Notre Dame Journal of Formal Logic
Volume 32, Number 4, Fall 1991

Inequality in Constructive Mathematics

WIM RUITENBURG

Abstract We present difference relations as a natural generalization of inequality in constructive mathematics. Differences on a set S are defined as binary relations on all powers S^n simultaneously, satisfying axiom schemas generalizing the ones for inequality. The denial inequality and the apartness relation are special cases of a difference relation. Several theorems in constructive algebra are given that unify and generalize well-known results in constructive algebra previously employing special cases of difference relations. Finally, we discuss extended differences for a set S as collections of relations defined on all powers S^X simultaneously.

Introduction In mathematics the natural generalization of equality is equivalence. A theory with equivalence involves the reflexive, symmetric, and transitive equivalence, and functions and relations respecting this equivalence. In constructive mathematics the same theory with equivalence relations works without difficulty. For inequality the situation is more complicated. There are different versions of constructive inequality that only in classical mathematics are equal to the one standard inequality. Examples are: denial inequality, where $x \neq y$ if and only if it is not true that $x = y$, that is, $\neg x = y$; and tight apartness, whose axiomatization we will present later on. The natural inequality on the set of real numbers $\mathbf{R}$, defined by $r \neq s$ if and only if $|r - s| > 1/n$ for some natural number n, is a tight apartness. Tight apartness and denial inequality are independent; a tight apartness need not be a denial inequality, a denial inequality need not be a tight apartness. We know of no definition of a binary relation on a set S, generalizing both denial inequality and apartness, that allows for a substantial constructive theory of inequality.

There are several theorems in algebra and elsewhere that hold if we use denial inequality as the intended inequality, and that also hold if we use a tight apartness as the intended inequality. Sometimes there may even be a third version of inequality that makes the theorem work. For each of these cases we need a new proof to establish our result. For a uniform treatment of such theorems we

Received April 26, 1990; revised July 12, 1990

present a generalization of the inequalities mentioned above, called a *difference*. Rather than defining a binary relation on a set S, a difference is a collection of binary relations defined on all powers S^n simultaneously. Then for some theorems we only need a difference to establish the conclusion. In Section 2 we present examples of theorems that have generalizations employing differences instead of denial inequality or apartness.

To illustrate why inequality is more troublesome than equality when we generalize to a constructive context, we consider the problem in the context of some first-order language with equality $=$. Besides the logical axiom schemata and rules concerning the logical operators and constants we have for equality the axiom schemas

$$\top \vdash x = x$$

$$x = y \vdash Ax \to Ay,$$

where in the last schema the variables x, y are not bound by a quantifier of A. If $=$ is an equivalence relation, then A is any formula built up from functions and relations that preserve the equivalence. It is well-known that we may restrict Ax to atomic formulas and equations $f = g$. The general case follows from this subcollection. The schemas above work in constructive mathematics as well as in classical mathematics.

From the schemas for equality we derive the obvious axiom schemas for inequality $\neq$ by reversing the entailments:

$$x \neq x \vdash \bot$$

$$Ay \vdash x \neq y \vee Ax,$$

where in the last schema the variables x, y are not bound by a quantifier of A.

The schemas for inequality are just fine in classical mathematics. Unfortunately, the introduction of a disjunction in a rule for a generalized inequality is unacceptable in constructive mathematics. In general, even denial inequality fails to obey the schemas.

To find a way out, suppose that Ax is the equation $f(x) \neq t$, where $f: S \to S$ and $x, t \in S$. Classically that gives

$$f(y) \neq t \vdash x \neq y \vee f(x) \neq t.$$

Then one inequality introduces a disjunction of two inequalities. Repeated application implies that, unless we somehow interfere, we end up with disjunctions of inequalities, a prospect unacceptable in constructive mathematics. The partial solution proposed in this paper is to replace the introduction of disjunctions like

$$x_1 \neq y_1 \vee \cdots \vee x_n \neq y_n$$

by introducing differences among sequences of elements:

$$\langle x_1, \ldots, x_n \rangle \neq \langle y_1, \ldots, y_n \rangle.$$

This seems to be the best that one can hope for without introducing disjunctions, but it requires an extension from a definition of $\neq$ on a set S to a definition of $\neq$ involving all powers S^n. The axiom schema involves functions $\mathbf{f}: S^m \to S^n$ only.

In Section 1 we show that in a first-order context the logical motivation presented above provides us with a natural generalization of the notion of inequality. In Section 2 we demonstrate the necessity and sufficiency of *difference* in elementary algebra. In Section 3 we hint at a more general formulation of *difference*, employing all powers S^X rather than only finite powers S^n.

1 Difference relations We define difference relations and strong extensionality in a way motivated by the discussion in the Introduction, and show that they satisfy the right properties. This presents us with the problem that the original definition, though well-motivated, lacks the elegance of a compact set of axioms. Fortunately, with Propositions 1.5 and 1.6, we are able to reduce the complicated definition below to a set of six axioms for difference, and a simple schema for strong extensionality.

From here on we use boldface letters to represent sequences of elements. Let S be a set, and let Λ be a set of partial functions $\mathbf{f}: S^m \to S^n$ between powers of S. Using partial functions rather than total functions is useful for later when we discuss functions $f: S \to T$ between different sets with difference relations. Then $E(\Lambda)$ denotes the smallest set of partial functions between powers of S that includes Λ, all projections $\pi_i: S^n \to S$, and is closed under composition and products. The set $E = E(\varnothing)$ is called the set of *elementary maps*. So elementary maps $\mathbf{f}: S^m \to S^n$ are such that for all i the coordinate maps $\pi_i \mathbf{f}: S^m \to S$ are projections.

A *difference* on S consists of relations $\neq_n$ on the powers S^n, all usually written $\neq$, satisfying the axiom schemata

(1) $$\langle \mathbf{x}, a\rangle \neq \langle \mathbf{y}, a\rangle \to \mathbf{x} \neq \mathbf{y};$$

(2) $$\mathbf{f}(\mathbf{y}) \neq \mathbf{t} \to \langle \mathbf{x}, \mathbf{f}(\mathbf{x})\rangle \neq \langle \mathbf{y}, \mathbf{t}\rangle,$$

where $a \in S$, $\mathbf{x},\mathbf{y} \in S^m$, $\mathbf{f}: S^m \to S^n \in E$, and $\mathbf{t} \in S^n$. We tacitly assume that $\mathbf{f}(\mathbf{y})$, $\mathbf{f}(\mathbf{x})$, etc. are defined when they occur in formulas. A difference is called *proper* if it satisfies the additional axiom schema

(3) $$\neg(\langle\ \rangle \neq \langle\ \rangle).$$

A set Λ is *strongly extensional* with respect to a difference relation $\neq$ if (2) holds for all $\mathbf{f} \in E(\Lambda)$.

There are two questions that we must answer to justify our definition of difference: Does it provide us with a useful theory; and does it provide us with a natural generalization of the notion of nonequivalence? We start with a quick look at the second question by looking at the complement of difference and at the complement of nonequivalence.

A difference induces relations $\sim$ on the sets S^n defined by $\mathbf{x} \sim \mathbf{y} \leftrightarrow \neg \mathbf{x} \neq \mathbf{y}$. We say $\mathbf{x}$ is *nearby* $\mathbf{y}$ if $\mathbf{x} \sim \mathbf{y}$. Then $\sim$ satisfies the schemas

$$\langle\ \rangle \sim \langle\ \rangle \text{ if } \neq \text{ is proper;}$$

$$\mathbf{x} \sim \mathbf{y} \to \langle \mathbf{x}, a\rangle \sim \langle \mathbf{y}, a\rangle; \text{ and}$$

$$\langle \mathbf{x}, \mathbf{f}(\mathbf{x})\rangle \sim \langle \mathbf{y}, \mathbf{t}\rangle \to \mathbf{f}(\mathbf{y}) \sim \mathbf{t},$$

where $a \in S$, $\mathbf{x},\mathbf{y} \in S^m$, $\mathbf{f}: S^m \to S^n \in E$, and $\mathbf{t} \in S^n$. The relation $\sim$ is symmetric (see Proposition 1.1) and, if $\neq$ is proper, reflexive; but $\sim$ need not be transitive. In Examples 1.14 and 1.15 we present models showing that even in classical mathematics it is possible to have elements x, y, z such that x and y are nearby, y and z are nearby, but $x \neq z$. So differences are essentially more general than the complements of equivalence relations. Nearness is *stable*, that is, $\neg\neg \mathbf{x} \sim \mathbf{y}$ implies $\mathbf{x} \sim \mathbf{y}$.

A difference is an *inequivalence* if its nearness relation is an equivalence relation on each of the powers S^n. From Proposition 1.16 it follows that a difference is an inequivalence if and only if it is proper and its corresponding nearness relation satisfies

(4) $$\mathbf{x} \sim \mathbf{y} \text{ if and only if } \forall i (x_i \sim y_i).$$

A difference is an *inequality* if it satisfies

$$\mathbf{x} \sim \mathbf{y} \text{ if and only if } \neg\neg \mathbf{x} = \mathbf{y}$$

for all n and $\mathbf{x}, \mathbf{y} \in S^n$. Obviously, inequalities are inequivalences.

Many natural examples of difference relations are derived from equivalence relations. One easily verifies that each equivalence relation $\equiv$ induces an inequivalence by

$$\mathbf{x} \neq \mathbf{y} \leftrightarrow \neg \mathbf{x} \equiv \mathbf{y},$$

where $\mathbf{x} \equiv \mathbf{y}$ is short for $\forall i (x_i \equiv y_i)$. The relation $\sim$ is the double negation of $\equiv$. The set Λ of all partial functions that preserve the equivalence is a strongly extensional set. One example is the *empty inequivalence*, where $\equiv$ is the maximum equivalence relation and the underlying set is one single equivalence class. The derived relation $\sim$ is identical to $\equiv$. Another example is the *denial inequality*, where $\equiv$ is the minimum equivalence relation, that is, $\equiv$ is the equality relation $=$. All partial functions respect equality and the maximal equivalence relation. So the set of all partial functions is strongly extensional with respect to the empty inequivalence as well as the denial inequality.

Proposition 1.1 *Differences are symmetric.*

Proof: From $\mathbf{y} \neq \mathbf{t}$ we get $\langle \mathbf{x}, \mathbf{x} \rangle \neq \langle \mathbf{y}, \mathbf{t} \rangle$ for all $\mathbf{x}$. Substitute $\mathbf{t}$ for $\mathbf{x}$ and apply (1) repeatedly to get $\mathbf{t} \neq \mathbf{y}$.

