural numbers would then seem to be *incomplete*, and similar worries arise for the other abstraction principles.

In fact, *nothing* prevents the number of the planets from being identical to Julius Caesar. Indeed they will be equal in some models and distinct in other ones. Nothing much is to be made of this fact, for HP is silent about it.

According to the present view, abstraction principles are extra-logical tools devised to accomplish two distinct, but specifically mathematical, tasks:

- 1 The selection of first-order representatives for second-order equivalence classes.
- 2 The imposition of cardinality constraints on the relative sizes of the first-order and the second-order domain.

The status of abstraction principles as extra-logical tools is consistent with the failure of their logical invariance. Similarly, any worries about the ontological status of the special objects they deliver (i.e., *abstracta*) disappear, as anything — anything *at all* — can play the role of these *abstracta*, as long as the choice respects the equivalence relation.

Numbers, as *abstracta* delivered by HP, are no longer regarded as logical objects on this view. “Number” is not a logical notion (although *cardinality* might well be).