

Generalizations of the Isomorphism Theorems

Ari Feiglin

In this lecture we will discuss certain generalizations of the three isomorphism theorems from algebra, and the relation of these generalizations to their algebraic counterparts. The lecture will discuss certain basic notions of universal algebra, without delving too deep.

1 The Isomorphism Theorems

Recall the following three group isomorphism theorems:

Theorem 1.1 (The First Group Isomorphism Theorem)

Let $f: G \longrightarrow H$ be a group homomorphism, then $G/\ker f \cong \text{im } f$.

Theorem 1.2 (The Second Group Isomorphism Theorem)

Let G be a group, $H \leq G$ a subgroup, and $N \trianglelefteq G$ a normal subgroup. Then N is normal in HN and $H \cap N$ is normal in H and

$$HN/N \cong H/(H \cap N)$$

Theorem 1.3 (The Third Group Isomorphism Theorem)

Let G be a group, $N \trianglelefteq G$ a normal subgroup. Then subgroups of G/N are precisely H/N for $N \leq H \leq G$, and normal subgroups are K/N for $N \leq K \trianglelefteq G$, furthermore

$$G/N / K/N \cong G/K$$

This family of three theorems have corresponding results in ring theory:

Theorem 1.4 (The First Ring Isomorphism Theorem)

Let $f: R \longrightarrow S$ be a ring homomorphism, then $R/\ker f \cong \text{im } f$.

Theorem 1.5 (The Second Ring Isomorphism Theorem)

Let R be a group, $S \leq R$ a subring, and $I \trianglelefteq R$ an ideal. Then I is an ideal in $S + I$ and $S \cap I$ is an ideal in S and

$$S + I / I \cong S / S \cap I$$

Theorem 1.6 (The Third Ring Isomorphism Theorem)

Let R be a ring, $I \trianglelefteq R$ an ideal. Then subrings of R/I are precisely S/I for $I \leq S \leq R$, and ideals are J/I

for $I \leq J \trianglelefteq R$, furthermore

$$R/I \big/_{K/I} \cong R/J$$

The statements and proofs of the ring isomorphism theorems are very similar to those of the group isomorphism theorems. This begs the question, can we generalize these theorems enough to encompass groups, rings, modules, and more? We turn to universal algebra for this.

2 Signatures and Structures

We begin by defining what signatures and structures are:

Definition 2.1

A **signature** is an object σ consisting of two disjoint sets: a set \mathcal{F} of function symbols, and \mathcal{R} of relation symbols. In addition, it has a function $\text{ar}: (\mathcal{F} \cup \mathcal{R}) \longrightarrow \mathbb{N}_{\geq 0}$, called the **arity function**. Functions with zero arity are called **constants**, and relations with zero arity are called **propositions**.

Definition 2.2

Let σ be a signature, then a σ -**structure** is an object \mathcal{A} consisting of the following three items:

- (1) A non-empty set A , called the **domain** of \mathcal{A} .
- (2) For every function symbol $f \in \sigma$, a function $f^{\mathcal{A}}: A^{\text{ar } f} \longrightarrow A$.
- (3) For every relation symbol $r \in \sigma$, a relation $r^{\mathcal{A}} \subseteq A^{\text{ar } r}$.

For example, let $\sigma = \{\circ\}$ where \circ is a binary function symbol. Then σ -structures are *magmas*, this includes semigroups, monoids, and groups. We can also have a signature $\sigma = \{\circ, {}^{-1}, 1\}$, where ${}^{-1}$ is a unary function symbol, and 1 is a constant.

So for example we can define a model $\mathcal{Z} = (\mathbb{Z}, +, -, 0)$ as such a σ -model.

We can also take a signature of ordered groups $\sigma = \{\circ, <\}$ where $<$ is a binary relation. \mathcal{Z} with the standard order is a σ -structure.

All these examples are pretty well structured, they satisfy some axioms regarding the signature (e.g. $\forall x, y, z. (x \circ y) \circ z = x \circ (y \circ z)$). But we need not have a σ -structure satisfy anything, all we require is that it associates to each a symbol a function or relation. What the structure satisfies is the subject of mathematical logic and beyond the scope of this lecture.

Definition 2.3

Let \mathcal{A} be a σ -structure and $B \subseteq A$ such that for every $f \in \sigma$ of arity n and $\vec{b} \in B^n$, $f^{\mathcal{A}}(\vec{b}) \in B$, then we can make B into a σ -structure where $f^B = f^{\mathcal{A}}|_{B^n}$ and $r^B = r^{\mathcal{A}} \cap B^{\text{ar } r}$. Such a structure is called a **substructure** of \mathcal{A} .

So for example, the only substructures of \mathcal{Z} are subgroups $n\mathbb{Z}$. But if we take \mathcal{Z} over the reduced signature $\mathcal{Z} = (\mathbb{Z}, +)$ then \mathbb{N} is also a substructure. So the signature of a structure is very significant.

To state the isomorphism theorems, we need some sense of a homomorphism:

Definition 2.4

Let \mathcal{A} and \mathcal{B} be σ -structures, then a **homomorphism** from \mathcal{A} to \mathcal{B} is a function $h: A \longrightarrow B$ such that

- (1) For every $f \in \sigma$ and $\vec{a} \in A^n$, $h(f^{\mathcal{A}}\vec{a}) = f^{\mathcal{B}}(h\vec{a})$ where $h\vec{a} = (ha_1, \dots, ha_n)$. For constant symbols

this means $hc^A = c^B$.

- (2) For every $r \in \sigma$ then $r^B h\vec{a}$ if and only if there exists a $\vec{c} \in A^n$ such that $h\vec{a} = h\vec{c}$ and $r^A \vec{c}$. Ignore the definition for propositional symbols.

Bijjective homomorphisms are isomorphisms, etc.

Notice that if h is injective, then the condition on relations just says $r^B h\vec{a} \iff r^A \vec{a}$.

Lemma 2.5

The image of a homomorphism is a substructure of the codomain.

We also need some notion of quotienting:

Definition 2.6

Let \mathcal{A} be a σ -structure, then a **congruence** on \mathcal{A} is an equivalence relation \sim on A such that for every $f \in \sigma$, if $\vec{a} \sim \vec{b}$ (meaning $a_i \sim b_i$ for every i), then $f^A \vec{a} \sim f^A \vec{b}$.