Proposition 1.2 *Let Λ be a strongly extensional set of partial functions. Then for all* $\mathbf{f} \in E(\Lambda)$,

(5) $$\langle \mathbf{f}(\mathbf{x}), \mathbf{z} \rangle \neq \langle \mathbf{f}(\mathbf{y}), \mathbf{w} \rangle \to \langle \mathbf{x}, \mathbf{z} \rangle \neq \langle \mathbf{y}, \mathbf{w} \rangle.$$

Proof: From $\langle \mathbf{f}(\mathbf{x}), \mathbf{z} \rangle \neq \langle \mathbf{f}(\mathbf{y}), \mathbf{w} \rangle$ we get, using (2), $\langle \mathbf{p}, \mathbf{q}, \mathbf{f}(\mathbf{p}), \mathbf{q} \rangle \neq \langle \mathbf{x}, \mathbf{z}, \mathbf{f}(\mathbf{y}), \mathbf{w} \rangle$ for all $\mathbf{p}$ and $\mathbf{q}$ with $\mathbf{f}(\mathbf{p})$ defined. Substituting $\mathbf{y}$ for $\mathbf{p}$ and $\mathbf{w}$ for $\mathbf{q}$ gives us $\langle \mathbf{y}, \mathbf{w}, \mathbf{f}(\mathbf{y}), \mathbf{w} \rangle \neq \langle \mathbf{x}, \mathbf{z}, \mathbf{f}(\mathbf{y}), \mathbf{w} \rangle$. By repeated application of (1) we get $\langle \mathbf{y}, \mathbf{w} \rangle \neq \langle \mathbf{x}, \mathbf{z} \rangle$. So by Proposition 1.1, $\langle \mathbf{x}, \mathbf{z} \rangle \neq \langle \mathbf{y}, \mathbf{w} \rangle$.

Corollary 1.3

(6) $$\langle \mathbf{x}, a, a \rangle \neq \langle \mathbf{y}, b, b \rangle \to \langle \mathbf{x}, a \rangle \neq \langle \mathbf{y}, b \rangle;$$

(7) $$\mathbf{x} \neq \mathbf{y} \to \langle \mathbf{x}, a \rangle \neq \langle \mathbf{y}, b \rangle; \text{ and}$$

(8) $$\langle x_{\pi 1}, \ldots, x_{\pi n} \rangle \neq \langle y_{\pi 1}, \ldots, y_{\pi n} \rangle \to \langle x_1, \ldots, x_n \rangle \neq \langle y_1, \ldots, y_n \rangle,$$

where π is a permutation on $\{1, \ldots, n\}$.

Proposition 1.4

(9) $$\langle \mathbf{x}, a\rangle \neq \langle \mathbf{y}, b\rangle \rightarrow \langle \mathbf{x}, b\rangle \neq \langle \mathbf{y}, a\rangle;$$

(10) $$\langle \mathbf{x}, a\rangle \neq \langle \mathbf{y}, b\rangle \rightarrow \langle \mathbf{x}, a, c\rangle \neq \langle \mathbf{y}, c, b\rangle.$$

Proof: From the assumption of (9) we get $\langle \mathbf{z}, c, \mathbf{z}, c\rangle \neq \langle \mathbf{x}, a, \mathbf{y}, b\rangle$ for all $\mathbf{z}$ and c. Substitute $\mathbf{x}$ for $\mathbf{z}$ and b for c to get $\langle \mathbf{x}, b, \mathbf{x}, b\rangle \neq \langle \mathbf{x}, a, \mathbf{y}, b\rangle$. So by (8) and (1) we have $\langle \mathbf{x}, b\rangle \neq \langle \mathbf{y}, a\rangle$.

The assumption of (10) implies $\langle \mathbf{z}, c, \mathbf{z}, c\rangle \neq \langle \mathbf{x}, a, \mathbf{y}, b\rangle$ for all $\mathbf{z}$. Substitute $\mathbf{x}$ for $\mathbf{z}$ and use (8) and (1) to get $\langle \mathbf{x}, c, c\rangle \neq \langle \mathbf{y}, a, b\rangle$. So by (8) and (9) we have $\langle \mathbf{x}, a, c\rangle \neq \langle \mathbf{y}, c, b\rangle$.

Proposition 1.5 *Let $\neq$ be a relation on the powers S^n of a set S. Then $\neq$ is a difference if and only if the following conditions hold.*

(1) $$\langle \mathbf{x}, a\rangle \neq \langle \mathbf{y}, a\rangle \rightarrow \mathbf{x} \neq \mathbf{y};$$

(6) $$\langle \mathbf{x}, a, a\rangle \neq \langle \mathbf{y}, b, b\rangle \rightarrow \langle \mathbf{x}, a\rangle \neq \langle \mathbf{y}, b\rangle;$$

(7) $$\mathbf{x} \neq \mathbf{y} \rightarrow \langle \mathbf{x}, a\rangle \neq \langle \mathbf{y}, b\rangle;$$

(8) $$\langle x_{\pi 1}, \ldots, x_{\pi n}\rangle \neq \langle y_{\pi 1}, \ldots, y_{\pi n}\rangle \rightarrow \langle x_1, \ldots, x_n\rangle \neq \langle y_1, \ldots, y_n\rangle;$$

(9) $$\langle \mathbf{x}, a\rangle \neq \langle \mathbf{y}, b\rangle \rightarrow \langle \mathbf{x}, b\rangle \neq \langle \mathbf{y}, a\rangle; \text{ and}$$

(10) $$\langle \mathbf{x}, a\rangle \neq \langle \mathbf{y}, b\rangle \rightarrow \langle \mathbf{x}, a, c\rangle \neq \langle \mathbf{y}, c, b\rangle,$$

where (8) *holds for all permutations* π.

Proof: Clearly conditions (1) and (6) through (10) hold for a difference relation. Conversely, suppose we have relations $\neq$ on the powers S^n of a set S satisfying the conditions above. To prove (2), let $\mathbf{f}: S^m \rightarrow S^n$ be an elementary map such that $\mathbf{f}(\mathbf{y}) \neq \mathbf{t}$. The map $\mathbf{f}$ is a sequence of projections $(\pi_{\lambda 1}, \ldots, \pi_{\lambda n})$. So $\langle y_{\lambda 1}, \ldots, y_{\lambda n}\rangle \neq \mathbf{t}$. Repeated application of (8) and (10) yields $\langle y_{\lambda 1}, x_{\lambda 1}, \ldots, y_{\lambda n}, x_{\lambda n}\rangle \neq \langle x_{\lambda 1}, t_1, \ldots, x_{\lambda n}, t_n\rangle$. So by (8), $\langle \mathbf{f}(\mathbf{y}), \mathbf{f}(\mathbf{x})\rangle \neq \langle \mathbf{f}(\mathbf{x}), \mathbf{t}\rangle$. Applying (8) and (9) repeatedly, we get $\langle \mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{x})\rangle \neq \langle \mathbf{f}(\mathbf{y},)\mathbf{t}\rangle$. So by (6), (7), and (8) we get $\langle \mathbf{x}, \mathbf{f}(\mathbf{x})\rangle \neq \langle \mathbf{y}, \mathbf{t}\rangle$.

Proposition 1.5 has two applications. First, it replaces schema (2) by a short sequence of elementary rules. Second, it suggests natural ways for generalizing difference relations. Prime choices are generalizations $\neq$ satisfying the conditions of Proposition 1.5 but with (6) or (10) removed. The structure of Example 1.15.1 satisfies all the conditions of Proposition 1.5, except (6). On domain $S = \mathbf{Z}$, define $\mathbf{x} \neq \mathbf{y}$ by $|x_i - y_i| \geq 2$ for some i. Then $\neq$ is a generalized difference relation satisfying all conditions of Proposition 1.5, except (10).

The definition of strongly extensional sets of functions allows for the possibility that a set need not be strongly extensional even if all its members are. Fortunately, this does not happen. Theorem 1.6 expresses strong extensionality of sets in terms of the individual functions.

Theorem 1.6 *Let $\neq$ be a difference on S and let Λ be a set of partial functions between finite powers on S. Then Λ is a strongly extensional set if and only if each $\mathbf{f} \in \Lambda$ satisfies the schema*

(5) $$\langle \mathbf{f}(\mathbf{x}), \mathbf{z}\rangle \neq \langle \mathbf{f}(\mathbf{y}), \mathbf{w}\rangle \rightarrow \langle \mathbf{x}, \mathbf{z}\rangle \neq \langle \mathbf{y}, \mathbf{w}\rangle.$$

Proof: By Proposition 1.2, (5) follows from the strong extensionality of Λ. Conversely, suppose that (5) holds for all $\mathbf{f} \in \Lambda$. A trivial induction on the complexity of $\mathbf{f}$ shows that (5) holds for all $\mathbf{f} \in E(\Lambda)$. Now suppose $\mathbf{f}(y) \neq \mathbf{t}$ and $\mathbf{f}(\mathbf{x})$ exists, for some $\mathbf{f} \in E(\Lambda)$. Substitution in the schema $\mathbf{h} \neq \mathbf{t} \rightarrow \langle \mathbf{g},\mathbf{g} \rangle \neq \langle \mathbf{h},\mathbf{t} \rangle$ gives $\langle \mathbf{f}(\mathbf{x}),\mathbf{f}(\mathbf{x}) \rangle \neq \langle \mathbf{f}(\mathbf{y}),\mathbf{t} \rangle$. Applying (5) yields $\langle \mathbf{x},\mathbf{f}(\mathbf{x}) \rangle \neq \langle \mathbf{y},\mathbf{t} \rangle$.

By Theorem 1.6 we are justified to define a function $\mathbf{f}$ to be *strongly extensional* if it satisfies the schema (5).

Proposition 1.7 *Constant functions are strongly extensional.*

Proof: Let $\mathbf{f}$ be a constant function with value $\mathbf{a}$. Then $\langle \mathbf{a},\mathbf{z} \rangle \neq \langle \mathbf{a},\mathbf{w} \rangle$ implies $\mathbf{z} \neq \mathbf{w}$, and thus $\langle \mathbf{x},\mathbf{z} \rangle \neq \langle \mathbf{y},\mathbf{w} \rangle$.

By Theorem 1.6 we know that the collection of strongly extensional functions is closed under composition and product. Next we show that the collection is also closed under a natural form of implicit definition. Traditionally, a (partial) function h is implicitly defined by the (partial) functions f and g when $f(x, hy)$ and $g(x, hy)$ exist whenever hx and hy exist; when $f(x, hx) = g(x, hx)$ whenever hx exists; and when $f(x,p) = g(x,p) \wedge f(x,q) = g(x,q)$ implies $p = q$, for all x, p, q. In ring theory, for example, the partial function of multiplicative inverse is implicitly definable from multiplication and the constant 1. We show that functions that are implicitly defined in the way explained below are strongly extensional if the functions used in its construction are.

Let S be a set with difference $\neq$. A partial function $\mathbf{h}$ is *implicitly defined with respect to* $\neq$ if there exist strongly extensional partial functions $\mathbf{f}$ and $\mathbf{g}$ such that $\mathbf{f}(\mathbf{x},\mathbf{h}(\mathbf{y}))$ and $\mathbf{g}(\mathbf{x},\mathbf{h}(\mathbf{y}))$ are defined whenever $\mathbf{h}(\mathbf{x})$ and $\mathbf{h}(\mathbf{y})$ are defined, satisfying

$$\mathbf{f}(\mathbf{x},\mathbf{h}(\mathbf{x})) = \mathbf{g}(\mathbf{x},\mathbf{h}(\mathbf{x})) \text{ whenever } \mathbf{h}(\mathbf{x}) \text{ is defined; and}$$

$$\langle \mathbf{p},\mathbf{z} \rangle \neq \langle \mathbf{q},\mathbf{w} \rangle \rightarrow \langle \mathbf{f}(\mathbf{x},\mathbf{p}),\mathbf{f}(\mathbf{x},\mathbf{q}),\mathbf{z} \rangle \neq \langle \mathbf{g}(\mathbf{x},\mathbf{p}),\mathbf{g}(\mathbf{x},\mathbf{q}),\mathbf{w} \rangle$$

whenever $\mathbf{f}(\mathbf{x},\mathbf{p})$, $\mathbf{f}(\mathbf{x},\mathbf{q})$, $\mathbf{g}(\mathbf{x},\mathbf{p})$, and $\mathbf{g}(\mathbf{x},\mathbf{q})$ are defined.

