REFLECTIONS ON REFLECTION FOR INTUITIONISTIC SET THEORIES

Abstract. We study the well-known reflection principle of Zermelo–Fraenkel set theory in the context of intuitionistic Zermelo–Fraenkel set theory IZF. We show that the reflection principle is equivalent to Aczel's relation reflection scheme RRS. As applications, we give a new proof that relativized dependent choice is equivalent to the conjunction of the relation reflection scheme and dependent choice and we present an intuitionistic version of Feferman's ZFC/S, a conservative extension of ZFC which is useful as a foundation for category theory.

1. Introduction

The basic form of the reflection principle for Zermelo–Fraenkel set theory ZF is the following.

Theorem 1.1. *Let $\varphi(x_1, \ldots, x_n)$ be a formula in the language of set theory with free variables as indicated. Then* ZF *proves*

$$\forall M.\ \exists S \supseteq M.\ \forall x_1, \ldots, x_n \in S.\ \varphi(x_1, \ldots, x_n) \Leftrightarrow \varphi^S(x_1, \ldots, x_n),$$

where φ^S is the S-relativization of φ, obtained by substituting any occurrence of "$\forall x$" and "$\exists x$" by "$\forall x \in S$" and "$\exists x \in S$". Furthermore, the resulting set S may be supposed to be transitive, to be closed under subsets or even to be a stage V_α of the cumulative hierarchy; and given not a single formula φ but a finite list $\varphi_1, \ldots, \varphi_s$ of formulas, we may suppose that S reflects all of them.

The reflection principle is important both for philosophical and for practical reasons: Philosophically, it tells us that truth of any formula can already be checked in an initial segment of the universe. Practically, it allows to transfer results obtained for a restricted class of objects to all such objects. For instance, if we manage to verify (the S-relativization of) some group-theoretic statement for all groups contained in an arbitrary set X, then we may deduce by the reflection principle that the statement holds for all groups in the universe.

This observation has been used by Feferman to construct ZFC/S ("ZFC with smallness"), a conservative extension of ZFC which provides a useful foundation of category theory [4]. This system extends ZFC by the addition of a new constant symbol $\mathbb{S}$ together with axioms stating that $\mathbb{S}$ is transitive, closed under subsets and reflective with respect to every formula $\varphi(x_1, \ldots, x_n)$ of the original language:

$$\forall x_1, \ldots, x_n \in \mathbb{S}.\ \varphi(x_1, \ldots, x_n) \Leftrightarrow \varphi^{\mathbb{S}}(x_1, \ldots, x_n). \qquad (\star)$$

The system ZFC/S is conservative over ZFC because any given proof in ZFC/S uses only a finite number of instances of the axioms $(\star)$, whereby the reflection principle can be used to yield an honest set S which validates just the same equivalences and can hence be used in place of $\mathbb{S}$.

Elements of $\mathbb{S}$ are deemed "small", so that $\mathbb{S}$ is the set of all small sets. The system ZFC/S is useful as a foundation for category theory because it supports, without requiring new set-theoretical commitments such as the existence of Grothendieck universes, a native treatment of large structures. For instance, the category of all small sets can be formed entirely within ZFC/S, without resorting to classes. We invite the reader wanting to learn more about the merits of ZFC/S to study the survey by Shulman [8, Section 11].

To our naive eyes, the passage from ZFC to ZFC/S looked sufficiently innocent so that we set out to create an intuitionistic version of ZFC/S: To our mind, size issues were entirely separate from issues of constructivity, and we hence opined that they should be dealt with separately. Such a separation would not only lead to improved mental hygiene and better understanding, but would also allow us to use the benefits of ZFC/S in situations where the law of excluded middle and the axiom of choice are not available, such as in realizability semantics, sheaf semantics or quite generally topos semantics.

We expected this modification to be entirely straightforward. However, the situation turned out to be more subtle and we failed in our original goal of verifying the reflection principle for intuitionistic Zermelo–Fraenkel set theory IZF. The situation for CZF, the predicative subsystem of IZF commonly heralded as the largest common denominator of all flavors of constructive set theory, is even more mysterious.

