On being the "same" or "different": Introduction to Apartness

wren gayle romano

7 August 2014

1 Introduction

We often talk about values being "the same as" or "different from" one another. But how can we formalize these notions? In particular, how should we do so in a constructive setting?

Constructively, we lack a general axiom for double-negation elimination; therefore, every primitive notion gives rise to both strong (strictly positive) and weak (doubly-negated) propositions. Thus, from the denial of (weak) difference we can only conclude weak sameness. Consequently, in the constructive setting it is often desirable to take difference to be a primitive— so that, from the denial of strong difference we can in fact conclude strong sameness.

This ability "un-negate" sameness is the principal reason for taking difference to be one of our primitive notions. While nice in and of itself, it also causes the strong and weak notions of sameness to become logically equivalent (thm 1.4); enabling us to drop the qualifiers when discussing sameness.

But if not being different is enough to be considered the same, then do we still need sameness to be primitive? To simplify our reasoning, we may wish to have sameness be *defined* as the lack of difference. However, this is not without complications. Sameness has been considered a natural primitive for so long that it has accrued many additional non-propositional properties (e.g., the substitution principle). So, if we eliminate the propositional notion of primitive equality, we will need somewhere else to hang those coats.

The rest of the paper fleshes out these various ideas.

1.1 Preliminary definitions

We write $\_ = \_$ ***"(strictly) equal"*** for the (potentially primitive) notion of being the same, $\_ \neq \_$ ***"unequal"*** for the derived notion of not being equal, and $\_ \approx \_$ ***"weakly equal"*** for the derived notion of being not-not equal. That is, we assume the following standard definitions:[1]

$$x \neq y := \neg(x = y) \tag{1}$$
$$x \approx y := \neg(x \neq y) \tag{2}$$

In addition, we write $\_\#\_$ ***"(strictly) apart"*** for the (potentially undefined) primitive notion of being different. Consider a collection of atomic things. If we can split them apart, then clearly they're not all the same thing. Moreover, by assumption, if they cannot be split apart then they are in fact the same. That is, we take the following to be axiomatic whenever $\_\#\_$ is defined:

$$\forall x, y.\ x = y \leftrightarrow \neg(x \# y) \tag{3}$$

However, given unequal things, we may not necessarily be able to split them. For example, suppose we are given a bag of marbles and are told (by a trusted source) that the marbles have different colors. Our job is to sort the marbles into piles according to color. But, as it turns out, we're color-blind! So even though we know the colors are different, we have no way of using that knowledge to actually split the collection.

Informally, we can think about $\_\#\_$ as proving that a splitting exists (i.e., constructively: by actually exhibiting one). Whereas the double-negation of $\_\#\_$ is weaker in the usual way: it's easier to prove since we need not produce the witness, but it's less useful since we cannot extract a witness from the proof.

1.2 Why is this desirable?

Definition 1.1. A formula ϕ is called ***stable*** whenever we have that $\forall \vec{x}.\ \phi[\vec{x}] \leftrightarrow \neg\neg\phi[\vec{x}]$.

Lemma 1.2. *Any formula of the form $\neg\phi$ is stable.*

Corollary 1.3. *Denial of $\_\#\_$ is stable*

Theorem 1.4. *If $\_\#\_$ is defined, then $\_=\_$ is stable.*

Proof. For any x and y,

$$\begin{array}{rcll}
x = y & \leftrightarrow & \neg(x \# y) & (3) \\
\neg(x = y) & \leftrightarrow & \neg\neg(x \# y) & \text{MT} \\
\neg\neg(x = y) & \leftrightarrow & \neg\neg\neg(x \# y) & \text{MT} \\
\neg\neg(x = y) & \leftrightarrow & \neg(x \# y) & \text{compose with lemma 1.3} \\
\neg\neg(x = y) & \leftrightarrow & x = y & \text{compose with (3)}
\end{array}$$

∎

[1] N.B., some authors use the symbol $\_\neq\_$ to denote a tight apartnesses, whereas we will be writing tight apartnesses as $\_\#\_$ (not distinguishing them from non-tight apartnesses). We stick with the older convention, using $\_\neq\_$ to denote the denial inequality, to reduce confusion among readers not familiar with this literature. Moreover, should the need to symbolically distinguish tight and non-tight apartnesses arise, we suggest using $\_\#\_$ or $\_\not\neq\_$ for non-tight apartnesses (depending on the slash convention), mirroring the standard practice of using $\_\equiv\_$ for equivalences.