Proposition 1.8 *Partial functions that are implicitly defined with respect to a difference relation are strongly extensional.*

Proof: Let $\mathbf{h}(\mathbf{x}) = \mathbf{y}$ be implicitly defined with respect to a difference by the equation $\mathbf{f}(\mathbf{x},\mathbf{y}) = \mathbf{g}(\mathbf{x},\mathbf{y})$. Suppose $\langle \mathbf{h}(\mathbf{r}),\mathbf{z} \rangle \neq \langle \mathbf{h}(\mathbf{s}),\mathbf{w} \rangle$. For all $\mathbf{x}$ such that $\mathbf{h}(\mathbf{x})$ is defined we have $\langle \mathbf{f}(\mathbf{x},\mathbf{h}(\mathbf{r})),\mathbf{f}(\mathbf{x},\mathbf{h}(\mathbf{s})),\mathbf{z} \rangle \neq \langle \mathbf{g}(\mathbf{x},\mathbf{h}(\mathbf{r})),\mathbf{g}(\mathbf{x},\mathbf{h}(\mathbf{s})),\mathbf{w} \rangle$. Substitute $\mathbf{x} = \mathbf{r}$. Using $\mathbf{f}(\mathbf{r},\mathbf{h}(\mathbf{r})) = \mathbf{g}(\mathbf{r},\mathbf{h}(\mathbf{r}))$ we get $\langle \mathbf{f}(\mathbf{r},\mathbf{h}(\mathbf{s})),\mathbf{z} \rangle \neq \langle \mathbf{g}(\mathbf{r},\mathbf{h}(\mathbf{s})),\mathbf{w} \rangle$. By (2) we have $\langle \mathbf{s},\mathbf{f}(\mathbf{s},\mathbf{h}(\mathbf{s})),\mathbf{z} \rangle \neq \langle \mathbf{r},\mathbf{g}(\mathbf{r},\mathbf{h}(\mathbf{s})),\mathbf{w} \rangle$. By (8) and (10), $\langle \mathbf{s},\mathbf{f}(\mathbf{s},\mathbf{h}(\mathbf{s})), \mathbf{g}(\mathbf{s},\mathbf{h}(\mathbf{s})),\mathbf{z} \rangle \neq \langle \mathbf{r},\mathbf{g}(\mathbf{s},\mathbf{h}(\mathbf{s})),\mathbf{g}(\mathbf{r},\mathbf{h}(\mathbf{s})),\mathbf{w} \rangle$. Since $\mathbf{f}(\mathbf{s},\mathbf{h}(\mathbf{s})) = \mathbf{g}(\mathbf{s},\mathbf{h}(\mathbf{s}))$ and $\mathbf{g}$ is strongly extensional we have $\langle \mathbf{s},\mathbf{s},\mathbf{h}(\mathbf{s}),\mathbf{z} \rangle \neq \langle \mathbf{r},\mathbf{r},\mathbf{h}(\mathbf{s}),\mathbf{w} \rangle$. So $\langle \mathbf{r},\mathbf{z} \rangle \neq \langle \mathbf{s},\mathbf{w} \rangle$. Thus $\mathbf{h}$ is strongly extensional.

If there exists $\mathbf{x}$ such that $\mathbf{x} \neq \mathbf{x}$, then everything is different from everything in each S^n, as follows from Proposition 1.9 below.

Proposition 1.9 *For all* $\mathbf{x}$, $\mathbf{y}$, *and* $\mathbf{z}$ *we have* $\mathbf{x} \neq \mathbf{x} \rightarrow \mathbf{y} \neq \mathbf{z}$.

Proof: Suppose $\mathbf{x} \neq \mathbf{x}$ for some $\mathbf{x}$. Repeated application of (1) implies $\langle\ \rangle \neq \langle\ \rangle$. Repeated application of (7) then yields $\mathbf{y} \neq \mathbf{z}$ for all $\mathbf{y}$ and $\mathbf{z}$.

So a difference is proper if and only if it is contained in the denial inequality.

The tight apartness on the real numbers **R** was introduced by Brouwer [2] and subsequently axiomatized by Heyting in 1925 (see [6]). The following is a new way of defining apartness relations: By employing the notion of difference relation. An *apartness* is a proper difference relation satisfying the extra axiom schema

$$\mathbf{x} \neq \mathbf{y} \leftrightarrow (x_i \neq y_i \text{ for some } i),$$

for all n and $\mathbf{x},\mathbf{y} \in S^n$. By Proposition 1.16, an apartness is an inequivalence. By Propositions 1.4 and 1.9 an apartness must satisfy the well-known conditions

$$\neg a \neq a; \tag{11}$$

$$a \neq b \rightarrow b \neq a; \text{ and} \tag{12}$$

$$a \neq b \rightarrow (a \neq c \vee c \neq b). \tag{13}$$

An apartness relation is *tight* if $\neg a \neq b$ implies $a = b$. A tight apartness is an inequality. By Proposition 1.6, a function $\mathbf{f}$ is strongly extensional if and only if $\mathbf{f}(\mathbf{x}) \neq \mathbf{f}(\mathbf{y})$ implies that $x_i \neq y_i$ for some i. Properties (11), (12), and (13) suffice to reconstruct an apartness relation.

Proposition 1.10 *Let $\neq$ be a binary relation on S. Define $\neq$ on S^n by $\mathbf{x} \neq \mathbf{y}$ if and only if $x_i \neq y_i$ for some i. If $\neq$ satisfies* (1), (12), *and* (13), *then the extension to all S^n is a difference. If $\neq$ satisfies* (11), (12), *and* (13), *then it is an apartness.*

Proof: Clearly, (11) implies (1). As to (2), let $\mathbf{f}: S^m \rightarrow S^n$ be an elementary map, $\mathbf{y} \in S^m$, and $\mathbf{t} \in S^n$ such that $\mathbf{f}(\mathbf{y}) \neq \mathbf{t}$. So $\pi_i\mathbf{f}(\mathbf{y}) \neq t_i$ for some i. If $\pi_i\mathbf{f}$ is the projection on the j^{th} coordinate, then $y_j \neq t_i$. So $x_j \neq y_j$ as j^{th} coordinate of $\langle \mathbf{x},\mathbf{f}(\mathbf{x})\rangle \neq \langle \mathbf{y},\mathbf{t}\rangle$, or $x_j \neq t_i$ as $(m+i)^{\text{th}}$ coordinate of $\langle \mathbf{x},\mathbf{f}(\mathbf{x})\rangle \neq \langle \mathbf{y},\mathbf{t}\rangle$.

The standard example of a tight apartness relation is the one on the real line. Define $r \neq s$ if and only if there exists a positive natural number n such that $|r - s| > 1/n$.

A generalization of the apartness on **R** is the apartness on local rings. A *local ring* is a ring (satisfying the usual universal properties for rings) such that if $r + s$ is a unit, then r is a unit or s is a unit. A local ring is *nontrivial* if 1 is not equal to 0. A Heyting field is a nontrivial commutative local ring such that 0 is the only nonunit, that is, if r is not a unit, then $r = 0$. The real numbers form a Heyting field (see Mines et al. [10]).

Let R be a local ring. Define $r \neq s$ if and only if $r - s$ is a unit. If $r - s$ is a unit, then $s - r$ is a unit. So $\neq$ is symmetric. If $r \neq s$, then $r - t + t - s$ is a unit, so by the local ring property, $r - t$ is a unit or $t - s$ is a unit. Thus $r \neq t$ or $t \neq s$. If $r \neq r$, then 0 is a unit, so $s \neq t$ for all s and t. By Proposition 1.10 $\neq$ is a difference relation on R. It is an apartness on R if and only if R is nontrivial. If R is commutative, then $\neq$ is a tight apartness if and only if R is a Heyting field.

Unions and intersections of differences are again differences:

Proposition 1.11 *Let $\neq_i$, $i \in I$, be a collection of relations, each defined on all finite powers of S simultaneously. Define $\neq$ by $\mathbf{x} \neq \mathbf{y}$ if and only if $\mathbf{x} \neq_i \mathbf{y}$ for some $i \in I$. If all $\neq_i$ satisfy one of the properties* (1) *through* (3) *or* (6) *through*

(13) *then* $\neq$ *satisfies that same property. In particular, if all* $\neq_i$ *are differences, then so is* $\neq$*; if all* $\neq_i$ *are proper, then so is* $\neq$*; and if all* $\neq_i$ *are apartnesses, then so is* $\neq$*. If all* $\neq_i$ *are inequivalences, then so is* $\neq$.

Proof: The cases for conditions (1) through (3) and conditions (6) through (13) immediately follow from their logical form. Suppose that all $\neq_i$ are inequivalences, and suppose that $\mathbf{x} \sim \mathbf{y}$ and $\mathbf{y} \sim \mathbf{z}$. Then $\mathbf{x} \sim_i \mathbf{y}$ and $\mathbf{y} \sim_i \mathbf{z}$ for all i. So $\mathbf{x} \sim_i \mathbf{z}$ for all i, and thus $\mathbf{x} \sim \mathbf{z}$.

Proposition 1.12 *Let* $\neq_i$, $i \in I$, *be a collection of relations, each defined on all finite powers of S simultaneously. Define* $\neq$ *by* $\mathbf{x} \neq \mathbf{y}$ *if and only if* $\mathbf{x} \neq_i \mathbf{y}$ *for all* $i \in I$*. If all* $\neq_i$ *satisfy one of the properties* (1) *through* (3) *or* (6) *through* (12) *then* $\neq$ *satisfies that same property. In particular, if all* $\neq_i$ *are differences, then so is* $\neq$*; and if at least one* $\neq_i$ *is proper, then so is* $\neq$.

Proof: The cases for conditions (1) through (3) and conditions (6) through (12) immediately follow from their logical form.

Proposition 1.13 *Let* $\{\neq_i\}_i$ *be a collection of differences on a set. Then partial functions that are strongly extensional with respect to all* $\neq_i$ *are also strongly extensional with respect to their union and intersection.*

Proof: Suppose $\neq$ is the union of the differences $\neq_i$, and let $\mathbf{f}$ be strongly extensional with respect to all $\neq_i$. If $\langle \mathbf{f(x)},\mathbf{z}\rangle \neq \langle \mathbf{f(y)},\mathbf{w}\rangle$, then $\langle \mathbf{f(x)},\mathbf{z}\rangle \neq_i \langle \mathbf{f(y)},\mathbf{w}\rangle$ for some i. So $\langle \mathbf{x},\mathbf{z}\rangle \neq_i \langle \mathbf{y},\mathbf{w}\rangle$, and thus $\langle \mathbf{x},\mathbf{z}\rangle \neq \langle \mathbf{y},\mathbf{w}\rangle$. A similar argument works for the intersection case.