However, we succeeded in verifying the reflection principle for a slight extension of IZF:

Theorem 1.2. *The reflection principle is equivalent, over* IZF*, to Aczel's relation reflection scheme* RRS*.*

Aczel's relation reflection scheme will be reviewed below. It is validated not only by ZFC, but also by ZF, and furthermore by all known models of IZF, hence might be regarded as not entirely unconstructive, even though it's conjectured to be independent of IZF. As a result, the question whether the reflection principle holds for IZF remains open (though conjectured to be false), and for the stronger system IZF + RRS we can give a variant "with smallness" which can serve as a set-theoretic foundation for category theory.

Conjecture 1.3. The system IZF does not prove the reflection principle.

Definition 1.4. The system (IZF + RRS)/S is obtained from IZF + RRS by adding a constant symbol $\mathbb{S}$ together with axioms stating that $\mathbb{S}$ is transitive, closed under subsets and reflective for all formulas of the original language (that is, adding one instance of ($\star$) for each formula of IZF).

Corollary 1.5. *The system* (IZF + RRS)/S *is conservative over* IZF + RRS*.*

This note is organized as follows. Section 2 reviews the classical proof of the reflection principle in the context of ZF set theory. Section **??** reviews Aczel's relation reflection scheme and presents our main result. We conclude in Section 4 and Section **??** with two short applications.

Acknowledgments. XXX

2. Review of the classical reflection principle

A basic proof of the reflection principle for ZF runs as follows. Our proof of the reflection principle for IZF + RRS will follow the same outline.

Lemma 2.1. *Let $\varphi(u_1, \ldots, u_n, x)$ be a formula in the language of set theory with free variables as indicated. Then* ZF *proves*

$$\forall M.\ \exists S \supseteq M.\ \forall u_1, \ldots, u_n \in S.\\ (\exists x.\ \varphi(u_1, \ldots, u_n, x)) \Longrightarrow (\exists x \in S.\ \varphi(u_1, \ldots, u_n, x)).$$

Furthermore: (1) We may suppose that S is transitive. (2) We may suppose that S is closed under subsets. (3) Given a finite list $\varphi_1, \ldots, \varphi_s$ of formulas instead of the single formula φ, we may suppose that S bounds all of the formulas φ_i.

Proof. Given a class X (as commonly understood as the comprehension of a formula), we denote by $X^\sim$ its subclass $\{x \in X \mid \forall y \in X.\ \mathrm{rank}(x) \leq \mathrm{rank}(y)\}$, where the rank function refers to the stage in the cumulative hierarchy. The two fundamental properties of this construction are: This subclass is equal to a set,[1] and it is inhabited if and only if X is.

Starting with $S_0 := M$, we construct S_{k+1} from S_k as the union

$$S_{k+1} := S_k \cup \bigcup_{u_1, \ldots, u_n \in S_k} \{x \mid \varphi(u_1, \ldots, u_n, x)\}^\sim.$$

It is then easy to check that $S := \bigcup_{k \in \mathbb{N}} S_k$ is a set with the required property.

For addendum (1), we redeclare $S := \bigcup_{k \in \mathbb{N}} S_k^t$, where S_k^t is the transitive closure of S_k.

To ensure both addendum (1) and addendum (2), we redeclare $S := \bigcup_{k \in \mathbb{N}} P^\omega(S'_k)$, where $P^\omega(X) := \bigcup_{\ell \in \mathbb{N}} P^{(\ell)}(X)$ is the union of iterated powersets.

For addendum (3), we change the definition of S_{k+1} to include one summand for each formula φ_i. $\square$

The proof of Lemma 2.1 is mostly constructive; however, there is one issue with nontrivial ramifications: While IZF does show that the subclass $X^\sim$ of a given class X is a set, it does not verify that $X^\sim$ is inhabited if X is. This would amount to the constructive taboo that any inhabited set contains a rank-minimal element; and in fact, by a result of Friedman and Scedrov [**XXX**], no definable substitute for $X^\sim$ exists. The remedy presented in Section 3 will construct $X^\sim$ in a non-unique fashion and then deal with the resulting fallout that taking the union requires additional care.