The stability of $\_=\_$ proves it is logically-equivalent to $\_\approx\_$. This is sufficient if we are only concerned with provability— but constructively we also care about the proofs themselves. Thus, let's consider propositions to name the collection of their proofs. The stability proof demonstrates that $x = y$ is inhabited just in case $x \approx y$ is. Moreover, the proof turns out to also show that $x = y$ and $x \approx y$ are (category-theoretically) equivalent[2] as types. Still further, if $\_=\_$ is weakly decidable then the equivalence is in fact an isomorphism. In section 3 we seek to further strengthen the equivalence to an equality by defining $\_=\_$ in terms of $\_\#\_$, rather than considering $\_=\_$ to be a primitive.

2 Take 1

2.1 Partial Equivalence

A ***partial equivalence relation (PER)*** is any binary relation, E, with the following properties. In fact, there is some redundancy in this definition. Given (4), the other three properties are logically-equivalent.

$$\textbf{\textit{symmetric}} \quad \forall x,y.\ E(x,y) \to E(y,x) \tag{4}$$
$$\textbf{\textit{transitive}} \quad \forall x,y,z.\ E(x,y) \wedge E(y,z) \to E(x,z) \tag{5}$$
$$\textbf{\textit{right-euclidean}} \quad \forall x,y,z.\ E(x,y) \wedge E(x,z) \to E(y,z) \tag{6}$$
$$\textbf{\textit{left-euclidean}} \quad \forall x,y,z.\ E(x,y) \wedge E(z,y) \to E(x,z) \tag{7}$$

2.2 Total Equivalence

A ***total equivalence*** relation is any PER, E, which additionally has the following property. In fact, there is some redundancy in this definition. To prove that E is a total equivalence it is sufficient to prove just (4,5,8a) or (6,8a) or (7,8a).

$$\textbf{\textit{reflexive}} \quad \forall x.\ E(x,x) \tag{8a}$$

Total equivalence relations are not total in the sense of total orders, however they do enjoy the following totality properties.

$$\textbf{\textit{surjective}} \quad \forall y.\ \exists x.\ R(x,y) \tag{9}$$
$$\textbf{\textit{entire}}\text{, aka }\textbf{\textit{serial}} \quad \forall x.\ \exists y.\ R(x,y) \tag{10}$$

[2]Let the components of our proof be the functions $f : \forall x,y.\ x = y \to x \approx y$ and $g : \forall x,y.\ x \approx y \to x = y$. The composition $g \circ f$ reduces to $(\lambda p.\ p)$— i.e., it is *exactly* the identity function. However, the composition $f \circ g$ only reduces to an equivalent function, not to the identity function itself. Consequently, f and g only form a (category-theoretic) equivalence, not an isomorphism.

2.3 Strict Equivalence

A ***strict equivalence*** relation is any PER, E, which additionally has the following property. We call these relations strict because they contain the strict equality relation.

$$\textbf{\textit{equi-reflexive}} \quad \forall x,y.\ x = y \rightarrow E(x,y) \tag{8b}$$

2.4 Proper Equivalence

A ***proper equivalence*** relation is any relation which is both a total equivalence and a strict equivalence. In general, (8a) and (8b) are incommensurable. However, in many settings we can prove one from the other, as demonstrated below.

Theorem 2.1. *If* $\_=\_$ *obeys the substitution principle, then* (8a) *entails* (8b).