Local rings with inequality defined by $r \neq s$ if and only if $r - s$ is invertible are examples of structures that need not have a proper difference relation. The standard difference on a local ring is proper only if the ring is nontrivial. For some applications, however, it may be essential to have a proper difference. In that case Proposition 1.12 is useful: Intersect the existing difference with denial inequality to make it proper. All functions are strongly extensional with respect to the denial inequality. Then Proposition 1.13 guarantees that functions that are strongly extensional with respect to the original difference are still strongly extensional with respect to the intersection of the original difference with denial inequality.

Examples 1.14 Even in classical mathematics, intersections of inequivalences need not be inequivalences. So we use Proposition 1.12 to construct an example of a discrete set with a decidable difference relation that is not an inequivalence.

1.14.1. Consider the discrete set $S = \{a, b, c\}$ with differences $\neq_1$ and $\neq_2$ that are complements of the equivalence relations on S with partitions $\{\{a, b\},\{c\}\}$ and $\{\{a\},\{b, c\}\}$ respectively. Then the intersection $\neq$ of $\neq_1$ and $\neq_2$ is such that $a \neq c$, $a \sim b$, and $b \sim c$. So $\neq$ is a decidable difference that is not the complement of a transitive relation even though $\neq_1$ and $\neq_2$ are decidable apartnesses. Thus differences are essentially more general than complements of equivalence relations.

1.14.2. Even if a difference is such that for some n the associated nearness is an equivalence relation on S^i for all $i < n$, then it still need not be an inequivalence. Example: Let R be a nontrivial commutative ring, and set $S = R^n$. Define $\mathbf{x} \neq \mathbf{y}$ if and only if $S = \sum_i (x_i - y_i)R$. By Proposition 1.5 $\neq$ is a difference on S. Then $\mathbf{x} \sim \mathbf{y}$ for all $i < n$ and $\mathbf{x}, \mathbf{y} \in S^i$. But the nearness relation is not an equivalence in S^n, for if $e_1, \ldots, e_n$ is a basis of S, then $\langle 0,0,\ldots,0\rangle \sim \langle 0, e_2, \ldots, e_n\rangle$ and $\langle 0, e_2, \ldots, e_n\rangle \sim \langle e_1, e_2, \ldots, e_n\rangle$, but $\langle 0,0,\ldots,0\rangle \neq \langle e_1, e_2, \ldots, e_n\rangle$.

1.14.3. If $\neq$ is an apartness relation, then the schema

$$\mathbf{f}(\mathbf{x}) \neq \mathbf{f}(\mathbf{y}) \rightarrow \mathbf{x} \neq \mathbf{y}$$

suffices to show that $\mathbf{f}$ is strongly extensional. In general, the schema is insufficient as it is essentially weaker than (5). Let S be the discrete set of Example 1.14.1 with decidable difference $\neq$. Define $f: S \rightarrow S$ by $f(a) = a$, $f(b) = a$, and $f(c) = b$. Then the schema above holds since $f(x) \sim f(y)$ for all $x, y \in S$. But f is not strongly extensional since $\langle f(b), b\rangle \neq \langle f(c), c\rangle$ and $\langle b, b\rangle \sim \langle c, c\rangle$.

Examples 1.15 Let (S, d) be a set S with pseudometric d, that is, d is a function from S^2 to $\mathbf{R}$ such that $d(x,x) = 0$, $d(x,y) = d(y,x)$, and $d(x,z) \leq d(x,y) + d(y,z)$. It is well-known that a pseudometric induces an apartness relation on S by $s \neq t$ if and only if $d(s,t) > 0$. The apartness is tight if and only if the pseudometric is a metric. Let r be a real number. A *difference with resolution* $r \geq 0$ on S is a difference $\neq$ satisfying $a \neq b$ if $d(a,b) > r$, and $a \sim b$ if $d(a,b) < r$, for all $a, b \in S$. So the standard apartness on S is a difference with resolution 0. For each $r \geq 0$, do there exist differences with resolution r on S?

1.15.1. Before resolving this question, consider the following nonexample. Define $\mathbf{x} \neq_r \mathbf{y}$ if and only if $\sum_i d(x_i, y_i) > r$. Then $\neq_r$ satisfies the conditions of Proposition 1.5 except for condition (6). Functions $\mathbf{f}$ satisfy (5) if $\sum_j d(f_j(\mathbf{x}), f_j(\mathbf{y})) \leq \sum_i d(x_i, y_i)$. This nonexample suggests ways by which to generalize the notion of difference relation.

1.15.2. To construct differences with resolution r on S, we follow a less elegant route. A subset $X \subseteq S$ is *open* if for all $x \in X$ there exists $\epsilon > 0$ such that $B(x, \epsilon) \subseteq X$, where $B(x, \epsilon) = \{y \in S \mid d(x,y) < \epsilon\}$. For each pair of open sets $p = (A_p, B_p)$ such that $A_p \cup B_p = S$ we define the difference $\neq_p$ by $\mathbf{x} \neq_p \mathbf{y}$ if and only if there exists i such that $d(x_i, y_i) > 0$, and $x_i \in A$ and $y_i \in B$, or $x_i \in B$ and $y_i \in A$. We easily verify that $\neq_p$ is an apartness relation. For $A \subseteq S$ and $r \in \mathbf{R}$, define $d(A) < r$ to mean that $d(a,b) < r$ for all $a, b \in A$. Similarly, $d(A) > r$ means that $d(a,b) > r$ for some $a, b \in A$. A *cover* of S is a collection γ of pairs $p = (A_p, B_p)$ of open sets A_p and B_p with $A_p \cup B_p = S$, such that $\bigcup_{p\in\gamma} A_p = S$. By Proposition 1.11, the union $\neq_\gamma$ of the apartnesses $\neq_p$ is again an apartness. A cover γ has *refinement* r if $d(A_p) < r$ for all $p \in \gamma$. Clearly, if γ has refinement r, then $a \neq_\gamma b$ whenever $d(a,b) > r$. Let $\neq_r$ be the intersection of all covers $\neq_\gamma$ of refinement r. We leave it as an exercise to show that $\neq_r$ is a difference with resolution r. If $r \leq s$, then $(\neq_s) \subseteq (\neq_r)$.

Unfortunately, difference relations $\neq_r$ usually have few strongly extensional functions.

A nearness relation associated with an inequivalence is completely determined by its binary relation $\sim$ on S:

Proposition 1.16 *A difference is an inequivalence if and only if it is proper and satisfies*

$$\text{(4)} \qquad \mathbf{x} \sim \mathbf{y} \leftrightarrow \forall i(x_i \sim y_i).$$

Proof: Suppose $\neq$ is proper and satisfies (4). Obviously, $\sim$ is reflexive and symmetric. Let $\mathbf{x} \sim \mathbf{y}$ and $\mathbf{y} \sim \mathbf{z}$. Then by (4) $\langle \mathbf{x},\mathbf{y} \rangle \sim \langle \mathbf{y},\mathbf{z} \rangle$. Repeated application of (8) and (10) yields $\mathbf{x} \sim \mathbf{z}$. Conversely, suppose that the difference is an inequivalence. Reflexivity implies that $\neq$ is proper, and (7) and (8) imply $\mathbf{x} \sim \mathbf{y} \rightarrow \forall i(x_i \sim y_i)$. Suppose $\mathbf{x} \sim \mathbf{y}$ and $a \sim b$. It suffices to show $\langle \mathbf{x}, a \rangle \sim \langle \mathbf{y}, b \rangle$. This follows immediately from $\langle \mathbf{x}, a \rangle \sim \langle \mathbf{x}, b \rangle$ and $\langle \mathbf{x}, b \rangle \sim \langle \mathbf{y}, b \rangle$, and the transitivity of $\sim$.

Lemma 1.17 *A proper difference is an inequality if and only if $\neg a = b$ implies $\neg\neg a \neq b$, for all a, b.*

Proof: Suppose $\neq$ is a proper difference such that $\neg a = b$ implies $\neg\neg a \neq b$, for all a, b. From Proposition 1.9 it follows that $\mathbf{x} \neq \mathbf{x}$ implies $\perp$. So we have $\mathbf{x} \neq \mathbf{y} \rightarrow \neg \mathbf{x} = \mathbf{y}$. Assume $\neg \mathbf{x} = \mathbf{y}$. Then $\neg\neg\exists i \neg x_i = y_i$. So $\neg\neg\exists i \neg\neg x_i \neq y_i$. And thus $\neg\neg\exists i x_i \neq y_i$, hence $\neg\neg \mathbf{x} \neq \mathbf{y}$. So $\neq$ is an inequality.

The converse is trivial.

Corollary 1.18 *The union of a proper difference and an inequality is an inequality.*

Proof: Let $\neq$ be the union of a proper difference $\neq_1$ and an inequality $\neq_2$. By Proposition 1.11, $\neq$ is a proper difference. Suppose $\neg a = b$. Then $\neg\neg a \neq_2 b$, so $\neg\neg a \neq b$. So by Lemma 1.17 $\neq$ is an inequality.

There is no unique way to define what a strongly extensional relation is. In this paper we present two ways. One involves functions between sets with differences.

We may identify an n-ary relation on a set S with a function from S^n to $\Omega = \mathcal{P}\{0\}$, the truth value object. Following an approach along that line, an n-ary relation is a special case of a function $f: S \rightarrow T$ between sets with differences, be it that we have to choose a difference relation for Ω. If there exists a set $U = S \cup T$ with difference such that this difference with restriction to S and T is the difference of S and T respectively, then f is just a partial function $f: U \rightarrow U$. Instead of the union of S and T there may be difference maintaining embeddings of S and T into a set U with difference, that is, the differences on S and T are the same as those of U restricted to the images of S and T respectively. If such U exist, then define $f: S \rightarrow T$ to be *strongly extensional* if $f: U \rightarrow U$ is strongly extensional in the sense of Theorem 1.6. This definition of strong extensionality depends on our choice of U and on the difference on U.

In many cases there is a natural choice for U. If $f: S^m \rightarrow S^n$ is a map between powers of a set S with difference, then S^m and S^n are sets with differences induced by the difference of S. For all k, the embedding $f: S^k \rightarrow S^{k+1}$ defined by $f\langle \mathbf{x}, a \rangle = \langle \mathbf{x}, a, a \rangle$ maintains difference. So choose $U = S^p$ with $p = \max(m, n)$. There exist difference preserving maps of S^m and S^n into S^p. Then $f: S^m \rightarrow S^n$ is strongly extensional as defined in Theorem 1.6 if and only if $f: U \rightarrow U$ is strongly extensional in the sense of Theorem 1.6.

If there is no choice for U as described in the example above, the disjoint union can be an alternative. Let $U = S \amalg T$. We extend the difference relations of S and T to U, and consider the function $f: S \to T$ as partial function $f_U : U \to U$. Define $\neq$ on powers U^n by setting $\mathbf{x} \neq \mathbf{y}$ if and only if either for some i, x_i and y_i come from different sets S and T, or else, up to a permutation π of the indices, there exist $\mathbf{s}_1, \mathbf{s}_2 \in S^p$ and $\mathbf{t}_1, \mathbf{t}_2 \in T^q$ such that $\pi\mathbf{x} = \langle \mathbf{s}_1, \mathbf{t}_1 \rangle$, $\pi\mathbf{y} = \langle \mathbf{s}_2, \mathbf{t}_2 \rangle$, and $\mathbf{s}_1 \neq \mathbf{s}_2$ over S or $\mathbf{t}_1 \neq \mathbf{t}_2$ over T. The relation $\neq$ on U is called the *canonical extension* of the difference relations on S and T.