Theorem 2.2. *Let $\varphi(x_1, \ldots, x_n)$ be a formula in the language of set theory with free variables as indicated. Then* ZF *proves*

$$\forall M.\ \exists S \supseteq M.\ \forall x_1, \ldots, x_n \in S.\ \varphi(x_1, \ldots, x_n) \Leftrightarrow \varphi^S(x_1, \ldots, x_n),$$

where φ^S is the S-relativization of φ. Furthermore, the resulting set S may be supposed to be transitive and to be closed under subsets; and given not a single formula φ but a finite list $\varphi_1, \ldots, \varphi_s$ of formulas, we may suppose that S reflects all of them.

Proof. Let a set M be given. We obtain S by applying Lemma 2.1 to the list of all subformulas of φ. That the resulting set S has the required property can be checked by an induction on the structure of φ. The cases "$=$", "$\in$", "$\top$", "$\bot$", "$\wedge$", "$\vee$", "$\Rightarrow$"

---

[1] If X is empty, then this claim is trivial; if X is inhabited by some element x_0, then $X^\sim$ can be obtained using separation from $V_{\mathrm{rank}(x_0)+1}$; and in fact, by an argument using the set of truth values and unbounded separation, the claim can also be proven in IZF.

follow trivially from the induction hypothesis. The case "$\forall$" does not need to be treated since we may assume without loss of generality that all universal quantifiers in φ have been rewritten as "$\neg\exists\neg$".

The remaining case is where φ is of the form $\varphi \equiv (\exists x.\ \psi(u_1, \ldots, u_m, x))$. In this case, the claim is that

$$\forall u_1, \ldots, u_m \in S.\ (\exists x.\ \psi(u_1, \ldots, u_m, x)) \Longleftrightarrow (\exists x \in S.\ \psi^S(u_1, \ldots, u_m, x)).$$

This claim follows by the property of S guaranteed by Lemma 2.1 and by the induction hypothesis concerning the subformula $\psi^S(u_1, \ldots, u_m, x)$. $\square$

The proof of Theorem 2.2 is constructive with the sole exception of treating the case of universal quantifiers by appealing to the law of excluded middle. For the constructive proof in Section 3, we will instead appeal to a strengthened version of Lemma 2.1.

3. Constructivizing the reflection principle

Our constructive rendition of the reflection principle will require Aczel's relation reflection scheme RRS displayed in Definition 3.1. This result is the best possible, as we verify in Theorem 3.6 that conversely the reflection principle entails the relation reflection scheme. The relation reflection scheme first surfaced in the theory of coinductive definitions of classes and enjoys substantial stability properties, as it passes from the meta theory to XXX[all kinds of] models. Background on RRS can be found in Aczel's original article introducing it [1].

Definition 3.1 (RRS, Aczel's relation reflection scheme)**.** Let X and $R \subseteq X \times X$ be classes. Let $A \subseteq X$ be a subset. If $\forall x \in X.\ \exists y \in X.\ \langle x, y\rangle \in R$, then there is a set B such that $A \subseteq B \subseteq X$ and such that $\forall x \in B.\ \exists y \in B.\ \langle x, y\rangle \in R$.

For our purposes, Palmgren's multivalued dependent choice is slightly more convenient to work with than RRS. It is equivalent, over CZF and a fortiori over IZF, to RRS [7].

Definition 3.2 (MDC, Palmgren's multivalued dependent choice)**.** Let X and $R \subseteq X \times X$ be classes. Let $A \subseteq X$ be a subset. If $\forall x \in X.\ \exists y \in X.\ \langle x, y\rangle \in R$, then there is a function $f : \mathbb{N} \to P(X)$ (the class of all subsets of X) such that $A \subseteq f(0)$ and such that $\forall x \in f(k).\ \exists y \in f(k+1).\ \langle x, y\rangle \in R$ for every number k.

Even though superficially similar, the following lemma is *not* yet a constructivization of Lemma 2.1; these two lemmas differ in the set from which $u_1, \ldots, u_n$ are drawn.