If \sim is a congruence on \mathcal{A} then define the **quotient structure** \mathcal{A}/\sim as follows:

- (1) The domain is the partition A/\sim .
- (2) For $f \in \sigma$ and $a \in A$, $f^{A/\sim} \vec{a}_\sim = (f^A \vec{a})_\sim$, this is well defined by congruence.
- (3) For $r \in \sigma$, $r^{A/\sim} \vec{a}_\sim$ if and only if there exists a $\vec{b} \sim \vec{a}$ such that $r^A \vec{b}$.

Notice that if $N \trianglelefteq G$ is a normal subgroup, then $a \sim b \iff ab^{-1} \in N$ is a congruence. This is because if $(a_1, a_2) \sim (b_1, b_2)$ then $a_1 a_2 (b_1 b_2)^{-1} = a_1 a_2 b_2^{-1} b_1^{-1} \in a_1 N b_1^{-1} = a_1 b^{-1} N = N$. So $a_1 a_2 \sim b_1 b_2$, as required. Similar for $^{-1}$ (for those who know logic, this is because $^{-1}$ is definable by \circ).

Then the quotient group G/N is precisely the quotient structure G/\sim . Notice: $a_\sim b_\sim = (ab)_\sim$ which is the same as saying $aN bN = abN$, since $aN = a_\sim$.

3 The Isomorphism Theorems

Notice that if $h: \mathcal{A} \longrightarrow \mathcal{B}$ is a homomorphism, then we can define a congruence \sim_h on \mathcal{A} by $x \sim_h y$ if and only if $hx = hy$. This is indeed a congruence: if $\vec{x} \sim_h \vec{y}$ then $hf^A \vec{x} = f^B h\vec{x} = f^B h\vec{y} = hf^A \vec{y}$ so $f^A \vec{x} \sim_h f^A \vec{y}$. This congruence is called the *kernel* of the homomorphism.

Notice that if $h: G \longrightarrow H$ is a homomorphism of groups, then $a \sim_h b$ if and only if $ha = hb$ if and only if $h(ab^{-1}) = 1$, which is if and only if $ab^{-1} \in \ker h$. So $G/\ker h = G/\sim_h$ by our above discussion.

Theorem 3.1 (The First Isomorphism Theorem)

- (1) Let \sim be a congruence on \mathcal{A} , then $\rho: \mathcal{A} \longrightarrow \mathcal{A}/\sim$ defined by $a \mapsto a/\sim$ is an epimorphism.
- (2) Let $h: \mathcal{A} \longrightarrow \mathcal{B}$ be a homomorphism of σ -structures, then $\iota: \mathcal{A}/\sim_h \longrightarrow h\mathcal{A}$ defined by $a/\sim_h \mapsto ha$ is an isomorphism.

Proof:

- (1) This is obviously surjective. Now to prove it is a homomorphism:

$$\rho f\vec{a} = (f\vec{a})/\sim = f(\vec{a}/\sim) = f\rho\vec{a}$$

and $r\rho\vec{a}$ if and only if $r(\vec{a}/\sim)$ if and only if there exists a $\vec{b} \sim \vec{a}$ such that $r\vec{b}$, and this is if and only if $\rho\vec{b} = \rho\vec{a}$. So ρ is a homomorphism.

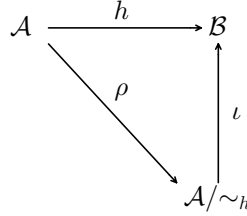
- (2) Notice that if $a/\sim_h = b/\sim_h$ then $ha = hb$ by definition, so ι is well-defined. ι is also trivially a bijection. It is a homomorphism since

$$\iota(f(\vec{a}/\sim_h)) = \iota(f\vec{a}/\sim_h) = hf\vec{a} = fh\vec{a} = f\iota(\vec{a}/\sim_h)$$

and

$$r\iota(\vec{a}/\sim_h) \iff rh\vec{a} \iff \exists \vec{b} \sim \vec{a}. r\vec{b} \iff r(\vec{a}/\sim_h) \quad \blacksquare$$

This gives us the classic first isomorphism theorem diagram:



This was pretty straightforward. Generalizing the next two isomorphism theorems requires a few more definitions.

Recall the second isomorphism theorem: if $H \leq G$ is a subgroup and $N \trianglelefteq G$ is a normal subgroup then $N \trianglelefteq HN \leq G$, $H \cap N \trianglelefteq H$, and $(HN)/N \cong H/(H \cap N)$. Our counterpart to the concept of normal subgroups in structures are congruences, which are relations on the underlying domain. So how do we define the product of a substructure and a congruence?

Well, notice that $a \in HN$ if and only if there exists an $h \in H$ such that $a \sim_N h$ (where \sim_N is the congruence induced by N). So we can also define $HN = \{a \in G \mid \exists h \in H. a \sim_N h\}$. We generalize this

Definition 3.2

Let \mathcal{A} be a σ -structure, $\mathcal{B} \leq \mathcal{A}$ a substructure, and \sim a congruence on \mathcal{A} . Define the \sim -closure of \mathcal{B} to be $\mathcal{B}_\sim = \{a \in A \mid \exists b \in B. a \sim b\}$.

And of course the induced congruence on H by $H \cap N$ is just $\sim_N \cap H^2$, i.e. $\sim_N|_H$. So we would expect

Theorem 3.3 (The Second Isomorphism Theorem)

Let σ be an algebraic signature, \mathcal{A} be a σ -structure, $\mathcal{B} \leq \mathcal{A}$ a substructure, and \sim a congruence on \mathcal{A} . Then

- (1) The \sim -closure \mathcal{B}_\sim is a substructure of \mathcal{A} ,
- (2) $\mathcal{B}_\sim/\sim \cong \mathcal{B}/\sim|_{\mathcal{B}}$. (The \sim under \mathcal{B}_\sim should be restricted to \mathcal{B}_\sim , but since $b \sim a \iff b \sim|_{\mathcal{B}} a$ for $b \in \mathcal{B}_\sim$, we omit the subscript.)

