Computational Recognition of Properties of Finite Algebras

Ralph McKenzie, George McNulty, and Ross Willard

Preamble

In 1993 Ralph McKenzie resolved several of the most intriguing and challenging problems concerning varieties generated by finite algebras with only finitely many basic operations. These accomplishments can be summarized as follows.

Accomplishment 1. *There is a finite algebra generating a residually countable inherently nonfinitely based variety.*

This refutes the R-S Conjecture that every finitely generated residually small variety should be residually very finite. It also refutes an old and provocative speculation that the finite algebras of finite type which generate residually small varieties might all be finitely based.

Accomplishment 2. *There is no algorithm for deciding which finite algebras generate residually finite varieties.*

Accomplishment 3. *There is no algorithm for deciding which finite algebras are finitely based.*

So Ralph McKenzie settled Tarski's celebrated Finite Basis Problem. Until McKenzie's breakthrough, no algbraically reasonable property of finite algebras was known to be undecidable. Indeed, a number of properties (generating a minimal variety, generating a congruence modular variety, etc.) were long known to be decidable.

Ralph McKenzie invented a robust technique for interpreting an arbitrary Turing machine $\mathcal{T}$ into a finite algebra $\mathbf{A}(\mathcal{T})$ so that the machine computations would be available in the variety generated by $\mathbf{A}(\mathcal{T})$. It seems likely that it can be used to demonstrate the undecidability of a wide assortment properties of varieties generated by finite algebras. Indeed, further undecidability results have already to obtained by Charles Latting, Ralph McKenzie, and Ross Willard.

These lectures are intended to provide a path to these accomplishments. There are only two main differences between McKenzie's exposition and the one found here. First, I organized the material into lectures which are each, more or less, amenable to presentation in fifty-minutes. The second and more significant deviation is that I followed the work of Ross Willard to prove that $\mathbf{A}(\mathcal{T})$ is finitely based when $\mathcal{T}$ halts. I added no new results (and I hope no errors either). This presentation is intended to be concrete, and to reveal how the ideas develop toward the ultimate results. See Ross Willard's work for a more abstract perspective.

These notes arose from five occasions on which I gave series of lectures on this material. My colleagues at LaTrobe University (1994), at the University of South Carolina (1994–95,1996–97), at the University of Hawaii (1995–96), at the Beijing Workshop in Logic, Universal Algebra, and Computer Science (1998), and in the Ulam Seminar at the University of Colorado (1998) all endured my struggles to talk reasonably about these results. Their criticisms and ideas have become part of these notes. Ralph McKenzie and Ross Willard both shared early drafts of their work with me. Of course, these lectures owe a large debt to Ralph McKenzie (say about 98%) who originated these spectacular results.

A Toolbox of Background Material

An algebra is conceived as a nonempty set equipped with a system of operations, each of finite rank.

Let A be a nonempty set and X be a set. We use A^X to denote the set of all functions from X into A. In particular when $n = \{0, 1, \ldots, n-1\}$ is a natural number, then A^n denotes the set of all n-tuples (or n-termed sequences) of elements of A. We refer to a function $F : A^n \to A$ as an **operation on A of rank** n. As long as A is not empty, every operation has a unique rank.

So an **algebra** is a system $\mathbf{A} = \langle A, F_i \rangle_{i \in I}$ where A is nonempty and F_i is an operation on A of finite rank, whenever $i \in I$. The set A is called the **universe** of $\mathbf{A}$, the each F_i is referred to as a **basic operation** of $\mathbf{A}$. The index set I is called the **set of operation symbols** of $\mathbf{A}$. For each algebra $\mathbf{A}$ the function $\rho : I \to \omega$ defined by setting $\rho(i)$ to the rank of F_i, for each $i \in I$, is called the **similarity type** or the **language** of $\mathbf{A}$. Algebras with the same similarity type are said to be **similar**.

For example, the symbols $+, -, \cdot, 0, 1$ are the ones ordinarily encountered in ring theory. The language of ring theory consists of these symbols and the information that $+$ and $\cdot$ have rank 2, that $-$ has rank 1, and that 0 and 1 have rank 0. Given particular symbols like these and an algebra $\mathbf{A}$ of the appropriate similarity type, we tend to write

$$
\begin{array}{llllll}
+^{\mathbf{A}} & \text{in place of } F_+ & \cdot^{\mathbf{A}} & \text{in place of } F_\cdot & -^{\mathbf{A}} & \text{in place of } F_- \\
0^{\mathbf{A}} & \text{in place of } F_0 & 1^{\mathbf{A}} & \text{in place of } F_1 & &
\end{array}
$$

and to display $\mathbf{A}$ as $\langle A, +^{\mathbf{A}}, -^{\mathbf{A}}, \cdot^{\mathbf{A}}, 0^{\mathbf{A}}, 1^{\mathbf{A}} \rangle$.