Proposition 1.19 *Let S and T be sets with proper difference relations, and let $U = S \amalg T$ be the disjoint union. Then the canonical extension $\neq$ to U is a proper difference relation whose restrictions to S and T are the differences on S and T respectively.*

Proof: Clearly, the canonical extension satisfies (1) and (3), and the restrictions of $\neq$ to S and T reproduce the original differences on them. Note that this requires the differences on S and T to be proper. Let $\mathbf{f}: U^m \to U^n$ be an elementary map such that $\mathbf{f}(\mathbf{y}) \neq \mathbf{t}$. Then $\mathbf{f}$ is a sequence of projections $(\pi_{\lambda 1}, \ldots, \pi_{\lambda n})$. So $\langle y_{\lambda 1}, \ldots, y_{\lambda n} \rangle \neq \mathbf{t}$. If $y_{\lambda i} \neq t_i$ for some i because they are from different sets S and T, then for the same reason $\langle x_{\lambda i}, x_{\lambda i} \rangle \neq \langle y_{\lambda i}, t_i \rangle$, and so by repeated application of (7), (8), and (10) $\langle \mathbf{x}, \mathbf{f}(\mathbf{x}) \rangle \neq \langle \mathbf{y}, \mathbf{t} \rangle$. Otherwise, suppose we have $\mathbf{x} \neq \mathbf{y}$ because for some i, x_i and y_i are from different sets S and T. Then by (7), $\langle \mathbf{x}, \mathbf{f}(\mathbf{x}) \rangle \neq \langle \mathbf{y}, \mathbf{t} \rangle$. Finally, suppose that for all i either both x_i and y_i are in S or both are in T, and that for all i either both $y_{\lambda i}$ and t_i are in S or both are in T. Then there exists a permutation π such that $\pi\mathbf{f}(\mathbf{y}) = \langle \mathbf{f}_S(\mathbf{y}), \mathbf{f}_T(\mathbf{y}) \rangle$ and $\pi\mathbf{t} = \langle \mathbf{t}_S, \mathbf{t}_T \rangle$, where $\mathbf{f}_S(\mathbf{y}) \neq \mathbf{t}_S$ over S or $\mathbf{f}_T(\mathbf{y}) \neq \mathbf{t}_T$ over T. Let $\mathbf{x}_S$, $\mathbf{x}_T$, $\mathbf{y}_S$, and $\mathbf{y}_T$ be the subsequences of $\mathbf{x}$ and $\mathbf{y}$ of elements that belong to S and T respectively. Then $\langle \mathbf{x}_S, \mathbf{f}_S(\mathbf{x}) \rangle \neq \langle \mathbf{y}_S, \mathbf{t}_S \rangle$ or $\langle \mathbf{x}_T, \mathbf{f}_T(\mathbf{x}) \rangle \neq \langle \mathbf{y}_T, \mathbf{t}_T \rangle$. After merging the two relations, we get $\langle \mathbf{x}, \mathbf{f}(\mathbf{x}) \rangle \neq \langle \mathbf{y}, \mathbf{t} \rangle$.

The assumption of properness is essential in Proposition 1.19. If we don't assume the differences on S and T are proper, we may not be able to derive (1) for the canonical extension $\neq$ on $S \amalg T$. If for example $s \neq s$ for some $s \in S$, then $\langle s, t \rangle \neq \langle s, t \rangle$ for all $t \in T$, and thus by (1) we would have $t \neq t$ for all $t \in T$. So the difference on T could not be proper either.

Another way to define strongly extensional n-ary relations is by returning to the original classical axiomatization of inequality:

$$x \neq x \vdash \bot$$

$$Ay \vdash x \neq y \vee Ax,$$

where in the last schema the variables x, y are not bound by a quantifier of A. We wish to replace the right-hand side of the second schema by a difference between sequences $\langle x, A \rangle \neq \langle y, x \rangle$. So defining strong extensionality for relations using sequences reduces to introducing a new constant A to S and extending the difference relation from S to $S \cup \{A\}$. Let R be an n-ary relation on S. Then R is a unary relation on S^n. Rather than defining strong extensionality of R over S, we define strong extensionality of R over S^n. So without loss of generality we define strong extensionality for unary relations only. Let R be a unary relation on a set S with difference. Then R is *strongly extensional* if there is an extension

$U = S \cup \{r\}$ with difference relation, such that the difference of U with restriction to S is the original difference of S, and such that for all $s \in S$, Rs holds if and only if $s \neq r$.

Example 1.20 Let S be a set with apartness, and let R be a unary relation on S satisfying

$$Rs \to (s \neq t \vee Rt). \tag{14}$$

Then R is strongly extensional: The apartness of S extends to be apartness on $S \cup \{r\}$ by setting $r \neq s$ if and only if Rs. Conversely, if the difference on $S \cup \{r\}$ is an apartness, then R satisfies (14).

2 Applications to algebra Groups and rings with differences are defined by the usual universal axioms together with the condition that the standard functions are strongly extensional. So a group G with difference consists of a set G with a difference relation, constant e, unary function ${}^{-1} : G \to G$, and binary function $\cdot : G \times G \to G$ such that ${}^{-1}$ and $\cdot$ are strongly extensional and such that for all $g, h, i \in G$ we have

$$g\cdot e = e\cdot g = g;$$

$$g\cdot(h\cdot i) = (g\cdot h)\cdot i; \text{ and}$$

$$g\cdot g^{-1} = g^{-1}\cdot g = e.$$

Proposition 2.1 *Let G be a group with a difference relation on the underlying set. Then G is a group with difference if and only if for all $a, b, x, \mathbf{c}, \mathbf{d}$ we have that $\langle a,\mathbf{c}\rangle \neq \langle b,\mathbf{d}\rangle$ implies $\langle ax,\mathbf{c}\rangle \neq \langle bx,\mathbf{d}\rangle$ and $\langle xa,\mathbf{c}\rangle \neq \langle xb,\mathbf{d}\rangle$.*

Proof: It suffices to show that multiplication and inverse are strongly extensional. Suppose $\langle ab,\mathbf{z}\rangle \neq \langle cd,\mathbf{w}\rangle$. Multiply by c^{-1} on the left and by b^{-1} on the right to get $\langle c^{-1}a,\mathbf{z}\rangle \neq \langle db^{-1},\mathbf{w}\rangle$. So $\langle c^{-1}a,1,\mathbf{z}\rangle \neq \langle 1,db^{-1},\mathbf{w}\rangle$. So after two more multiplications we arrive at $\langle a,b,\mathbf{z}\rangle \neq \langle c,d,\mathbf{w}\rangle$. Thus multiplication is strongly extensional.

The strong extensionality of the inverse follows from Proposition 1.8 with $f(x,y) = xy$ and $g(x,y) = 1$.

Let G be a group with normal subgroup N. Define $\neq_N$ by $\mathbf{x} \neq_N \mathbf{y}$ if and only if the normal subgroup generated by $\{\ldots, x_i y_i^{-1}, \ldots\}$ contains N. One easily verifies that $\neq_N$ satisfies the conditions of Propositions 1.5 and 2.1. So G with $\neq_N$ is a group with difference.

The following example of a group with difference was suggested to us by Fred Richman. It illustrates that there exists an elementary algebraic structure whose natural relation $\neq$ is a difference that cannot be shown to be an inequivalence. Let $\mathbf{Z}$ be the group of integers and let $\mathbf{N}$ be the set of natural numbers. Define Q to be the quotient group $Q = \mathbf{Z}^{\mathbf{N}}/\sum_{\mathbf{N}} \mathbf{Z}$. A natural way to define an inequality on Q would be to set $\mathbf{a} \neq 0$ if and only if there are infinitely many $n \in \mathbf{N}$ such that $\mathbf{a}(n)$ is not 0, that is, for all $m > 0$ there exists $n > m$ such that $\mathbf{a}(n)$ is not 0. Define $\neq$ by $\langle \mathbf{a}_1, \ldots, \mathbf{a}_n\rangle \neq 0$ if and only if there are infinitely many elements unequal to 0; and $\langle \mathbf{a}_1, \ldots, \mathbf{a}_n\rangle \neq \langle \mathbf{b}_1, \ldots, \mathbf{b}_n\rangle$ if and only if $\langle \mathbf{a}_1 - \mathbf{b}_1, \ldots, \mathbf{a}_n - \mathbf{b}_n\rangle \neq 0$. It is immediate from Propositions 1.5 and 2.1 that this makes Q a group with difference.

A subset $X \subseteq Y$ is *detachable* from Y if the union of X and its complement equals Y. Let E be the set of even numbers and let O be the set of odd numbers. So E and O are countably infinite detachable subsets of $\mathbf{N}$ such that $E \cup O = \mathbf{N}$, and $E \cap O = \varnothing$. Consider the principle EO: If $A \subseteq \mathbf{N}$ is a detachable subset such that $A \cap E$ and $A \cap O$ are not infinite, then A is not infinite.

Now assume that the difference $\neq$ on Q is an inequivalence. Let $A \subseteq \mathbf{N}$ be a detachable subset such that $A \cap E$ and $A \cap O$ are not infinite. Define $a, b \in Q$ by $a(n) = 1$ if and only if $2n \in A$, and $b(n) = 1$ if and only if $2n + 1 \in A$. Then $a \sim 0$ and $b \sim 0$. If $\neq$ is an inequivalence, then $\langle a, b\rangle \sim 0$. But this means that A is not infinite. So if $\neq$ is an inequivalence, then EO holds.

The principle EO is not derivable in constructive mathematics. In Section 4 we present a topos $\mathcal{E}_{\mathbf{G}}$ whose natural number object $\mathbf{N}$ has a detachable subset X such that both X and $\mathbf{N}\backslash X$ are not infinite. EO implies that there exists no such X: a detachable infinite subset $A \subseteq \mathbf{N}$ is isomorphic to $\mathbf{N}$, and $A \cap E$ and $A \cap O$ then are isomorphic to a partition X and $\mathbf{N}/X$ of detachable subsets of $\mathbf{N}$.

A ring with difference is a set R with difference satisfying the well-known universal axioms for zero, one, addition, and multiplication such that addition and multiplication are strongly extensional. A ring is *nontrivial* if $1 \neq 0$. The partial function of multiplicative inverse $f(x) = x^{-1}$ is implicitly defined by the equation $xy = 1$, hence by Proposition 1.8 is strongly extensional.