Lemma 3.3. *Let $\varphi(u_1, \ldots, u_n, x)$ be a formula in the language of set theory with free variables as indicated. Then* IZF *proves*

$$\begin{aligned}&\forall H.\ \exists H' \supseteq H.\ \forall u_1, \ldots, u_n \in H.\\ &\qquad (\exists x.\ \varphi(u_1, \ldots, u_n, x)) \implies (\exists x \in H'.\ \varphi(u_1, \ldots, u_n, x)) \quad \wedge\\ &\qquad (\forall x.\ \varphi(u_1, \ldots, u_n, x)) \impliedby (\forall x \in H'.\ \varphi(u_1, \ldots, u_n, x)).\end{aligned}$$

Furthermore: (1) We may suppose that H' is transitive. (2) We may suppose that H' is closed under subsets. (3) Given a finite list $\varphi_1, \ldots, \varphi_s$ of formulas instead of the single formula φ, we may suppose that H' has the displayed property for each of the formulas φ_i.

Proof. Let $\Omega := P(\{0\})$ be the set of truth values. For given elements $u_1, \ldots, u_n \in H$, we have

$$\begin{array}{ll} \forall a \in \{a \in \{0\} \mid \exists x.\ \varphi(u_1, \ldots, u_n, x)\}. & \exists x.\ \varphi(u_1, \ldots, u_n, x) \quad \text{and} \\ \forall p \in \{p \in \Omega \mid \exists x.\ (0 \in p \Leftrightarrow \varphi(u_1, \ldots, u_n, x))\}. & \exists x.\ (0 \in p \Leftrightarrow \varphi(u_1, \ldots, u_n, x)). \end{array}$$

Hence, by collection, there are sets C and D such that

$$\begin{array}{lcl} (\exists x.\ \varphi(u_1, \ldots, u_n, x)) & \Longrightarrow & (\exists x \in C.\ \varphi(u_1, \ldots, u_n, x)) \quad \text{and} \\ (\forall x.\ \varphi(u_1, \ldots, u_n, x)) & \Longleftarrow & (\forall x \in D.\ \varphi(u_1, \ldots, u_n, x)). \end{array}$$

The union $C \cup D$ satisfies both of these conditions at once.

Applying collection again, there is a set X such that for any $u_1, \ldots, u_n \in H$ there exists a set $E \in X$ such that

$$\begin{array}{lcl} (\exists x.\ \varphi(u_1, \ldots, u_n, x)) & \Longrightarrow & (\exists x \in E.\ \varphi(u_1, \ldots, u_n, x)) \quad \text{and} \\ (\forall x.\ \varphi(u_1, \ldots, u_n, x)) & \Longleftarrow & (\forall x \in E.\ \varphi(u_1, \ldots, u_n, x)). \end{array}$$

Hence the set $H' := M \cup \bigcup X$ has the required property.

To ensure that H' is transitive and closed under subsets, we simply pass first to its transitive closure and then compute the union of all its finitely-iterated powersets.

In order to accommodate more than a single formula φ, we add one summand in the definition of H' for each formula φ_i. $\square$

Lemma 3.4. *Let $\varphi(u_1, \ldots, u_n, x)$ be a formula in the language of set theory with free variables as indicated. Then* IZF + RRS *proves*

$$\begin{array}{l} \forall M.\ \exists S \supseteq M.\ \forall u_1, \ldots, u_n \in S. \\ \qquad \begin{array}{lcl} (\exists x.\ \varphi(u_1, \ldots, u_n, x)) & \Longrightarrow & (\exists x \in S.\ \varphi(u_1, \ldots, u_n, x)) \quad \wedge \\ (\forall x.\ \varphi(u_1, \ldots, u_n, x)) & \Longleftarrow & (\forall x \in S.\ \varphi(u_1, \ldots, u_n, x)). \end{array} \end{array}$$

Furthermore: (1) We may suppose that S is transitive. (2) We may suppose that S is closed under subsets. (3) Given a finite list $\varphi_1, \ldots, \varphi_s$ of formulas instead of the single formula φ, we may suppose that S bounds all of the formulas φ_i.

Proof. By MDC and Lemma 3.3, there is a function $f : \mathbb{N} \to P(V)$, where V is the universe, such that $M \in f(0)$ and such that for any number k and any set $H \in f(k)$, there is a set $H' \in f(k+1)$ such that H and H' are related as in the conclusion of Lemma 3.3.

We set $S := \bigcup \big(\bigcup_{k \in \mathbb{N}} f(k) \big)$. This set S has the required property; to verify this, it is useful to observe that given $u_1, \ldots, u_n \in S$, there exists a common number k such that $u_1, \ldots, u_n \in H$ for some $H \in f(k)$.