Proof:

- (1) Let $f \in \sigma$ and $\vec{a} \in \mathcal{B}_\sim$, then $\vec{a} \sim \vec{b} \in \mathcal{B}$, so $f\vec{a} \sim f\vec{b} \in \mathcal{B}$ so $f\vec{a} \in \mathcal{B}_\sim$.
- (2) Define

$$h: \mathcal{B} \longrightarrow \mathcal{B}_\sim/\sim, \quad hb = b/\sim$$

This is a homomorphism:

$$hf\vec{b} = (f\vec{b})/\sim = f(\vec{b}/\sim) = f(h\vec{b})$$

And its kernel is $a \sim_h b$ if and only if $a \sim b$, so the kernel is $\sim_h|_{\mathcal{B}}$ as required. \blacksquare

Why must the signature be algebraic? Well, otherwise h may not have been a homomorphism:

$$rh\vec{b} \iff r(\vec{b}/\sim) \iff \exists \vec{a} \sim \vec{b}. r\vec{a}$$

But suppose no $\vec{b} \in \mathcal{B}$ satisfies $r\vec{b}$, but there exists a $\vec{a} \sim \vec{b}$ such that $r\vec{a}$. Then we don't have that there exists a \vec{b}' such that $h\vec{b}' = h\vec{b}$ and $r\vec{b}'$, yet we do have $rh\vec{b}$.

For example, here is a simple counterexample: let our signature contain a single unary relation r and define $\mathcal{A} = \{a, b\}$, $\mathcal{B} = \{b\}$ where only $r^{\mathcal{A}}a$. Define $a \sim b$, so we have that $\mathcal{B}_{\sim} = \mathcal{A}$, so $\mathcal{B}_{\sim}/\sim = \{\{a, b\}\}$ and $\mathcal{B}/\sim|_{\mathcal{B}} = \{\{b\}\}$. But $r^{\mathcal{A}/\sim}\{a, b\}$ since $r^{\mathcal{A}}a$, but we don't have that $\mathcal{B}_{\sim}/\sim \cong \mathcal{B}/\sim|_{\mathcal{B}}$.

This shouldn't be surprising: quotienting should have weird effects on relations. For example, let us look at $(\mathbb{Z}, <)$: if we take the congruence \sim_n corresponding to $n\mathbb{Z}$ over \mathbb{Z} then $a/\sim_n < b/\sim_n$ if and only if there exists $a' \sim_n a$ and $b' \sim_n b$ such that $a' < b'$. But this is always true since the congruence classes are all infinite. So $<_{\mathbb{Z}/n\mathbb{Z}}$ is just the trivial relation.

This may not be all that surprising, after all how do you define an order on $\mathbb{Z}/n\mathbb{Z}$? But it does mean that while a structure may satisfy some theory, its quotient need not. Both of these examples have demonstrated this.

The issue is congruences don't take relations into account. Let us fix this.

Definition 3.4

A **total congruence** is a congruence \sim on a σ -structure \mathcal{A} such that for all \vec{a}, \vec{b} , if $\vec{a} \sim \vec{b}$ then $r\vec{a} \iff r\vec{b}$ for all $r \in \sigma$.

Notice that when quotienting by a total congruence, $r(\vec{a}/\sim) \iff r\vec{a}$ since if $\vec{a} \sim \vec{b}$ and $r\vec{b}$ then $r\vec{a}$ by definition.

Theorem 3.5 (The Second Isomorphism Theorem, Relational-Style)

Let \sim be a total congruence on \mathcal{A} (which need not be algebraic), and $\mathcal{B} \leq \mathcal{A}$ a substructure. Then

$$\mathcal{B}_{\sim}/\sim \cong \mathcal{B}/\sim|_{\mathcal{B}}$$

Proof: all we need to show is that the h defined before is a homomorphism. $r\vec{h}\vec{b}$ if and only if $r(\vec{b}/\sim)$ if and only if $r\vec{b}$. ■

For the third isomorphism theorem, we have the following: let $N \trianglelefteq G$, then

- (1) Every subgroup of G/N is of the form K/N for $N \leq K \leq G$ (and vice versa).
- (2) Every normal subgroup of G/N is of the form K/N for $N \leq K \trianglelefteq G$ (and vice versa).
- (3) If $N \leq K \trianglelefteq G$ then $(G/N)/(K/N) \cong G/K$.

But here $(G/N)/(K/N)$ is a quotient structure of a quotient structure, so we must somehow deal with this. Notice that

$$aN \cdot K/N = bN \cdot K/N \iff ab^{-1}N \in K/N \iff ab^{-1} \in K$$

so we have that $a/\sim_N \sim_{K/N} b/\sim_N$ if and only if $a \sim_K b$.

Definition 3.6

Let \mathcal{A} be a σ -structure, \sim_N and \sim_K be congruences such that $\sim_N \subseteq \sim_K$. Then define their **quotient congruence** to be $\sim_{K/N} = \sim_K/\sim_N$ over \mathcal{A}/\sim_N by

$$a/\sim_N \sim_{K/N} b/\sim_N \iff a \sim_K b$$

This is well-defined since if $a/\sim_N = b/\sim_N$ then $a \sim_K b$ as well. And this is obviously a congruence.

This helps us show the following:

Theorem 3.7 (The Third Isomorphism Theorem)

Let \mathcal{A} be a σ -structure, and \sim_N a congruence on \mathcal{A} .

- (1) Every substructure of \mathcal{A}/\sim_N is of the form \mathcal{B}/\sim_N for some substructure satisfying $\mathcal{B}_{\sim_N} = \mathcal{B}$ and vice versa.
- (2) Every congruence on \mathcal{A}/\sim_N is of the form $\sim_{K/N}$ for some congruence on \mathcal{A} , $\sim_K \supseteq \sim_N$, and vice versa.
- (3)

$$\mathcal{A}/\sim_N \Big/ \sim_{K/N} \cong \mathcal{A}/\sim_K$$

Proof:

- (1) Let \mathcal{C} be a substructure of \mathcal{A}/\sim_N , then define $B = \{a \in A \mid a/\sim_N \in C\}$. This defines a substructure of \mathcal{A} , since for $f \in \sigma$ and $\vec{a} \in B^n$, $(f\vec{a})/\sim_N = f(\vec{a})/\sim_N \in C$ so $f\vec{a} \in B$. And we can see that if $a \in B$ and $a \sim_N b$ then $a/\sim_N = b/\sim_N$, so $b/\sim_N \in C$ and so $b \in B$, thus $\mathcal{B}_{\sim_N} = \mathcal{B}$ as required. It is obvious that $\mathcal{B}/\sim_N = \mathcal{C}$.