The notions of subalgebra, homomorphism, and direct product of systems of algebras are commonplace in familiar branches of algebra. Below we frame them in a general setting.

Let $\mathbf{A}$ be an algebra and $B \subseteq A$. We say that B is a **subuniverse** of $\mathbf{A}$ if B is closed with respect to all be basic operations of $\mathbf{A}$. This means that if F is a basic operation of $\mathbf{A}$ and r is the rank of F, and $b_0, b_1, \ldots, b_{r-1} \in B$, then $F(b_0, b_1, \ldots, b_{r-1}) \in B$. If B is a nonempty subuniverse of $\mathbf{A}$, then the algebra $\mathbf{B}$ similar to $\mathbf{A}$ obtained by restricting each of the basic operations of $\mathbf{A}$ to B is said to be a **subalgebra** of $\mathbf{A}$. Notice that it is possible for the empty set to be a subuniverse of certain algebras. But no subalgebra is permitted to have an empty universe.

Let $\mathbf{A}$ and $\mathbf{B}$ be similar algebras, and let $h : A \to B$. The function h is a **homomorphism** from $\mathbf{A}$ into $\mathbf{B}$ provided for every operation symbol Q and all $a_0, a_1, \ldots, a_{r-1} \in A$

$$
h\left(Q^{\mathbf{A}}(a_0, a_1, \ldots, a_{r-1})\right) = Q^{\mathbf{B}}(h(a_0), h(a_1), \ldots, h(a_{r-1}))
$$

where r is the rank of Q. If, in addition, h is one-to-one then h is an **embedding** of $\mathbf{A}$ into $\mathbf{B}$. If, on the other hand, h is onto B, then $\mathbf{B}$ is a **homomorphic image** of $\mathbf{A}$. In the event that h is both one-to-one and onto B, then h is an **isomorphism** and $\mathbf{A}$ and $\mathbf{B}$ are said to be **isomorphic**.

Homomorphisms from $\mathbf{A}$ into $\mathbf{A}$ are called **endomorphisms** of $\mathbf{A}$, and isomorphisms from $\mathbf{A}$ onto $\mathbf{A}$ are called **automorphisms** of $\mathbf{A}$.

Let I be any set and let $\langle A_i : i \in I\rangle$ be a system of nonempty sets. By a **choice function** for this system we mean a function $f : I \to \bigcup_I A_i$ such that the value of f at i (which we denote both as $f(i)$ and f_i) must belong to A_i for each $i \in I$. The direct product $\prod_I A_i$ of the system $\langle A_i : i \in I\rangle$ of sets is the set of all choice functions for the system. Now suppose $\langle \mathbf{A}_i : i \in I\rangle$ is a system of similar algebras. By the direct product $\mathbf{A} = \prod_I \mathbf{A}_i$ we mean the algebra with universe $A = \prod_I A_i$ and with the basic operations defined coordinatewise as follows: Let Q be an operation symbol. Let r be the rank of Q, and let $f^{(0)}, f^{(1)}, \ldots, f^{(r-1)} \in A$. Then

$$Q^{\mathbf{A}}(f^{(0)}, f^{(1)}, \ldots, f^{(r-1)}) = \langle Q^{\mathbf{A}_i}(f_i^{(0)}, f_i^{(1)}, \ldots, f_i^{(r-1)}) : i \in I\rangle.$$

Observe that it is possible to take I to be empty. The direct product of an empty system of sets is the set $\{0\} = \{\varnothing\} = 1$. The product of the empty system of similar algebras is an algebra with only one element.

Let $\mathcal{K}$ be a class of similar algebras. By $H\ \mathcal{K}$ we mean the class of all homomorphic images of algebras belonging to $\mathcal{K}$. By $S\ \mathcal{K}$ we mean the class of all algebras isomorphic to subalgebras of algebras belong to $\mathcal{K}$. By $P\ \mathcal{K}$ we mean the class of all algebras isomorphic to direct products of systems of algebras belonging to $\mathcal{K}$.