Proof. Fix any x and y, and assume that $x = y$. By (8a) we have that $E(x,x)$. By the substitution principle with the schema $E(x,\_)$ we can use the proof of $x = y$ to rewrite the proof of $E(x,x)$ into a proof of $E(x,y)$. Therefore we have proven $\forall x,y.\ x = y \rightarrow E(x,y)$. ■

Remark. The standard definition of an “equivalence” is what we are calling a total equivalence. The standard definition omits (8b) because it is a trivial theorem in the standard setting. However, we highlight the fact that (8b) is required because it exposes the fact that the substitution principle is crucial to the proof. Later on when we move to settings where $\_=\_$ no longer obeys the substitution principle, we must include (8b) in the definition of equivalences, or else we must do without the inferences it provides.

Theorem 2.2. *Suppose our proof language is a dependently typed programming language. If the type family* $\_=\_$ *is defined such that only proofs of type* $x = x$ *can be produced, and if we have dependent case analysis on proofs of the* $\_=\_$ *type family, then* (8a) *entails* (8b).

Proof. Fix any x and y, and assume that $x = y$. Dependent case analysis on the proof of $x = y$ will unify x and y to a new variable z. We can now use (8a) to prove $R(z,z)$. Upon concluding the case analysis, our proof will be reconsidered as belonging to the type $R(x,y)$. ■

Remark. The overall idea of this proof is basically the same as for thm 2.1. The difference lies in the exact mechanics of how we get our proofs from (8a) to be of the appropriate type. The unification process behind dependent case analysis is a much simpler mechanism than the substitution principle, though it will be less familiar to most readers.

Theorem 2.3. *If* $\_=\_$ *satisfies* (8a)*, then* (8b) *entails* (8a) *(for any relation).*

Theorem 2.4. *If the meta-theoretical propositions* $(*)$ *and* $(**)$ *are valid, then* $\_=\_$ *satisfies* (8a).

For all α, α is definitionally-equal to α. $(*)$

For all α and β, if α is definitionally-equal to β then α is propositionally-equal to β. $(**)$

Proof. “Definitional-equality” refers to our meta-theoretical notion of equality, whereas “propositional-equality” refers to the $\_ = \_$ relation in our theory. This distinction is common in dependently typed programming languages, and so we make use of their terminology here even though it does not necessarily make sense outside of context.

For all x, we have that x is definitionally-equal to itself via $(*)$ at $\alpha := x$. From this, we have that x is propositionally-equal to itself via $(**)$ at $\alpha := x$; $\beta := x$. Thus, we have proven $x = x$ for all x. And this is all we need to show in order to prove $\forall x.\ x = x$. ■

This is by no mean an exhaustive list of how one might be able to prove (8b) from (8a), or vice versa. It is just a short list of the most common justifications. In settings with different metatheoretical machinery there may be different proofs, or no proofs at all. We only wish to highlight the need for some such machinery if we do not take both (8a) and (8b) to be equally axiomatic.

2.5 Equality

An ***equality*** relation is any proper equivalence, E, which additionally satisfies either of the following two properties (which are logically-equivalent by (4)):

$$\textbf{\textit{antisymmetric}} \quad \forall x,y.\ E(x,y) \land E(y,x) \to x = y \tag{11}$$

$$\textbf{\textit{co-reflexive}} \quad \forall x,y.\ E(x,y) \to x = y \tag{12}$$

Notably, for the definition of equality relations, we need only assume that there is some relation $\_ = \_$ —we do not need it to have any properties in particular.

Theorem 2.5. *If any equality relation exists, then $\_ = \_$ is an equality.*

Proof. By (8b) and (12), $\_ = \_$ is logically-equivalent to that equality. ■

So whenever it’s worth talking about equality relations, we know we have one; which is nice. Unfortunately, this tells us nothing about what it *means* to be an equality, to be a notion of sameness. We’re saying there is some way to make finer distinctions than equivalence gives us, but we cannot say what that way is. *How* does being an equality allow finer distinctions? What distinctions are being made? Assuming an equality exists is begging the question!