And conversely, if $\mathcal{B}_{\sim_N} = \mathcal{B}$ then \mathcal{B}/\sim_N is well-defined (since \mathcal{B} contains all its equivalence classes), and so from definition it is a substructure.

- (2) Let \sim be a congruence on \mathcal{A}/\sim_N then define $a \sim_K b \iff (a/\sim_N) \sim (b/\sim_N)$, this obviously contains \sim_N . This a congruence: if $\vec{a} \sim_K \vec{b}$ then

$$(\vec{a}/\sim_N) \sim (\vec{b}/\sim_N) \implies f(\vec{a}/\sim_N) \sim f(\vec{b}/\sim_N) \implies (f\vec{a}/\sim_N) \sim (f\vec{b}/\sim_N) \implies f\vec{a} \sim_K f\vec{b}$$

And by definition $\sim_{K/N} = \sim$.

- (3) Let us define

$$h: \mathcal{A} \longrightarrow \mathcal{A}/\sim_N \Big/ \sim_{K/N}, \quad a \mapsto (a/\sim_N)/\sim_{K/N}$$

This is indeed a homomorphism:

$$hf\vec{a} = (f\vec{a}/\sim_N)/\sim_{K/N} = f((\vec{a}/\sim_N)/\sim_{K/N}) = fh\vec{a}$$

And $rh\vec{a}$ is equivalent to $r(\vec{a}/\sim_N)/\sim_{K/N}$ which is equivalent to there existing a $\vec{b}/\sim_N \sim_{K/N} \vec{a}/\sim_N$ such that $r(\vec{b}/\sim_N)$. This is equivalent to $\vec{b} \sim_K \vec{a}$ such that $r(\vec{b}/\sim_N)$. This is equivalent to there existing a $\vec{c} \sim_N \vec{b}$ such that $r\vec{c}$. Now, since $\vec{c} \sim_N \vec{b} \sim_K \vec{a}$, we have $\vec{c} \sim_K \vec{a}$ by transitivity, so $h\vec{c} = (\vec{c}/\sim_N)/\sim_{K/N} = (\vec{a}/\sim_N)/\sim_{K/N} = h\vec{a}$. So if $rh\vec{a}$ then there exists a \vec{c} with $h\vec{c} = h\vec{a}$ and $r\vec{c}$.

The converse is trivial (in general, since $h\vec{a} = h\vec{b}$ and $r\vec{b}$ means $rh\vec{b}$ so $rh\vec{a}$). ■

4 Applications

4.1 Rings

In this section, we will prove the well-known special cases of the isomorphism theorems for rings and modules. All we need to show is that congruences, closures, and quotient congruences can be associated with their corresponding notions in these objects.

So for rings, we need to show that congruences can be associated with ideals. Let R be a ring, \sim a congruence. Then we claim that $I = 0/\sim$ is an ideal: indeed $0 \in I$, and if $a, b \sim 0$ then $a + b \sim 0 + 0 = 0$. Finally if $r \in R$ and $a \in I$ then $r \sim r$ and $a \sim 0$ so $ra \sim r0 = 0$ so $ra \in I$, as required. Similar for ar .

Now if I is an ideal, then $a \sim b \iff a - b \in I$ is a congruence. This is because if f is a homomorphism such that $\ker f = I$, then

$$a \sim_f b \iff f(a) = f(b) \iff f(a - b) = 0 \iff a - b \in I$$

So $\sim_f = \sim$ and in particular \sim is a congruence.

Now suppose R is a ring, $S \leq R$ a subring, and I an ideal. Let \sim be the congruence associated with I , then we claim $S + I = S_\sim$. Indeed, if $s + i \in S + I$ then $s + i \sim s$ so $s + i \in S_\sim$, thus $S + I \subseteq S_\sim$. And if $s \sim r$ then $r = s + i$.

And finally if $I \trianglelefteq R$ is an ideal and $I \leq J$ is another ideal, then the congruence associated with J/I is the quotient congruence of J by the congruence of I . I.e. we have that $\sim_{J/I} = \sim^J / \sim_I$. So we must show that $a \sim_J b \iff (a / \sim_I) \sim_{J/I} (b / \sim_I)$. This is $a - b \in J \iff (a - b + I) \in J/I$ which is just $a - b \in J$ as required.

4.2 Boolean Algebras

Definition 4.1

A **Boolean algebra** is a $\{\wedge, \vee, \neg, 0, 1\}$ -structure satisfying the following axioms:

- (1) \vee, \wedge are associative and commutative
- (2) $a \vee (a \wedge b) = a$ and $a \wedge (a \vee b) = a$
- (3) \vee, \wedge distribute: $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$ and $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$
- (4) $a \vee 0 = a$ and $a \wedge 1 = a$
- (5) $a \vee \neg a = 1$ and $a \wedge \neg a = 0$

Suppose \mathcal{B} is a boolean algebra and \sim a congruence on it. Then let $I = 0 / \sim$ then it has the following properties:

- (1) If $a, b \in I$ then $a \vee b \sim 0 \vee 0 = 0$ so $a \vee b \in I$
- (2) If $a \in I$ and $x \in \mathcal{B}$ then $a \wedge x \sim 0 \wedge x = 0$ so $a \wedge x \in I$

These two properties define what is called a *Boolean Ideal*, and every boolean ideal defines a congruence: let I be a boolean ideal, then define $a \sim b \iff a \Delta b \in I$. It can be seen by computation that this is indeed a congruence. So there is a correspondence between boolean ideals and congruences. A *Prime Ideal* is a filter such that for every $a \in \mathcal{B}$ either a or $\neg a \in I$, equivalently it is a maximal ideal.

If we look at $F = 1 / \sim$, this has dual properties:

- (1) If $a, b \in F$ then $a \wedge b \sim 1 \wedge 1 = 1$ so $a \wedge b \in F$
- (2) If $a \in F$ and $x \in \mathcal{B}$ then $a \vee x \sim 1 \vee x = 1$ so $a \vee x \in F$.

A subset of \mathcal{B} satisfying these properties is called a *Filter*. A filter F defines a congruence by $a \sim b \iff a \leftrightarrow b \in F$. An *Ultrafilter* is a filter such that for every $a \in \mathcal{B}$, either a or $\neg a \in F$, equivalently it is a maximal filter.