Proposition 2.2 *Let R be a ring with a difference relation on the underlying set. Then R is a ring with difference if and only if for all $a, b, x, \mathbf{c}, \mathbf{d}$ we have that $\langle a,\mathbf{c}\rangle \neq \langle b,\mathbf{d}\rangle$ implies $\langle a + x,\mathbf{c}\rangle \neq \langle b + x,\mathbf{d}\rangle$, and $\langle ab,\mathbf{c}\rangle \neq \langle 0,\mathbf{d}\rangle$ implies $\langle b,\mathbf{c}\rangle \neq \langle 0,\mathbf{d}\rangle$ and $\langle a,\mathbf{c}\rangle \neq \langle 0,\mathbf{d}\rangle$.*

Proof: By Proposition 2.1 the additive abelian group is a group with difference. Suppose $\langle ab,\mathbf{z}\rangle \neq \langle cd,\mathbf{w}\rangle$. Then $\langle ab, ad,\mathbf{z}\rangle \neq \langle ad, cd,\mathbf{w}\rangle$. So $\langle a(b - d), (a - c)d,\mathbf{z}\rangle \neq \langle 0,0,\mathbf{w}\rangle$, and thus $\langle b - d, a - c,\mathbf{z}\rangle \neq \langle 0,0,\mathbf{w}\rangle$. So $\langle a, b,\mathbf{z}\rangle \neq \langle c, d,\mathbf{w}\rangle$.

The abelian group Q above is a ring with difference with multiplication $\mathbf{a}\cdot\mathbf{b} = \mathbf{c}$ with $\mathbf{c}(n) = \mathbf{a}(n)\mathbf{b}(n)$ for all n.

Let R be a ring, I a two-sided ideal of R. Define $\neq_I$ by $\mathbf{x} \neq_I \mathbf{y}$ if and only if the ideal $\sum_i R(x_i - y_i)R$ contains I. We immediately see from Propositions 1.5 and 2.2 that this makes R a ring with difference $\neq_I$.

Proposition 2.3 *Let R be a ring with difference, and let $n > 0$. Then we have*

(i) $$\langle ax_1, \ldots, ax_n,\mathbf{y}\rangle \neq 0 \rightarrow \langle a,\mathbf{y}\rangle \neq 0;$$

(ii) $$\mathbf{x} \neq 0 \rightarrow 1 \neq 0;$$

(iii) $$\langle a^n,\mathbf{y}\rangle \neq 0 \rightarrow \langle a,\mathbf{y}\rangle \neq 0; \textit{ and}$$

(iv) $$\langle a^n, b + ac,\mathbf{x}\rangle \neq 0 \rightarrow \langle a, b,\mathbf{x}\rangle \neq 0.$$

Proof: For (i) we have $\langle ax_1, \ldots, ax_n,\mathbf{y}\rangle \neq 0 \rightarrow \langle ax_1, \ldots, ax_n,\mathbf{y}\rangle \neq \langle 0x_1, \ldots, 0x_n, 0\rangle$. So $\langle a,\mathbf{y}\rangle \neq 0$.

(ii) follows immediately from (i).

By (i), $\langle a^{n+1},\mathbf{y}\rangle \neq 0$ implies $\langle a^n,\mathbf{y}\rangle \neq 0$. Repeated application yields (iii).

For (iv), $\langle a^n, b + ac,\mathbf{x}\rangle \neq \langle 0^n, 0 + 0c, 0, \ldots, 0\rangle$, so $\langle a, b,\mathbf{x}\rangle \neq 0$.

The polynomial ring $R[X]$ over a commutative ring R with difference is defined in the usual way. It remains to construct a difference on $R[X]$. $R[X]$ can be considered as a subset of $\bigcup_{n\in\mathbf{N}} R^n$, and so borrows the difference from R by defining $\langle f_1,\ldots,f_n\rangle \neq \langle g_1,\ldots,g_n\rangle$ if and only if the sequences of coefficients differ over R, that is,

$$\langle a_{01},\ldots,a_{m1},\ldots,a_{0n},\ldots,a_{mn}\rangle \neq \langle b_{01},\ldots,b_{m1},\ldots,b_{0n},\ldots,b_{mn}\rangle,$$

where $f_i = a_{0i} + \cdots + a_{mi}X^m$ and $g_i = b_{0i} + \cdots + b_{mi}X^m$. We easily see that the addition and multiplication operations of $R[X]$ are strongly extensional since they are built up from the addition and multiplication operations of R.

We say $deg\, f \leq n$ if $f = a_0 + \cdots + a_nX^n$ for some $a_i \in R$. We say $deg\, f \geq n$ if $f = g + hX^n$ for some $g, h \in R[X]$ with $deg\, g \leq n-1$ and $h \neq 0$. Let $g = b_0 + \cdots + b_m X^m$ for some $b_i \in R$. We say $deg\, f \leq deg\, g$ if for all k, $\langle a_k,\ldots,a_n\rangle \neq 0$ implies $\langle b_k,\ldots,b_m\rangle \neq 0$. We say $deg\, f < deg\, g$ if for all k, $\langle a_k,\ldots,a_n\rangle \neq 0$ implies $\langle b_{k+1},\ldots,b_m\rangle \neq 0$.

The definition of integral domain presents us with the problems of establishing what structures we want to be integral domains, and what properties we should be able to derive for integral domains. The ring $\mathbf{Z}$ of integers and the ring $\mathbf{R}$ of real numbers with apartness must be integral domains; integral domains must have quotient fields, where a field is an integral domain such that a is invertible whenever $a \neq 0$; and polynomial rings in one variable over integral domains must be integral domains.

A commutative ring with difference is an *integral domain with difference* if it satisfies:

(1) $$1 \neq 0;$$

(2) $$a \neq 0 \wedge ab = 0 \rightarrow b = 0;$$

(3) $$a \neq 0 \wedge b \neq 0 \rightarrow ab \neq 0;$$

(4) $$\mathbf{x} \neq 0 \wedge \langle \ldots, x_ib, \ldots \rangle = 0 \rightarrow b = 0; \text{ and}$$

(5) $$\mathbf{x} \neq 0 \wedge \mathbf{y} \neq 0 \rightarrow \langle \ldots, x_iy_j, \ldots \rangle \neq 0.$$

A *field with difference* is an integral domain with difference satisfying

(6) If $a \neq 0$ then a is invertible.

Clearly, (4) implies (2), and (5) implies (3). Let $R = \mathbf{Z}[X, Y, Z]/(XZ, YZ, Z^2)$, and let $I = XR + YR$, the ideal generated by X and Y. Define $\mathbf{x} \neq \mathbf{y}$ if and only if the ideal $\sum_i (x_i - y_i)R$ contains some power I^n of I. Then R is a commutative ring with difference. We have $a \neq 0$ if and only if $a = 1 + rZ$ or $-1 + rZ$ for some $r \in R$. So $a \neq 0$ if and only if a is a unit. We easily verify that R satisfies (1), (2), (3), (5), and (6). But (4) fails since $\langle X, Y\rangle \neq 0$ and $\langle XZ, YZ\rangle = 0$.

Let $R = \mathbf{Z}[X, Y]$. Define $\mathbf{x} \neq \mathbf{y}$ if and only if the ideal $\sum_i (x_i - y_i)R$ contains the ideal $I = XR + YR$. Then R is a commutative ring with difference, and $a \neq 0$ if and only if $a = 1$ or -1. So we easily verify that R satisfies (1), (2), (3), (4), and (6). But (5) fails because $\langle X, Y\rangle \neq 0$ while $\langle X^2, XY, Y^2\rangle \sim 0$.

Let $\mathbf{Z}$ be the ring of integers. The prime ideals $2\mathbf{Z}$ and $3\mathbf{Z}$ induce the usual decidable equivalence relations $\sim_2$ and $\sim_3$ on $\mathbf{Z}$ with corresponding difference relations $\neq_2$ and $\neq_3$. The standard ring operations preserve the equivalences, so

$\mathbf{Z}$ is an integral domain with difference with respect to $\neq_2$ as well as with respect to $\neq_3$. Let $\neq$ be the intersection of $\neq_2$ and $\neq_3$. Then by Proposition 1.13 $\mathbf{Z}$ is an integral domain with difference with respect to $\neq$. Note that the decidable relation $\neq$ is not an inequivalence as $2 \sim 0$ and $3 \sim 0$, while $\langle 2,3 \rangle \neq 0$.

Proposition 2.4 *Let R be a commutative ring with difference satisfying* (1), (2), *and* (3). *If $\neq$ is an apartness, then R is an integral domain. If $\neq$ is denial inequality and equality is stable, that is, $\neg\neg a = b$ implies $a = b$, then R is an integral domain.*

Proof: The case for apartness is trivial.

Suppose that $\neq$ is denial inequality and $=$ is stable. If $\mathbf{x} \neq 0$ and $\langle \ldots, x_i y, \ldots \rangle = 0$, then $\neg\neg\exists i(x_i \neq 0 \wedge x_i y = 0)$. So $\neg\neg y = 0$, and thus $y = 0$. That proves (4). Let $\mathbf{x}$ and $\mathbf{y}$ be such that $\mathbf{x} \neq 0$, $\mathbf{y} \neq 0$, and $\langle \ldots, x_i y_j, \ldots \rangle = 0$. Then for all i and j we have $\neg\neg(x_i = 0 \vee y_j = 0)$. So for all i, $\neg\neg(x_i = 0 \vee \mathbf{y} = 0)$. Thus $\neg\neg(\mathbf{x} = 0 \vee \mathbf{y} = 0)$. Contradiction. Thus R satisfies (5).

So not only the ring $\mathbf{Z}$ and the ring $\mathbf{R}$ with apartness, but even the ring of real numbers $\mathbf{R}$ with denial inequality are integral domains with difference.

From ([10], p. 47) we know that (1), (2), and (3) are necessary and sufficient to embed a commutative ring with difference in a field. The quotient field Q of an integral domain R is constructed by localizing to the set $S = \{s \in R \mid s \neq 0\}$. Then S is a multiplicative set because of (1) and (3), and R embeds in Q because of (2). The difference on Q is defined by $\langle x_1/s_1, \ldots, x_n/s_n \rangle \neq 0$ over Q if and only if $\langle x_1, \ldots, x_n \rangle \neq 0$ over R. Obviously, this relation satisfies (1), (4), and (5).

It remains to present the motivations for (4) and (5) in the definition of integral domains with difference. Suppose $R[X]$ is a commutative ring satisfying (2). Then for all $f = \sum_i x_i X^i \neq 0$ and $y \in R$ such that $fy = 0$, we have $y = 0$. So R satisfies (4). Suppose $R[X]$ is a commutative ring satisfying (3). Let $f = \sum_i x_i X^i$ and $g = \sum_j y_j X^j$ be such that $f \neq 0$ and $g \neq 0$. Then $fg \neq 0$. So

$$\left\langle x_0 y_0, \ldots, \sum_k x_k y_{h-k}, \ldots, x_m y_n \right\rangle \neq 0.$$

Using the strong extensionality of addition we get $\langle \ldots, x_i y_j, \ldots \rangle \neq 0$. Thus R satisfies (5). So if polynomial rings $R[X]$ over integral domains R must be integral domains themselves, then (4) and (5) are necessary. With Proposition 2.7 we establish that (1) through (5) are sufficient.