In order to ensure addendum (1) and (2), we apply MDC in a slightly different way such that for any number k and any set $H \in f(k)$, there is a set $H' \in f(k+1)$ such that H and H' are related as in the conclusion of Lemma 3.3 and such that H' is transitive and closed under subsets. Even though it cannot be expected that for any number k, any set $H \in f(k)$ will be transitive and closed under subsets, the union S will. $\square$

Theorem 3.5. *Let $\varphi(x_1, \ldots, x_n)$ be a formula in the language of set theory with free variables as indicated. Then* IZF + RRS *proves*

$$\forall M.\ \exists S \supseteq M.\ \forall x_1, \ldots, x_n \in S.\ \varphi(x_1, \ldots, x_n) \Leftrightarrow \varphi^S(x_1, \ldots, x_n),$$

where φ^S is the S-relativization of φ. Furthermore, the resulting set S may be supposed to be transitive and to be closed under subsets; and given not a single *formula φ but a finite list $\varphi_1, \ldots, \varphi_s$ of formulas, we may suppose that S reflects all of them.*

Proof. The proof of Theorem 2.2 carries over. The only difference is that instead of Lemma 2.1, Lemma 3.4 has to be used, and that the case for the universal quantifier has to be treated just as the case for the existential quantifier is. $\square$

Theorem 3.6. *Over* IZF*, each instance of Aczel's relation reflection scheme* RRS *can be deduced from suitable instances of the assumption that, given a finite list of formulas, for every set M there is a set $S \supseteq M$ reflecting the given formulas.*

Proof. We verify the equivalent MDC instead of RRS. Let X and $R \subseteq X \times X$ be classes. Let $A \subseteq X$ be a subset and suppose $\forall x \in X.\ \exists y \in X.\ \langle x, y\rangle \in R$.

By assumption, there is a set $S \supseteq A$ which reflects the three formulas "$x \in X$", "$\langle x, y\rangle \in R$" and "$\exists y \in X.\ \langle x, y\rangle \in R$".

By separation, the class $S \cap X$ is a set, and hence we may set $f(k) := S \cap X$ for all numbers k. Then $A \subseteq f(0)$; any given $x \in f(k)$, by assumption and the reflecting property of S there exists $y \in S$ such that $y \in X$ (hence $y \in f(k+1)$) and $\langle x, y\rangle \in R$. $\square$

4. A new proof of RDC = RRS + DC

When he introduced RRS, Aczel proved that over CZF without subset collection, relative dependent choice RDC is equivalent to the conjunction of RRS and dependent choice DC [1, Theorem 2.4]. Using reflection, we can provide a new proof of this fact, although over the much stronger base theory IZF instead of CZF$^-$. The idea is that reflection allows to reduce RDC to DC.

Definition 4.1 (DC, dependent choice)**.** Let X and $R \subseteq X \times X$ be sets. Let $x_0 \in X$. If $\forall x \in X.\ \exists y \in X.\ \langle x, y\rangle \in R$, then there is a function $f : \mathbb{N} \to X$ such that $f(0) = x_0$ and such that $\langle f(k), f(k+1)\rangle \in R$ for all numbers k.

Definition 4.2 (RDC, relative dependent choice)**.** Let X and $R \subseteq X \times X$ be classes. Let $x_0 \in X$. If $\forall x \in X.\ \exists y \in X.\ \langle x, y\rangle \in R$, then there is a function $f : \mathbb{N} \to X$ such that $f(0) = x_0$ and such that $\langle f(k), f(k+1)\rangle \in R$ for all numbers k.

Proposition 4.3. *Over* IZF*,* RDC *is equivalent to* RRS + DC*.*

Proof. It's clear that RDC implies RRS + DC.

Conversely, assume RRS and DC. In order to verify RDC, let classes X and $R \subseteq X \times X$ be given. Let $x_0 \in X$ and assume $\forall x \in X.\ \exists y \in X.\ \langle x, y\rangle \in R$.

By Theorem 3.5, reflection is available; hence there is a set $S \supseteq \{x_0\}$ which reflects the three formulas "$x \in X$", "$\langle x, y\rangle \in R$" and "$\forall x.\ x \in X \Rightarrow \exists y.\ (y \in X \wedge \langle x, y\rangle \in R)$".