Now, it can be shown that a substructure of a boolean algebra is itself a boolean algebra (for those who know logic, this is because the theory of boolean algebras is a \forall -theory). So let $\mathcal{C} \leq \mathcal{B}$ be a subboolean algebra, I a boolean ideal, and \sim its congruence. Then $\mathcal{C}_\sim = \mathcal{C} \Delta I = \{c \Delta i \mid c \in \mathcal{C}, i \in I\}$ since

$$x \in \mathcal{C}_\sim \iff \exists c \in \mathcal{C}. x \Delta c \in I \iff \exists c \in \mathcal{C}. i \in I. x = c \Delta i \iff x \in \mathcal{C} \Delta I$$

So we have that by the second isomorphism theorem,

$$\mathcal{C} / \mathcal{C} \cap I \cong \mathcal{C} \Delta I / I$$

And by the third isomorphism theorem, every subboolean algebra of \mathcal{B} / I is of the form \mathcal{C} / I for $I \subseteq \mathcal{C}$, and an ideal is of the form J / I for $I \subseteq J$. This is because if \sim is a congruence on \mathcal{B} / I then there exists some $J \supseteq I$ ideal such that $a / I \sim b / I \iff a \sim_J b$. So $a / I \sim 0 / I$ if and only if $a \sim_J 0$ if and only if $a \in J$ so the induced ideal is J / I . And then the third isomorphism says

$$\mathcal{B} / I / J / I \cong \mathcal{B} / J$$

5 Ideals

Notice that in all the objects discussed thus far there is some notion of an ideal. In each instance, an ideal is just a congruence class e/\sim which satisfies the following properties:

- (1) e/\sim uniquely defines the congruence, i.e. if $e/\sim = e'/\sim$ then $\sim = \sim'$.
- (2) For any substructure \mathcal{B} , $e/\sim \subseteq \mathcal{B}$ if and only if $\mathcal{B}_\sim = \mathcal{B}$.

How can we get these properties in a general structure? Notice that for all of these structures, and any congruence which defines an ideal I , $a \sim b$ if and only if $t(a, b) \in I$ for some function t which can be defined using the function symbols in the signature. Such a function is called a *term* (or more specifically, a *term function*).

Definition 5.1

Let σ be a signature and V some collection of variables (we assume such a collection is global, we don't really care too much what it is). Then a σ -**term** is defined recursively as follows:

- (1) Every constant $c \in \sigma$ is a term.
- (2) Every variable $x \in V$ is a term.
- (3) If $f \in \sigma$ has arity n , and t_1, \dots, t_n are terms then $ft_1 \cdots t_n$ is a term.

We denote the set of all σ -terms by \mathcal{T}_σ , and we omit the σ when the signature is understood. For a term $t \in \mathcal{T}_\sigma$, we write $t = t(\bar{x})$ to mean that the variables in t are contained in \bar{x} .

If \mathcal{A} is a σ -structure, $t(\bar{x})$ a σ -term, and $\bar{a} \in \mathcal{A}^n$, we define $t(\bar{a}) \in \mathcal{A}$ to be the result of substituting \bar{a} for \bar{x} in t . Recursively this is defined as follows:

- (1) If $t = c$ then $t(\bar{a}) = c^{\mathcal{A}}$.
- (2) If $t = x_i$ then $t(\bar{a}) = a_i$.
- (3) Otherwise $t = ft_1 \cdots t_n$ and so $t(\bar{a}) = f^{\mathcal{A}}t_1(\bar{a}) \cdots t_n(\bar{a})$.

Lemma 5.2

If $\mathcal{B} \leq \mathcal{A}$ are σ -structures, $t(\bar{x})$ a σ -term, if $\bar{b} \in \mathcal{B}$ then $t(\bar{b}) \in \mathcal{B}$.

Proof: this is done by term induction, it's not complicated. ■

Lemma 5.3

If \sim is a congruence on \mathcal{A} , $t(\bar{x})$ a term, and $\bar{a} \sim \bar{b} \in \mathcal{A}$ then $t(\bar{a}) \sim t(\bar{b})$.

Proof: again by term induction. ■

Definition 5.4

Call a σ -structure \mathcal{A} **idealized** if there exists a constant $e \in \sigma$, and two σ -terms \mathfrak{t} and $\bar{\mathfrak{t}}$ such that:

- (1) For every congruence \sim : $a \sim b \iff \mathfrak{t}(a, b) \in e/\sim$.
- (2) For every a, b , $\bar{\mathfrak{t}}(\mathfrak{t}(\bullet, b), b) = a$. This means that $\bar{\mathfrak{t}}(\mathfrak{t}(\bullet, b), b) = \text{id}$, so $\bar{\mathfrak{t}}(\bullet, b)$ is surjective and $\mathfrak{t}(\bullet, b)$ is injective for every $b \in \mathcal{A}$.
- (3) $\mathfrak{t}(\bullet, b)$ is surjective (so it is a bijection), this is equivalent to $\bar{\mathfrak{t}}(\bullet, b)$ being a bijection since their composition is the identity.

Call e/\sim an **ideal** of \sim .

- Groups are idealized: $\mathfrak{t}(a, b) = ab^{-1}$ and $\bar{\mathfrak{t}}(a, b) = ab$.
- Rings are idealized: $\mathfrak{t}(a, b) = a - b$ and $\bar{\mathfrak{t}}(a, b) = a + b$.
- Boolean algebras are idealized: $\mathfrak{t}(a, b) = \bar{\mathfrak{t}}(a, b) = a \triangle b$ and $e = 0$, or $\mathfrak{t}(a, b) = \bar{\mathfrak{t}}(a, b) = a \leftrightarrow b$ and $e = 1$.

Theorem 5.5

Let \mathcal{A} be idealized, then there is a bijective correspondence between congruences and ideals.

Proof: map \sim to e/\sim . This is obviously surjective by the definition of ideals, and it is injective since if $e/\sim = e'/\sim$ then

$$a \sim b \iff \mathfrak{t}(a, b) \in e/\sim = e'/\sim \iff a \approx b$$

so this correspondence is bijective. ■

This means that if I is an ideal of \mathcal{A} , we can write \mathcal{A}/I for \mathcal{A}/\sim where $I = e/\sim$ and there can be no confusion.