It does, however, make one thing clear. Of all the strict equivalences, equalities make the *finest* distinctions. This is guaranteed by (8b). Supposing we could make distinctions finer than $\_ = \_$ recognizes, any strict equivalence must erase those distinctions by relating all the points that $\_ = \_$ relates. Thus, by definition, any relation making finer distinctions cannot be a strict equivalence. Now, whether that is desirable or not is another matter. Our reason for defining total equivalences is to allow us to discuss equivalences which may be finer than equalities.

When it comes to provability, one equality is as good as any other. However, as discussed in section 1.2, just because E and $\_ = \_$ are logically-equivalent, that doesn’t mean they’re category-theoretically equivalent. For example, we typically assume the proposition $x = y$ has at most one proof. Whereas, $E(x,y)$ could quite easily admit more than one proof; indeed, equality relations regularly do! Thus, in proof-relevant contexts (e.g., computational contexts) it is still important to distinguish equality relations from one another.

2.6 (Weak) Apartness

A ***(weak) apartness*** relation is any binary relation, R, satisfying the following three properties. In the intuitionistic setting (13,14) imply (4′); however, in minimal logic this is not the case, so (4′) is required for the definition.

$$\textbf{\textit{irreflexive}} \quad \forall x.\ \neg R(x,x) \tag{13}$$

$$\text{symmetric} \quad \forall x,y.\ R(x,y) \to R(y,x) \tag{4$'$}$$

$$\textbf{\textit{co-transitive}} \quad \forall x,y,z.\ R(x,z) \to R(x,y) \vee R(y,z) \tag{14}$$

Mirroring the distinction between (8a) and (8b), we could also introduce the property:

$$\textbf{\textit{equi-irreflexive}} \quad \forall x,y.\ x = y \to \neg R(x,y) \tag{13b}$$

However, this property isn't terribly interesting. In the one direction: if $\_ = \_$ satisfies (8a), then (13b) entails (13). And in the usual setting where (8a) entails (8b), the same mechanism can be used to prove that (13) entails (13b).

Theorem 2.6. *For any apartness R, the relation $E(x,y) := \neg R(x,y)$ is a total equivalence.*

Proof. Under the definition of E: (8a) is identical to (13); (4) is modus tollens of (4′); and (5) is equivalent to modus tollens of (14) by De Morgan's law. ■

Lemma 2.7. *The equivalence induced by an apartnesses is stable.*

2.7 Tight Apartness

Assuming $\_ = \_$ is an equality, we say that an apartness is ***tight***[3] just in case it satisfies either of the following properties (which are logically-equivalent by (4′)):

$$\textbf{\textit{tight}} \quad \forall x,y.\ \neg R(x,y) \to x = y \tag{15}$$

$$\textbf{\textit{connected}} \quad \forall x,y.\ \neg(R(x,y) \vee R(y,x)) \to x = y \tag{16}$$

Theorem 2.8. *An apartness is tight iff the induced equivalence is a (possibly improper) equality.*

Proof. For the equivalence $\neg R$: (15) is identical to (12); and (16) is equivalent to (11) by De Morgan's law. ■

[3] N.B., some authors refer to this definition as "strict apartness" or "strong apartness", however both are non-standard. Moreover, these terms are better used to describe other structures (respectively: the primitive apartness, and strong apartnesses of section 3).

Unfortunately, with the above definitions, the tightness property assumes a primitive notion of equality. Consequently, we cannot define our notion of equality in terms of the notion of tight apartness, since we can't formulate the tightness condition which distinguishes equalities from equivalences. Attempting to sidestep this problem by requiring $\neg R$ to be antisymmetric instead of requiring R to be connected will not work, since the definition of antisymmetry also requires primitive equality.

3 Take 2

To eliminate our reliance on primitive equality, we can rephrase things in terms of a primitive apartness. For clarity, whenever referring to some property P we have redefined, we will say $\mathrm{P}^{(old)}$ to refer to the old/original definition and will say $\mathrm{P}^{(new)}$ to refer to the new/current definition.