The syntax appropriate to a given language, in this setting, begins with the concept of terms. We reserve a denumberably infinite sequence of distinct **variables**: $x_0, x_1, x_2, \ldots$ and we will always suppose that the variables and operation symbols of our languages can be concatenated in an unambiguous manner. More precisely, we suppose that no variable or operation symbol can be obtained as a finite string of other symbols chosen from among the variables and the operation symbols of the language. Then the set of **terms** is defined by the following recursion:

i. Every variable is a term.

ii. If Q is an operation symbol, r is the rank of Q, and $t_0, t_1, \ldots, t_{r-1}$ are terms, then $Qt_0t_1 \ldots t_{r-1}$ is a term.

Just as the operation symbols denoted the basic operations of an algebra, so the terms denote certain functions on an algebra. Let $\mathbf{A}$ be an algebra. The set of **term functions** of $\mathbf{A}$ is defined by the following recursion:

i. If t is x_i, then $t^{\mathbf{A}} : A^{\omega} \to A$ is the i^{th} projection function.

ii. If $t = Qt_0t_1 \ldots t_{r-1}$, where Q is an operation symbol of rank r and $t_0, t_1, \ldots, t_{r-1}$ are terms, then $t^{\mathbf{A}} : A^{\omega} \to A$ is defined so that for all $a = \langle a_0, a_1, a_2, \ldots\rangle \in A^{\omega}$

$$t^{\mathbf{A}}(a) = Q^{\mathbf{A}}(t_0^{\mathbf{A}}(a), t_1^{\mathbf{A}}(a), \ldots, t_{r-1}^{\mathbf{A}}(a)).$$

Notice that while $t^{\mathbf{A}}$ has been defined as an operation on A of rank ω, it can really only depend on finitely many inputs, since only finitely many variables actually occur in t.

An **equation** is just an ordered pair of terms. For terms s and t we use the suggestive notation $s \approx t$ for the ordered pair (s, t) of terms. Given an algebra $\mathbf{A}$ and terms s and t, all of the same similarity type, we say that $s \approx t$ is **true** in $\mathbf{A}$ and that $\mathbf{A}$ is a **model** of $s \approx t$ if and only if $s^{\mathbf{A}} = t^{\mathbf{A}}$. To denote this important relationship between algebras and equations we use

$$\mathbf{A} \models s \approx t.$$

Let $\mathcal{K}$ be any class of similar algebras and let Σ be any set of equations of the same similarity type as $\mathcal{K}$. We say that Σ is true in $\mathcal{K}$, and that $\mathcal{K}$ is a class of models of Σ provided $\mathbf{A} \models s \approx t$ for every algebra $\mathbf{A} \in \mathcal{K}$ and every equation $s \approx t \in \Sigma$. Our notation in this case is $\mathcal{K} \models \Sigma$. The class of all models of Σ is denoted by $\operatorname{Mod}\Sigma$.

A class $\mathcal{V}$ of similar algebras is a **variety** if and only if it is closed with respect to the formation of homomorphic images, subalgebras, and direct products.

THEOREM A (Birkhoff's HSP Theorem). *Let $\mathcal{K}$ be a class of similar algebras. The following are equivalent:*

i. *$\mathcal{K}$ is a variety.*
ii. *$\mathcal{K} = HSP\ \mathcal{K}$.*
iii. *There is a set Σ of equations so that $\mathcal{K} = \operatorname{Mod}\Sigma$.*

An algebra is **locally finite** if each of its finitely generated subalgebras is finite. A variety is **locally finite** if each of its algebras is locally finite.

THEOREM B. *Every variety generated by a finite algebra is locally finite.*

THEOREM C. *An infinite locally finite algebra* $\mathbf{A}$ *generates a locally finite variety if and only if for each natural number n there is a finite upper bound on the cardinalities of the n-generated subalgebras of* $\mathbf{A}$.

A set Σ of equations is a **base** for the variety $\mathcal{V}$ (or the algebra $\mathbf{A}$) provided $\mathcal{V}$ (or $HSP\ \mathbf{A}$) is the class of all models of Σ. An algebra or a variety is said to be **finitely based** if and only if it has a finite base.