Proposition 5.6

If \mathcal{A} is idealized, then for any congruence \sim , all of its cosets (congruence classes) have the same cardinality.

Proof: notice that for the identity congruence, $\mathfrak{t}(b, b) \in e/=$ means $\mathfrak{t}(b, b) = e$. So $\mathfrak{t}(\bullet, b)$ maps $b/\sim \longrightarrow e/\sim$ since $\mathfrak{t}(b', b) \sim \mathfrak{t}(b, b) = e$. It is a bijection, so $|b/\sim| = |e/\sim|$. ■

Theorem 5.7

Let \mathcal{A} be idealized, then $\mathcal{B}_\sim = \bar{\mathfrak{t}}(e/\sim, \mathcal{B})$ for any set $\mathcal{B} \subseteq \mathcal{A}$. Therefore $e/\sim \subseteq \mathcal{B}$ if and only if $\mathcal{B}_\sim = \mathcal{B}$.

Proof: let $a \in \mathcal{B}_\sim$, so there exists a $b \in \mathcal{B}$ such that $a \sim b$. So $\mathfrak{t}(a, b) \in e/\sim$, then $\bar{\mathfrak{t}}(\mathfrak{t}(a, b), b) = a$ and this is in $\bar{\mathfrak{t}}(e/\sim, \mathcal{B})$. And for $\bar{\mathfrak{t}}(e', b) \in \bar{\mathfrak{t}}(e/\sim, \mathcal{B})$, notice that $\mathfrak{t}(\bar{\mathfrak{t}}(e', b), b) = e' \in e/\sim$. Thus $\bar{\mathfrak{t}}(e', b) \sim b \in \mathcal{B}$. So $\bar{\mathfrak{t}}(e', b) \in \mathcal{B}_\sim$.

So if $e/\sim \subseteq \mathcal{B}$ then $\mathcal{B}_\sim = \bar{\mathfrak{t}}(e/\sim, \mathcal{B}) \subseteq \bar{\mathfrak{t}}(\mathcal{B}, \mathcal{B}) \subseteq \mathcal{B}$. And since $\mathcal{B} \subseteq \mathcal{B}_\sim$ we have the required result. Conversely, since $e \in \mathcal{B}$ as it is a constant, we always have that $e/\sim \subseteq \mathcal{B}_\sim = \mathcal{B}$. ■

Notice that $\{a\}_\sim = a/\sim$, so we have that $a/\sim = \bar{\mathfrak{t}}(I, a)$. Thus we can view congruence classes as algebraic cosets:

$$\mathcal{A}/\sim = \{\bar{\mathfrak{t}}(I, a) \mid a \in \mathcal{A}\}$$

Let us write $I\mathcal{B}$ for $\bar{\mathfrak{t}}(I, \mathcal{B})$. Now, we'd like to state the following:

Theorem 5.8 (The Idealized Second Isomorphism Theorem)

Let \mathcal{A} be idealized, $\mathcal{B} \leq \mathcal{A}$ a substructure, and $I \trianglelefteq \mathcal{A}$ an ideal. Then I is an ideal of $I\mathcal{B}$ and $I \cap \mathcal{B}$ is an ideal of \mathcal{B} and

$$I\mathcal{B}/I \cong \mathcal{B}/I \cap \mathcal{B}$$

Unfortunately, just because \mathcal{A} is idealized doesn't mean that its substructures are. So writing $\mathcal{B}/(I \cap \mathcal{B})$ has no meaning. In order to get this result, we turn to model theory. The following is quite tedious and may have to be left for another lecture. But first let us generalize the third isomorphism theorem for ideals

Theorem 5.9 (The Idealized Third Isomorphism Theorem)

Let \mathcal{A} be idealized, $I \trianglelefteq \mathcal{A}$ an ideal. Then

- (1) \mathcal{A}/I is idealized.
- (2) Substructures of \mathcal{A}/I are of the form \mathcal{B}/I for $I \subseteq \mathcal{B}$.
- (3) Ideals of \mathcal{A}/I are of the form J/I for $I \subseteq J \trianglelefteq \mathcal{A}$.
- (4) $\mathcal{A}/I \big/_{J/I} \cong \mathcal{A}/J$

Proof: Let \sim_N be the congruence associated with I . Let $\sim_{K/N}$ be a congruence of \mathcal{A}/I , then Now suppose $t(a/I, b/I) \sim_{K/N} e/I$, that means that $t(a, b)/I \sim_{K/N} e/I$, so $t(a, b) \sim_K e$ meaning $a \sim_K b$. And so $a/I \sim_{K/N} b/I$. So \mathcal{A}/I is idealized, furthermore its ideals are $(e/I)/\sim_{K/N}$ and $e/I \sim_{K/N} e'/I$ if and only if $e \sim_K e'$, so the ideal is of the form J/I where J is the ideal associated with \sim_K . (2) and (4) follow immediately. ■

6 Theories Allowing For Ideals

Let us quickly define the groundwork for first order logic. We will gloss over the details due to lack of time.

Definition 6.1

Given a signature σ , we define \mathcal{L} the set of all formulas recursively as follows:

- (1) If t, s are σ -terms then $t = s \in \mathcal{L}$.
- (2) If $r \in \sigma$ is an n -ary relation and t_1, \dots, t_n are σ -terms then $rt_1 \dots t_n \in \mathcal{L}$.
- (3) If $x \in V$ is a variable and $\varphi \in \mathcal{L}$ then $\forall x \varphi \in \mathcal{L}$.
- (4) If $\varphi, \psi \in \mathcal{L}$ then $(\varphi \wedge \psi), \neg \varphi \in \mathcal{L}$.

Definition 6.2

Let $\varphi \in \mathcal{L}$ be a formula, and suppose $x \in V$ has an occurrence in φ . This occurrence is **bound** if it occurs within the scope of the quantifier $\forall x$, and otherwise **free**. Let $free\varphi$ be all the variables which occur free in φ . Write $\varphi = \varphi(\bar{x})$ if $free\varphi \subseteq \bar{x}$.