We assume that $\_\#\_$ is an apartness$^{(old)}$:

$$\text{irreflexive} \quad \forall x.\ \neg(x \# x) \tag{$13^\sharp$}$$

$$\text{symmetric} \quad \forall x,y.\ x \# y \to y \# x \tag{$4^\sharp$}$$

$$\text{co-transitive} \quad \forall x,y,z.\ x \# z \to x \# y \vee y \# z \tag{$14^\sharp$}$$

3.1 (Strong) Apartness

A ***(strong) apartness*** is any relation, R, satisfying:

$$\textbf{\textit{strongly irreflexive}} \quad \forall x,y.\ R(x,y) \to x \# y \tag{13^*}$$

$$\text{symmetric} \quad \forall x,y.\ R(x,y) \to R(y,x) \tag{$4'$}$$

$$\text{co-transitive} \quad \forall x,y,z.\ R(x,z) \to R(x,y) \vee R(y,z) \tag{14}$$

Lemma 3.1. *$\_\#\_$ is an apartness$^{(new)}$.*
Proof. (13^*) is the identity function; ($4'$) is ($4^\sharp$); (14) is ($14^\sharp$). ■

Lemma 3.2. *Every apartness$^{(new)}$ is irreflexive, and hence an apartness$^{(old)}$.*
Proof. By composing (13^*) with ($13^\sharp$). ■

Consider the following definition:

$$x = y \ := \neg(x \# y) \tag{17}$$

If we take (17) to be true, then (3) follows immediately, and so does not need to be taken as the defining characteristic of $\_\#\_$. Requiring (17) to hold is stronger than requiring (3) to hold, but it does allow us to completely eliminate the need for primitive $\_=\_$.

Lemma 3.3. *Under definition* (17)*, every apartness$^{(new)}$ is equi-irreflexive.*
Proof. Under the definition, (13b) is modus tollens of (13^*). ■

3.2 Tight (Strong) Apartness

An apartness$^{(new)}$ R is called ***tight*** just in case it satisfies either of the following properties (which are logically-equivalent by (4′)):

$$\textbf{\textit{strongly tight}} \quad \forall x,y.\ x \# y \to R(x,y) \tag{15*}$$
$$\textbf{\textit{strongly connected}} \quad \forall x,y.\ x \# y \to R(x,y) \lor R(y,x) \tag{16*}$$

Theorem 3.4. *$\_\#\_$ is a tight$^{(new)}$ apartness.*

Proof. (15*) is the identity function. ■

Theorem 3.5. *Under definition* (17)*, every tight$^{(new)}$ apartness is tight$^{(old)}$.*

Proof. Under the definition, (15,16) are modus tollens of (15*,16*) respectively. ■

3.3 Equality

An ***equality*** relation is any proper equivalence relation, E, which additionally satisfies either of the following two properties (which are logically-equivalent by (4)):

$$\text{antisymmetric} \quad \forall x,y.\ E(x,y) \land E(y,x) \to \neg(x \# y) \tag{11′}$$
$$\text{co-reflexive} \quad \forall x,y.\ E(x,y) \to \neg(x \# y) \tag{12′}$$

Corollary 3.6. *Under definition* (17)*, every equality$^{(new)}$ is an equality$^{(old)}$.*

Corollary 3.7. *If R is tight$^{(new)}$ then $\neg R$ is an equality$^{(new)}$.*

However, just as defining equality relations in terms of $\_=\_$ begs the question of what it means to be an "equality", so too defining tight apartnesses in terms of $\_\#\_$ begs the question of what it means to be "tight".

4 Decidability

Lemma 4.1. *If E is an equality and R is an apartness, then $\forall x,y.\ E(x,y) \to \neg R(x,y)$.*

Proof. For any x and y,

$$\begin{array}{rr}
E(x,y) \to \neg(x \# y) & (12') \\
x \# y \to \neg E(x,y) & \text{CP} \\
R(x,y) \to x \# y & (13^*) \\
R(x,y) \to \neg E(x,y) & \text{composition} \\
E(x,y) \to \neg R(x,y) & \text{CP}
\end{array}$$

■

Lemma 4.2. *If E is a strict equivalence and R is a tight apartness, then* $\forall x,y.\ \neg R(x,y) \to E(x,y)$.