Lemma 2.5 *Let R be a commutative ring with difference satisfying* (5). *Then*

(i) $$\langle a_1, \ldots, a_n \rangle \neq 0 \to \langle a_1^m, \ldots, a_n^m \rangle \neq 0;$$

(ii) $$\langle a, \mathbf{x} \rangle \neq 0 \wedge \langle b, \mathbf{x} \rangle \neq 0 \to \langle ab, \mathbf{x} \rangle \neq 0.$$

Proof: $\langle a_1, \ldots, a_n \rangle \neq 0$ implies $\langle \ldots, a_i a_j, \ldots \rangle \neq 0$. Repeated application of Proposition 2.3(i) yields $\langle a_1, \ldots, a_{n-1}, a_n^2 \rangle \neq 0$. Iteration of this process yields (i).

If $\langle a, \mathbf{x} \rangle \neq 0$ and $\langle b, \mathbf{x} \rangle \neq 0$, then (5) implies

$$\langle ab, ax_1, \ldots, ax_n, bx_1, \ldots, bx_n, \ldots, x_i x_j, \ldots \rangle \neq 0.$$

Repeated application of Proposition 2.3(i) yields $\langle ab, \mathbf{x} \rangle \neq 0$.

Lemma 2.6 *Proposition* 2.5(ii) *is equivalent to* (5).

Proof: Let $\mathbf{x} \times \mathbf{y} = \langle \ldots, x_i y_j, \ldots \rangle$, and $\mathbf{t}^{-i}$ the sequence $\mathbf{t}$ with t_i removed. Suppose $\mathbf{x} \neq 0$ and $\mathbf{y} \neq 0$, and let $\mathbf{z} = \langle \mathbf{x}, \mathbf{y} \rangle$. Then $\langle \mathbf{x} \times \mathbf{y}, \mathbf{z}^{-i} \rangle \neq 0$ for all i. So by Proposition 2.5(ii), $\langle \mathbf{x} \times \mathbf{y}, \mathbf{z}^{-i,-j} \rangle \neq 0$ for all $i < j$. Applying Proposition 2.5(ii) to this new collection by comparing all sequences that differ in one coordinate gives $\langle \mathbf{x} \times \mathbf{y}, \mathbf{z}^{-i,-j,-k} \rangle \neq 0$ for all $i < j < k$. After sufficiently many applications of this operation we obtain $\mathbf{x} \times \mathbf{y} \neq 0$.

Proposition 2.7 *If R is a commutative ring with difference satisfying one of the properties* (1) *or* (5), *then* $R[X]$ *satisfies the same property. If R satisfies both* (4) *and* (5), *then so does* $R[X]$. *If R is an integral domain with difference, then so is* $R[X]$.

Proof: The case for (1) is trivial.

Suppose R satisfies (5). Let A be an $n \times n$ matrix and $\mathbf{b} \in R^n$ such that $d = \det A \neq 0$ and $\mathbf{b} \neq 0$. Let A' be the adjoint of A, that is, $AA' = A'A = dI$. Then $A'A\mathbf{b} = d\mathbf{b} \neq 0$. From the strong extensionality of A' we obtain $A\mathbf{b} \neq 0$. So if $\det A \neq 0$ and $\mathbf{b} \neq 0$, then $A\mathbf{b} \neq 0$. Let $f, g \in R[X]$, $\mathbf{h} \in R[X]^n$ be such that $\langle f, \mathbf{h} \rangle \neq 0$ and $\langle g, \mathbf{h} \rangle \neq 0$. Then $f = \sum_i a_i X^i$ and $g = \sum_j b_j X^j$ for certain $a_i, b_j \in R$. Identify polynomials of degree at most p with vectors in R^p. Then the coefficients of $fg = \sum_k c_k X^k$ form the vector $A\mathbf{b}$, where

$$A = \begin{bmatrix} a_0 & 0 & 0 & \cdots & 0 \\ a_1 & a_0 & 0 & \cdots & 0 \\ a_2 & a_1 & a_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_m & a_{m-1} & a_{m-2} & \cdots & a_0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_n & a_{n-1} & a_{n-2} & \cdots & a_{n-m} \\ 0 & a_n & a_{n-1} & \cdots & a_{n-m+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_n \end{bmatrix}$$

and

$$\mathbf{b} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

So we must show that $\langle A\mathbf{b}, \mathbf{h} \rangle \neq 0$. Let A_i be the $(m + 1) \times (m + 1)$ submatrix of A with the a_i on the diagonal, and let $d_i = \det A_i$. Then $d_i = a_i^{m+1} + \sum_{j<i} a_j p_j(\mathbf{a})$ for some p_j. Now R satisfies (5), so $\langle a_0^{m+1}, \ldots, a_n^{m+1}, \mathbf{h} \rangle \neq 0$. So by a finite induction on n, using Proposition 2.3(iv), $\mathbf{d} = \langle d_0, \ldots, d_n, \mathbf{h} \rangle \neq 0$. Let A'_i be the adjoint of A_i. So $A'_i \langle c_i, \ldots, c_{i+m} \rangle^T = d_i \mathbf{b}$. There exists a linear

map $F = \langle \dots, A'_i\pi_i, \dots \rangle : R^{m+n+1} \to R^{(m+1)(n+1)}$ such that F is a strongly extensional map satisfying $FA\mathbf{b} = \langle \dots, d_i b_j, \dots \rangle$ and $F0 = 0$. So $\langle FA\mathbf{b}, \mathbf{h} \rangle \neq \langle F0, 0 \rangle$. Hence $\langle A\mathbf{b}, \mathbf{h} \rangle \neq 0$. Thus $\langle fg, \mathbf{h} \rangle \neq 0$.

Suppose R satisfies (4) and (5). Let $\langle f_1, \dots, f_m \rangle \neq 0$ and g be such that $\langle f_1 g, \dots, f_m g \rangle = 0$, for $f_i, g \in R[X]$. We may identify g and all f_i with vectors in R^{n+1} for some n. Then $f_i g$ is a vector in R^{2n+1}, and $f_i g = A_i \mathbf{b}$, where A_i is a $(2n+1) \times (n+1)$-matrix as above, and $\mathbf{b}$ is an $(n+1) \times 1$ vector associated with g. Let A_{ij} be the $(n+1) \times (n+1)$ submatrix of A_i with the j^{th} coefficient on the diagonal, and set $d_{ij} = \det A_{ij}$. Then $\langle \dots, d_{ij}, \dots \rangle \neq 0$. Apply a sequence of elementary maps F_i as above. Then $\langle \dots, d_{ij} b_k, \dots \rangle = 0$. So $\mathbf{b} = 0$. Thus $g = 0$.

3 Differences for all powers In Section 2 we were just able to extend the difference from a ring R to the polynomial ring $R[X]$ because $R[X] \subseteq \bigcup_{n \in \mathbf{N}} R^n$. Extending the difference to the power series ring $R[[X]]$ requires a substantial extension of the definition of difference: define $\neq$ on all powers S^X simultaneously rather than on finite powers S^n only. The definition presented in this section follows the 'finite' version of Section 1.

A generalized (proper) difference $\neq$ on a set S is defined on all powers S^X simultaneously. It satisfies axiom schemata that are straightforward generalizations of Section 1(1), Section 1(2), and Section 1(3).

We generalize Axiom (1) of Section 1 as follows: Let $X = Y \cup Z$. If f is a function with domain X, then we write f_Y and f_Z for the functions restricted to the subdomains Y and Z respectively. The generalization of Section 1(1) now reads: for all X, Y, Z such that $X = Y \cup Z$, and all $f, g : X \to S$, we have:

(1) If $f \neq g$ and $f_Z = g_Z$, then $f_Y \neq g_Y$.

For a generalization of Axiom (2) of Section 1 we must extend our definition of elementary function. Let S be the set for which we define a difference tion. For each function $f \colon Y \to X$ there is a corresponding map $f^* : S^X \to S^Y$ defined by $f^*(g) = gf$. The elementary maps of Section 1, defined between finite powers of S, are of the form $f^* : S^m \to S^n$, where f is a function from $n = \{0, \dots, n-1\}$ to $m = \{0, \dots, m-1\}$. More generally, *elementary maps* between S^X and S^Y are defined as the maps f^*, with $f \colon Y \to X$. The generalization of Section 1(2) now reads: For all sets A and B, $f \colon S^A \to S^B$ an elementary map, $x, y \in S^A$, and $t \in S^B$,

(2) if $fy \neq t$, then $\langle x, fx \rangle \neq \langle y, t \rangle$, where $\langle x, fx \rangle, \langle y, t \rangle \in S^{A \amalg B}$.

Proper differences satisfy

(3) $\langle\ \rangle \neq \langle\ \rangle$ is false,

where $\langle\ \rangle$ is the unique element of $S^0 = 1$.

We define nearness $\sim$ by $f \sim g$ if and only if $\neg f \neq g$. An *inequivalence* is a proper difference such that for all sets $X = Y \cup Z$ and $f, g : X \to S$, if $f_Y \sim g_Y$ and $f_Z \sim g_Z$, then $f \sim g$.

A proper difference is an *apartness* if for all X and $f, g : X \to S$, if $f \neq g$, then $f(x) \neq g(x)$ for some $x \in X$. Clearly, an apartness is an inequivalence.

For each collection Λ of partial functions between powers of S, $E(\Lambda)$ is the smallest subcategory of partial maps between powers of S that includes Λ and the elementary maps. The collection $E = E(\emptyset)$ of elementary maps itself forms a subcategory. We define Λ to be a collection of *strongly extensional* maps if all (partial) maps of $E(\Lambda)$ satisfy (2). As in Section 1, we easily show that Λ is strongly extensional if and only if for all $f: S^X \rightharpoonup S^Y \in \Lambda$, all $Z, x, y \in S^X$, and $z, w \in S^Z$,

$$\langle fx, z\rangle \neq \langle fy, w\rangle \text{ implies } \langle x, z\rangle \neq \langle y, w\rangle.$$

There is a canonical way to extend differences defined on the finite powers S^n to differences on all powers S^X. Let $\neq$ be a difference on all finite powers. For all X define $\neq$ on S^X by $f \neq g$ if and only if there is an $n \in \mathbf{N}$ and a map $e: \{1, \ldots, n\} \to X$ such that $fe \neq ge$, that is, $\langle fe(1), \ldots, fe(n)\rangle \neq \langle ge(1), \ldots, ge(n)\rangle$. We call this the *infinite extension* of $\neq$. The extension preserves strong extensionality of functions.

Proposition 3.1 *The infinite extension of a difference relation on the finite powers S^n is a difference. If the finite difference is proper, an inequivalence or an apartness, then so is the infinite extension.*

Proof: Let $f, g: X \to S$ be maps such that $f \neq g$, and suppose $X = Y \cup Z$ such that $f_Z = g_Z$. There is a map $e: \{1, \ldots, n\} \to X$ such that $fe \neq ge$. Since $\neq$ is a difference on the finite powers S^n, we can remove all coordinates i for which $e(i) \in Z$, because for them $fe(i) = ge(i)$. So there is a subsequence generated by a map $d: \{1, \ldots, m\} \to Y$ for some $m \leq n$ such that $fd \neq gd$. Thus $f_Y \neq g_Y$. So $\neq$ satisfies (1).

Suppose $fy \neq t$ for $y \in S^A$, $f = g^*: S^A \to S^B$ elementary, and $t \in S^B$. So $yge \neq te$ for some $e: \{1, \ldots, n\} \to B$. Then $\langle x, fx\rangle\langle ge, e\rangle = \langle xge, fxe\rangle \neq \langle yge, te\rangle = \langle y, t\rangle\langle ge, e\rangle$. Thus $\langle x, fx\rangle \neq \langle y, t\rangle$. So $\neq$ satisfies (2).