Hence $\forall x \in X \cap S.\ \exists y \in X \cap S.\ \langle x, y\rangle \in R$. By DC, there is a choice function $f : \mathbb{N} \to X \cap S$ such that $f(0) = x_0$ and $\langle f(k), f(k+1)\rangle \in R$ for all numbers k. This is a function of the kind required by RDC. $\square$

5. An intuitionistic version of Feferman's ZFC/S

XXX: repeat definition and conservativity statement?

The system (IZF + RRS)/S is also interesting from the point of view of topos theory. Any topos supports an internal language which can be used to reason about the objects and morphisms of the topos in a naive element-based language, allowing us to pretend that the objects are plain sets (or types) and that the morphisms are plain maps between those sets ([2, Chapter 6], [3, Section 1.3], [5, Chapter 14], [6, Chapter VI]).

Given a topos $\mathcal{E}$ and a formula φ in its internal language, we write "$\mathcal{E} \models \varphi$" to mean that φ holds in $\mathcal{E}$. As a special case, truth in the topos Set, the category of all sets, coincides with truth in the background theory; symbolically: $\mathrm{Set} \models \varphi$ iff φ.

However, in the context of (IZF or) ZFC, it is difficult to make this claim precise. Because in ZFC the category Set of all sets can not be coded as a set, ZFC cannot define truth in Set. We must therefore resort to a meta theory in order to express this claim, for instance by stating that primitive recursive arithmetic PRA proves that for any formula φ of ZFC, ZFC proves that $(\mathrm{Set} \models \varphi) \Leftrightarrow \varphi$, where "$\mathrm{Set} \models \varphi$" is to be unrolled by PRA.

An alternative is offered by ZFC+I, ZFC plus the existence of a strongly inaccessible cardinal. In this system, there is a Grothendieck universe U, we can form the category Set_U of all sets of U as an honest set, can define truth in Set_U and prove, within the system, that for any formula φ, $(\mathrm{Set}_U \models \varphi) \Leftrightarrow (U \models \varphi)$. This even holds for formulas of the full infinitary language of toposes, which allows for infinite disjunctions and infinite conjunctions; this could not be mimicked by resorting to PRA as indicated above.

However, truth in a Grothendieck universe U need not be related to actual truth. A solution is provided by ZFC/S and by (IZF + RRS)/S. In these systems, we can form the topos $\mathrm{Set}_{\mathbb{S}}$, define truth in it, prove for all formulas φ that $(\mathrm{Set}_{\mathbb{S}} \models \varphi) \Leftrightarrow (\mathbb{S} \models \varphi)$; and reflection for $\mathbb{S}$ ensures that for each (external, standard) formula φ, the system proves "$(\mathrm{Set}_{\mathbb{S}} \models \varphi) \Leftrightarrow \varphi$".

References

[1] P. Aczel. "The relation reflection scheme". In: *MLQ Math. Log. Q.* 54.1 (2008), pp. 5–11.

[2] F. Borceux. *Handbook of Categorical Algebra: Volume 3, Sheaf Theory.* Encyclopedia Math. Appl. Cambridge University Press, 1994.

[3] O. Caramello. *Theories, Sites, Toposes: Relating and studying mathematical theories through topos-theoretic 'bridges'.* Oxford University Press, 2018.

[4] S. Feferman. "Set-theoretical foundations of category theory". In: *Reports of the Midwest Category Seminar III.* Vol. 106. Lecture Notes in Math. Springer, 1969, pp. 201–247.

[5] R. Goldblatt. *Topoi: The Categorical Analysis of Logic.* Vol. 98. Stud. Logic Found. Math. Elsevier, 1984.

[6] S. Mac Lane and I. Moerdijk. *Sheaves in Geometry and Logic: a First Introduction to Topos Theory.* Universitext. Springer, 1992.

[7] E. Palmgren. "A small variation of Aczel's relation reflection scheme". 2007. URL: `http://www2.math.uu.se/~palmgren/MVDC.pdf`.

[8] M. Shulman. "Set theory for category theory". 2008. URL: `https://arxiv.org/abs/0810.1279`.