Definition 6.3

Let $\varphi(\bar{x})$ be a formula, \mathcal{A} a structure, and \bar{a} a sequence of elements from \mathcal{A} . Then we write $\mathcal{A} \models \varphi(\bar{a})$ to mean that φ holds in \mathcal{A} when we valuate \bar{x} as \bar{a} . Formally:

- (1) If $\varphi = t(\bar{x}) = s(\bar{x})$, then $\mathcal{A} \models \varphi(\bar{a})$ if and only if $t(\bar{a}) = s(\bar{a})$.
- (2) If $\varphi = rt_1(\bar{x}) \dots t_n(\bar{x})$, then $\mathcal{A} \models \varphi(\bar{a})$ if and only if $r^{\mathcal{A}}t_1(\bar{a}) \dots t_n(\bar{a})$.
- (3) If $\varphi = \forall y \psi(\bar{x}, y)$, then $\mathcal{A} \models \varphi(\bar{a})$ if and only if for every $b \in \mathcal{B}$, $\mathcal{A} \models \psi(\bar{a}, b)$.
- (4) $\mathcal{A} \models (\varphi \wedge \psi)(\bar{a})$ if and only if $\mathcal{A} \models \varphi(\bar{a}), \psi(\bar{a})$.
- (5) $\mathcal{A} \models \neg \varphi(\bar{a}) \iff \mathcal{A} \not\models \varphi(\bar{a})$.

\models is called the **satisfaction relation**.

Definition 6.4

Let X be a set of \mathcal{L} -formulas, and $\varphi \in \mathcal{L}$. Then we say $X \models \varphi$ if and only if for every $\mathcal{A} \models X$, $\mathcal{A} \models \varphi$. This is called the **consequence relation**.

Definition 6.5

An \mathcal{L} -theory T is a set of \mathcal{L} sentences (formulas with no free variables) such that $T \models \varphi \implies \varphi \in T$. I.e. it is deductively closed. Let X be a set of sentences, then there exists a smallest theory which contains it (its deductive close: $T_X = \{\varphi \mid X \models \varphi\}$). T_X is said to be **axiomatized** by X .

Now for some model-theoretic results which we will not prove:

Theorem 6.6 (The Compactness Theorem)

Let X be a set of \mathcal{L} -formulas, and $\varphi \in \mathcal{L}$. Then $X \models \varphi$ if and only if there exists a finite $X_0 \subseteq X$ such that $X_0 \models \varphi$.

Definition 6.7

A \forall **formula** is a formula of the form $\forall \vec{x} \varphi$ where φ is quantifier-free. A theory is a \forall -theory if it can be axiomatized by a set of \forall -sentences.

Theorem 6.8

A theory is invariant under substructures (i.e. $\mathcal{B} \subseteq \mathcal{A} \models T \implies \mathcal{B} \models T$) if and only if it is a \forall -theory.

Definition 6.9

A **positive formula** is one constructed only by $\wedge, \vee, \forall, \exists$ (i.e. without \neg).

Theorem 6.10

A theory is invariant under homomorphic images if and only if it can be axiomatized by positive sentences.

Definition 6.11

Say that a positive theory T **allows for weak ideals** if there exists a constant e and a term t such that

$$T \models \forall a, b (t(a, b) = e \leftrightarrow a = b)$$

And T **allows for ideals** if it is also universal and there exists a constant e and terms t, \bar{t} such that

$$T \models \forall a, b (\bar{t}(t(a, b), b) = a), \quad T \models \forall a, b (t(\bar{t}(a, b), b) = a), \quad T \models \forall a (t(a, a) = e)$$

If T allows for ideals then $\bar{t}(e, b) = \bar{t}(t(b, b), b) = b$. And so if T allows for ideals then it allows for weak ideals: if $t(a, b) = e$ then $\bar{t}(t(a, b), b) = \bar{t}(e, b) = b$. And $\bar{t}(t(a, b), b) = a$, so $a = b$.

Definition 6.12

Call \mathcal{A} **weakly idealized** by e, t if for every congruence, $a \sim b \iff t(a, b) \sim e$.

Theorem 6.13

If T allows for weak ideals, then all of its models are weakly idealized. Similarly if T allows for ideals, then all of its models are idealized.

Proof: let $\mathcal{A} \models T$ be a model of T and let \sim be a congruence on \mathcal{A} . Then the canonical homomorphism $f: \mathcal{A} \rightarrow \mathcal{A}/\sim$ is surjective and thus since T is positive, $\mathcal{A}/\sim \models T$. Now,

$$a \sim b \iff fa = fb \iff \mathfrak{t}(fa, fb) = e \iff f\mathfrak{t}(a, b) = e = fe \iff \mathfrak{t}(a, b) \sim e$$

as required. And so if T allows for ideals, its models are obviously all idealized. ■

If T allows for ideals then it is closed under substructures, so substructures of its models are also idealized. So we can now, in good conscience, restate the second isomorphism theorem:

Theorem 6.14 (The Idealized Second Isomorphism Theorem)

Let T allow for ideals, $\mathcal{A} \models T$, $\mathcal{B} \leq \mathcal{A}$ a substructure, and $I \trianglelefteq \mathcal{A}$ an ideal. Then I is an ideal of $I\mathcal{B}$ and $I \cap \mathcal{B}$ is an ideal of \mathcal{B} and

$$I\mathcal{B}/I \cong \mathcal{B}/I \cap \mathcal{B}$$

Since the theory of groups, rings, modules, and boolean algebras all allow for ideals, we have successfully generalized all the isomorphism theorems in a way which takes ideals into account.

7 The Chinese Remainder Theorem

Recall the Chinese Remainder Theorem:

Theorem 7.1 (Chinese Remainder Theorem)

Let R be a ring, and $I_1, \dots, I_n \trianglelefteq R$ be pairwise comaximal ideals. Then

$$R/I_1 \cap \dots \cap I_n \cong R/I_1 \times \dots \times R/I_n$$

In order to generalize this, we need to define what it means for two ideals to be comaximal. Recall that I, J are comaximal if and only if $I + J = R$. Recall that $I + J$ is just the smallest ideal containing both I and J . Fortunately, we know the following

Lemma 7.2

Let $\{\theta_i\}_{i \in I}$ be a family of congruences, then the smallest congruence containing all of them is

$$\bigvee_{i \in I} \theta_i = \{(a, b) \mid \text{There exists a sequence } a = a_0, \dots, a_n = b \text{ such that } (a_j, a_{j+1}) \in \theta_i \text{ for some } i \in I\}$$

This is called the **join** of the congruences.