Clearly, if a finite difference is proper, then so is its infinite extension.

Suppose the finite difference is an inequivalence, and let $X = Y \cup Z$ be sets and $f, g: X \to S$ such that $f_Y \sim g_Y$ and $f_Z \sim g_Z$. If $f \neq g$, then $fe \neq ge$ for some $e: n \to X$. There are p, q such that $p + q = n$, $e_p: p \to Y$ and $e_q: q \to Z$. Then $fe_p \sim ge_p$ and $fe_q \sim ge_q$. So $fe \sim ge$. Contradiction. Thus $f \sim g$.

The case for apartness is trivial.

If $\neq$ is the denial inequality on the finite powers S^n, then its canonical extension as defined above usually is not the denial inequality on infinite powers S^X.

Example: the denial inequality on the set $\mathbf{N}$ of natural numbers is the well-known discrete inequality, while the infinite extension to $\mathbf{N}^{\mathbf{N}}$ is the apartness relation defined by $f \neq g$ if and only if $f(n) \neq g(n)$ for some n.

A map $f: S^X \to S^Y$ is *strongly extensional* with respect to a difference if for all Z and $v, w \in S^Z$, if $\langle fx_1, v\rangle \neq \langle fx_2, w\rangle$ in $S^{Y \cup Z}$, then $\langle x_1, v\rangle \neq \langle x_2, w\rangle$ in $S^{X \cup Z}$.

Obviously, if $f: S^m \to S^n$ is strongly extensional with respect to a difference relation on the finite powers, then it is also strongly extensional with respect to the infinite extension.

Proposition 3.2 *Let $R[[X]]$ be the power series ring over a commutative ring R with difference. The difference on $R[[X]] = R^{\mathbf{N}}$ is the infinite extension of the difference on R. If R satisfies Section 2(5), then so does $R[[X]]$.*

Proof: Let $f = \sum_i a_i X^i$, $g = \sum_j b_j X^j \in R[[X]]$, $\mathbf{h} \in R[[X]]^n$ be such that $\langle f, \mathbf{h}\rangle \neq 0$ and $\langle g, \mathbf{h}\rangle \neq 0$. Let $f_m = \sum_{i \leq m} a^i X^i$ and $g_n = \sum_{j \leq n} b_j X^j$. By Proposition 2.7 there are m, n such that $\langle f_m g_n, \mathbf{h}\rangle \neq 0$. Write $fg = \sum_i c_i X^i$. We prove by induction on $m + n$ that $\langle c_0, \dots, c_{m+n}, \mathbf{h}\rangle \neq 0$. If $m + n = 0$, then $\langle c_0, \mathbf{h}\rangle \neq 0$. Induction step: If $\langle f_m g_n, \mathbf{h}\rangle \neq 0$, then $\langle c_0, \dots, c_{m+n}, \dots, a_i b_j, \dots, \mathbf{h}\rangle \neq 0$, where i, j are all pairs such that $i + j \leq m + n$ and $i < m$ or $j < n$. So $\langle c_0, \dots, c_{m+n}, f_{m-1}, g_{n-1}, \mathbf{h}\rangle \neq 0$, where $f_{-1} = g_{-1} = 0$. And thus by induction $\langle c_0, \dots, c_{m+n}, \mathbf{h}\rangle \neq 0$.

In general, if R is an integral domain with difference, then $R[[X]]$, with the infinite extension as difference relation, may not satisfy Section 2(4). Let $R = \mathbf{Z}[S, Z_0, Z_1, Z_2, \dots]/J$, where J is the ideal generated by SZ_0 and $SZ_{i+1} + Z_i$, for all i. Define $\mathbf{x} \neq \mathbf{y}$ if and only if the ideal $\sum_i (x_i - y_i)R$ equals R. Then R is an integral domain with difference. Let $f, g \in R[[X]]$ be defined by $f = S + X$ and $g = \sum_i Z_i X^i$. Then $f \neq 0$, $fg = 0$, but g is not identical to 0. So $R[[X]]$ does not satisfy Section 2(4).

Acknowledgment — The idea of a generalized notion of inequality arose following suggestions by William Julian during the writing of [10]. Thanks are due to Fred Richman for suggesting one of the examples, and to John Simms for some helpful discussions on set theory.

REFERENCES

[1] Barwise, J., ed., "Handbook of Mathematical Logic," *Studies in Logic and the Foundations of Mathematics*, vol. 90, North-Holland, 1977.

[2] Brouwer, L., *Collected Works, Vol. 1, Philosophy and Foundations of Mathematics*, edited by A. Heyting, North-Holland, 1975.

[3] van Dalen, D., H. Doets, and H. de Swart, *Sets: Naive, Axiomatic and Applied, International Series in Pure and Applied Mathematics*, vol. 106, Pergamon Press, 1978.

[4] Diaconescu, R., "Axiom of choice and complementation," *Proceedings of the American Mathematical Society*, vol. 51 (1975), pp. 176–178.

[5] Fourman, M., "The logic of topoi," pp. 1053–1090 in *Studies in Logic and the Foundations of Mathematics*, vol. 90, North Holland, 1977.

[6] Heyting, A., *Intuitionism, an Introduction, Studies in Logic and the Foundations of Mathematics*, vol. 34, North-Holland, 1956.

[7] Johnstone, P., "Topos theory," *London Mathematical Society Monographs*, vol. 10, Academic Press, 1977.

[8] Johnstone, P., *Notes on Logic and Set Theory*, Cambridge University Press, 1987.

[9] Lambek, J. and P. Scott, *Introduction to Higher Order Categorical Logic*, *Cambridge Studies in Advanced Mathematics*, 7, Cambridge University Press, 1986.

[10] Mines, R., F. Richman, and W. Ruitenburg, *A Course in Constructive Algebra*, *Universitext*, Springer, 1988.

Department of Mathematics, Statistics and Computer Science
Marquette University
Milwaukee, WI 53233

4 Appendix: A topos model We construct a topos $\mathcal{E}$ whose natural number object $\mathbf{N}$ has a detachable subset X such that both X and $\mathbf{N}\setminus X$ are not infinite, where a subset $Y \subseteq \mathbf{N}$ is infinite if for all m there exists $n > m$ such that $n \in Y$. We hasten to add that the construction of the topos model itself uses principles from classical logic and set theory.

All languages that we consider are for a higher-order logic as described in Fourman [5] or Lambek and Scott [9], with additional type constants and function constants. We construct a sequence of higher-order languages L_i, theories $\{T_i \mid i \in \mathbf{N}\}$ for the languages L_i, and topos models $\{\mathcal{E}_i \mid i \in \mathbf{N}\}$ for the theories T_i.

Let L_0 be the language with extra type constant N, extra function symbol $s: N \to N$, and extra constant symbol 0 of type N. Let T_0 be the theory of higher-order logic for L_0 with the Axiom Schema of Choice (epimorphisms split), implying excluded middle Diaconescu [4], and the additional schema: $(N, s, 0)$ is a natural number object in L_0 ([5] or Johnstone [7] or [9]). Obviously, T_0 has a topos model contained in the category of sets $\mathcal{S}$. Define exp_λ for all ordinals λ by $exp_0 = \aleph_0$, $exp_{\alpha+1} = 2^{exp_\alpha}$, and $exp_\lambda = \bigcup_{\alpha<\lambda} exp_\alpha$ for limit ordinals λ. Set $\mathcal{E}_0 = V_\lambda$ with λ a regular cardinal bigger than exp_ω, where V_λ is an initial segment of the cumulative hierarchy (see van Dalen [3], p. 168, or Johnstone [8], p. 71).

Suppose L_i and T_i have been defined and a model $\mathcal{E}_i$ constructed. Define L_{i+1} as the extension of L_i obtained by adding constant symbols for all elements of the natural number object $N_i \in |\mathcal{E}_i|$, plus one more symbol c_{i+1}. Define T_{i+1} as the extension of T_i by adding all properties for the constants satisfied by the corresponding elements of N_i in $\mathcal{E}_i$, plus the axiom schema $c_{i+1} > n$ for all constants n of N_i. Set $\mathcal{E}_{i+1} = \mathcal{E}_i^{N_i}/F$, where F is an ultrafilter on N_i that exists and is free in $\mathcal{E}_i$. The category $\mathcal{E}_{i+1}$ is a subcategory of $\mathcal{E}_i$ with embedding $\sigma_i : \mathcal{E}_{i+1} \to \mathcal{E}_i$, and is a topos with natural number object $N_{i+1} = N_i^{N_i}/F$. For c_{i+1} choose the diagonal element $(\mathrm{id} : N_i \to N_i)/F$. Then $\mathcal{E}_{i+1}$ is a model of T_{i+1}.

Consider the sequence of categories

$$\dots \xrightarrow{\sigma_2} \mathcal{E}_2 \xrightarrow{\sigma_1} \mathcal{E}_1 \xrightarrow{\sigma_0} \mathcal{E}_0,$$

where the σ_i are the inclusion functors. Note that the σ_i are left exact. We use the glueing construction as described in [7], p. 109, to construct a new topos. Let $\mathcal{E} = \prod_i \mathcal{E}_i$. Let $\mathbf{G} = (G, \epsilon, \delta)$ be the comonad on $\mathcal{E}$ defined by

$$G\Big(\prod_i A_i\Big) = \prod_i \prod_{j\geq i} A_j;$$

$$^{\varepsilon}\textstyle\prod_i A_i = \displaystyle\prod_i \pi_1 : G\Big(\prod_i A_i\Big) \to \prod_i A_i; \text{ and}$$

$$^{\delta}\textstyle\prod_i A_i = \displaystyle\prod_i \prod_{j\geq i} \prod_{k\geq j} \pi_k : G(A) \to G^2(A).$$

The functor G is left exact. So by ([7], Theorem 2.32) the category $\mathcal{E}_{\mathbf{G}}$ of co-algebras is a topos.

The objects of $\mathcal{E}_{\mathbf{G}}$ are most easily described as sequences

$$A = (A_0 \xrightarrow{a_0} A_1 \xrightarrow{a_1} A_2 \xrightarrow{a_2} \dots),$$

where $A_i \in |\mathcal{E}_i|$ and a_i is a morphism of $\mathcal{E}_i$. Morphisms $f: A \to B$ consist of sequences $f = (f_0, f_1, f_2, \dots)$, where the $f_i: A_i \to B_i$ are such that $b_{i+1} f_i = f_{i+1} a_i$. We easily see that $N = (N_0, N_1, N_2, \dots)$ is the natural number object of $\mathcal{E}_{\mathbf{G}}$. Let $X = (X_0, X_1, X_2, \dots)$ be the subobject of N defined by $X_{2i} = \{n \in N_{2i} \mid c_{2j-1} \leq n \leq c_{2j}$ for some $j \leq i\}$ and $X_{2i+1} = \{n \in N_{2i+1} \mid c_{2j-1} \leq n \leq c_{2j}$ for some $j \leq i$ or $c_{2i+1} \leq n\}$. We easily verify:

Theorem 4.1 *X is a detachable subobject of N such that neither X nor $N \backslash X$ is infinite.*