Proof. By composing (15) with (8b). ■

Lemma 4.3. *If there exists a tight apartness R, then every equality E is stable.*

Proof. For any x and y,

$$\begin{array}{rr} E(x,y) \to \neg R(x,y) & \text{lemma 4.1} \\ R(x,y) \to \neg E(x,y) & \text{CP} \\ \neg\neg E(x,y) \to \neg R(x,y) & \text{MT} \\ \neg R(x,y) \to E(x,y) & \text{lemma 4.2} \\ \neg\neg E(x,y) \to E(x,y) & \text{composition} \end{array}$$
■

Theorem 4.4. *For any equality E and tight apartness R, the following properties are equivalent:*

(A) R is weakly decidable

(B) E is weakly decidable

(C) E is decidable

Proof. By double-negation introduction, if a relation is decidable then it is weakly decidable. Thus, C entails B; and if A entails C (shown below) then A entails B. For the remaining four implications, we give explicit proof terms. We have the following theorems of minimal logic:

$$\begin{array}{rcl} \mathit{exchange} & : & \forall P,Q.\ (\forall \vec{x}.\ P\vec{x} \vee Q\vec{x}) \to (\forall \vec{x}.\ Q\vec{x} \vee P\vec{x}) \\ \_\vee\_ & : & \forall P,P',Q,Q'.\ (\forall \vec{x}.\ P\vec{x} \to P'\vec{x}) \to (\forall \vec{x}.\ Q\vec{x} \to Q'\vec{x}) \to (\forall \vec{x}.\ P\vec{x} \vee Q\vec{x}) \to (\forall \vec{x}.\ P'\vec{x} \vee Q'\vec{x}) \\ \_\circ\_ & : & \forall P,Q,R.\ (\forall \vec{x}.\ Q\vec{x} \to R\vec{x}) \to (\forall \vec{x}.\ P\vec{x} \to Q\vec{x}) \to (\forall \vec{x}.\ P\vec{x} \to R\vec{x}) \\ \mathit{id} & : & \forall P.\ (\forall \vec{x}.\ P\vec{x} \to P\vec{x}) \\ \mathit{CP} & : & \forall P,Q.\ (\forall \vec{x}.\ P\vec{x} \to \neg Q\vec{x}) \to (\forall \vec{x}.\ Q\vec{x} \to \neg P\vec{x}) \\ \mathit{MT} & : & \forall P,Q.\ (\forall \vec{x}.\ P\vec{x} \to Q\vec{x}) \to (\forall \vec{x}.\ \neg Q\vec{x} \to \neg P\vec{x}) \end{array}$$

and the lemmas 4.1 and 4.2:

$$\begin{array}{rcl} \mathit{e2nr} & : & \forall x,y.\ E(x,y) \to \neg R(x,y) \\ \mathit{nr2e} & : & \forall x,y.\ \neg R(x,y) \to E(x,y) \end{array}$$

Thus, using the types:

$$\begin{array}{rcl} A & : & \forall x,y.\ \neg\neg R(x,y) \vee \neg R(x,y) \\ B & : & \forall x,y.\ \neg\neg E(x,y) \vee \neg E(x,y) \\ C & : & \forall x,y.\ E(x,y) \vee \neg E(x,y) \end{array}$$

the four remaining proof terms are as follows:

$$\begin{aligned} C &:= (nr2e \vee MT\ e2nr)(exchange\ A) \\ A &:= (MT\ nr2e \vee MT(CP\ e2nr))(exchange\ B) \\ C &:= ((nr2e \circ MT(CP\ e2nr)) \vee id)\ B \\ A &:= (MT\ nr2e \vee e2nr)(exchange\ C) \end{aligned}$$

■