If $\{I_i\}_{i \in I}$ is a family of ideals, let us denote their join to be the ideal of the join of the congruences associated with I_i by $\bigvee_{i \in I} I_i$. This is simply

$$\bigvee_{i \in I} I_i = \{a \mid \text{There exists a sequence } e = a_0, \dots, a_n = a \text{ such that } \mathfrak{t}(a_j, a_{j+1}) \in \bigcup_{i \in I} I_i\}$$

Furthermore, instead of S_{\sim_I} , we will denote the \sim_I -closure of a set S by S_I .

Lemma 7.3

Let \mathcal{A} be idealized, and $I \trianglelefteq \mathcal{A}$ an ideal. Then $\bar{\mathfrak{t}}(I, I) = \bar{\mathfrak{t}}(I, e) = \bar{\mathfrak{t}}(e, I) = I$.

Proof: this is since

$$\begin{aligned}\bar{\mathfrak{t}}(I, I) &= I_I = I \\ \bar{\mathfrak{t}}(I, e) &= \{e\}_I = I \\ \bar{\mathfrak{t}}(e, I) &= I_{\{e\}} = I\end{aligned}$$

as required. ■

Lemma 7.4

Let \mathcal{A} be idealized, and $I, J \trianglelefteq \mathcal{A}$ ideals. Then $I_J = J_I$, i.e. $\bar{\mathfrak{t}}(I, J) = \bar{\mathfrak{t}}(J, I)$.

Proof: we know that

$$\bar{\mathfrak{t}}(I, J) \sim_J \bar{\mathfrak{t}}(I, e) = I$$

so $\bar{\mathfrak{t}}(I, J) \subseteq I_J = \bar{\mathfrak{t}}(J, I)$. By symmetry we have the other direction as well. ■

Lemma 7.5

$I \vee J = \bar{\mathfrak{t}}(I, J)$.

Proof: we know that

$$I \vee J = \{a \mid \exists e = a_0, \dots, a_n = a: \mathfrak{t}(a_i, a_{i+1}) \in I \cup J\}$$

So assume $a \in I \vee J$, we will prove by induction that $a_i \in \bar{\mathfrak{t}}(I, J)$. For $a_0 = e$, this is trivial. Now, $a_i \in \bar{\mathfrak{t}}(I, J)$ and $\mathfrak{t}(a_{i+1}, a_i) \in I \cup J$. Thus

$$a_{i+1} = \bar{\mathfrak{t}}(\mathfrak{t}(a_{i+1}, a_i), a_i) \in \bar{\mathfrak{t}}(I \cup J, \bar{\mathfrak{t}}(I, J)) = \bar{\mathfrak{t}}(I, \bar{\mathfrak{t}}(I, J)) \cup \bar{\mathfrak{t}}(J, \bar{\mathfrak{t}}(I, J))$$

Now,

$$\bar{\mathfrak{t}}(I, \bar{\mathfrak{t}}(I, J)) \sim_I J$$

so $\bar{\mathfrak{t}}(I, \bar{\mathfrak{t}}(I, J)) \subseteq J_I = \bar{\mathfrak{t}}(I, J)$. By symmetry (since $\bar{\mathfrak{t}}(I, J) = \bar{\mathfrak{t}}(J, I)$), we have that $\bar{\mathfrak{t}}(J, \bar{\mathfrak{t}}(I, J)) \subseteq \bar{\mathfrak{t}}(I, J)$. Thus $I \vee J \subseteq \bar{\mathfrak{t}}(I, J)$.

Conversely, $\bar{\mathfrak{t}}(I, J) \sim_{I \vee J} \bar{\mathfrak{t}}(e, e) = e$ so $\bar{\mathfrak{t}}(I, J) \subseteq I \vee J$. ■

So the following definition is motivated

Definition 7.6

Let \mathcal{A} be idealized, and $I, J \trianglelefteq \mathcal{A}$ ideals. They are **comaximal** if $\bar{\mathfrak{t}}(I, J) = \mathcal{A}$.

Theorem 7.7 (The Idealized Chinese Remainder Theorem)

Let I, J be comaximal ideals, then

$$\mathcal{A}/_{I \cap J} \cong \mathcal{A}/_I \times \mathcal{A}/_J$$

Proof: let us define

$$h: \mathcal{A}/_{I \cap J} \longrightarrow \mathcal{A}/_I \times \mathcal{A}/_J$$

Its kernel is all $a \in \mathcal{A}$ such that $(a/I, a/J) = (e/I, e/J)$ i.e. $a \in I \cap J$. So all that remains is to show that h is surjective. Let $a, b \in \mathcal{A}$, then $a, b \in I_J, J_I$ so $a \sim_I j \in J$ and $b \sim_J \bar{\mathfrak{t}}(i, e) \in I$ for some $i \in I$ since $\bar{\mathfrak{t}}(I, e) = I$, thus

$$\bar{\mathfrak{t}}(i, j)/_I = \bar{\mathfrak{t}}(e, j)/_I = j/_I = a/_I$$

and

$$\bar{\mathfrak{t}}(i, j) / J = \bar{\mathfrak{t}}(i, e) / J = b / J$$

Thus $h\bar{\mathfrak{t}}(i, j) = (a/I, b/J)$ as required. So h is indeed surjective, and we have the desired result. ■

Let us translate this to the world of boolean algebras: Notice that a filter F is \mathcal{B} if and only if $0 \in F$, and similarly an ideal is \mathcal{B} if and only if $1 \in I$. Thus

Theorem 7.8

Let \mathcal{B} be a boolean algebra, and $F_1, F_2 \subseteq \mathcal{B}$ filters of \mathcal{B} . Then if there exists elements $a \in F_1, b \in F_2$ such that $a \leftrightarrow b = 0$ then

$$\mathcal{B} / F_1 \cap F_2 \cong \mathcal{B} / F_1 \times \mathcal{B} / F_2$$

Similarly if $I_1, I_2 \subseteq \mathcal{B}$ are ideals such that there exists $a \in I_1, b \in I_2$ such that $a \Delta b = 1$ then

$$\mathcal{B} / I_1 \cap I_2 \cong \mathcal{B} / I_1 \times \mathcal{B} / I_2$$