Let $\mathcal{V}$ be a variety and let n be a natural number. We denote by $\mathcal{V}^{(n)}$ the variety based on the following set of equations:

$$\{s \approx t : \text{ no more than } n \text{ distinct variables occur in } s \approx t \text{ and } \mathcal{V} \models s \approx t\}.$$

THEOREM D. *Let $\mathcal{V}$ be a variety and let n be a natural number.* $\mathbf{A} \in \mathcal{V}^{(n)}$ *if and only if every subalgebra of* $\mathbf{A}$ *with no more than n generators belongs to $\mathcal{V}$.*

THEOREM E (Birkhoff). *If $\mathcal{V}$ is a locally finite variety of finite type and n is a natural number, then $\mathcal{V}^{(n)}$ is finitely based.*

An algebra $\mathbf{A}$ is **inherently nonfinitely based** provided:

i. $\mathbf{A}$ has only finitely many basic operations,
ii. $\mathbf{A}$ belongs to some locally finite variety,
iii. $\mathbf{A}$ belongs to no locally finite variety which is finitely based.

Likewise, a variety $\mathcal{V}$ is **inherently nonfinitely based** provided it has a finite similarity type, is locally finite, and is included in no finitely based locally finite variety. As a consequence, if $\mathbf{A}$ is inherently nonfinitely based and $\mathbf{A} \in HSP\ \mathbf{B}$, where $\mathbf{B}$ is a finite algebra, then $\mathbf{B}$ is also inherently nonfinitely based. This property is a strong, even contagious, failure of an algebra to be finitely based.

The concept of an inherently nonfinitely based variety can be framed in an entirely algebraic way, as the next theorem demonstrates.

THEOREM F. *For a locally finite variety $\mathcal{V}$ of a finite similarity type, the following conditions are equivalent:*

i. *$\mathcal{V}$ is inherently nonfinitely based.*

ii. *The variety $\mathcal{V}^{(n)}$ is not locally finite for any natural number n.*
iii. *For arbitrarily large natural numbers n, there is a non-locally-finite algebra $\mathbf{B}_n$ whose n-generated subalgebras belong to $\mathcal{V}$.*

Let $\mathbf{A}$ be an algebra and $0 \in A$. 0 is an **absorbing** element provided $F(a_0, a_1, \ldots, a_{r-1}) = 0$, whenever F is a basic operation of $\mathbf{A}$, r is the rank of F, and $0 \in \{a_0, a_1, \ldots, a_{r-1}\} \subseteq A$. Any element of A which is not an absorbing element is called a **proper** element of A. Every basic operation F of $\mathbf{A}$ of rank r can be construed as a set of $r+1$-tuples. Those $r+1$ tuples which belong to F and which have no absorbing element among their entries comprise the **proper part** of F. Notice that an algebra can have at most one absorbing element, if the algebra has a basic operation of rank at least 2.

Theorem G (Baker, McNulty, and Werner). *Let $\mathbf{A}$ be an infinite locally finite algebra with finitely many basic operations, with an absorbing element 0, and with an automorphism σ such that*
i. $\{0\}$ *is the only σ-orbit of $\mathbf{A}$ which is finite,*
ii. *the proper part of every basic operation of $\mathbf{A}$ is partitioned by σ into only finitely many orbits, and*
iii. $\sigma(a) = \lambda(a)$ *for some proper element a and some nonconstant unary polynomial λ of $\mathbf{A}$.*

Then $\mathbf{A}$ is inherently nonfinitely based.

Let $\mathbf{A}$ and h be a homomorphism with domain A. By the **kernel** of h we mean

$$\ker h = \{(a,b) : a,b \in A \text{ and } h(a) = h(b)\}$$

It is easy to see that the kernel of a homomorphism is an equivalence relation on A such that for each basic operation F of $\mathbf{A}$ and all $a_0, b_0, a_1, b_1, \ldots, a_{r-1}, b_{r-1} \in A$ where r is the rank of F,

$$\text{If } (a_i, b_i) \in \ker h \text{ for all } i < r, \text{ then } (F(a_0, a_1, \ldots, a_{r-1}), F(b_0, b_1, \ldots, b_{r-1})) \in \ker h.$$

By a **congruence relation** on $\mathbf{A}$ we means any θ such that

- θ is an equivalence relation on A, and
- for each basic operation F of $\mathbf{A}$ and all $a_0, b_0, a_1, b_1, \ldots, a_{r-1}, b_{r-1} \in A$ where r is the rank of F,

$$\text{If } (a_i, b_i) \in \theta \text{ for all } i < r, \text{ then } (F(a_0, a_1, \ldots, a_{r-1}), F(b_0, b_1, \ldots, b_{r-1})) \in \theta.$$

Thus, the kernel of each homomorphism is a congruence relation. It turns out that, conversely, every congruence relation is also the kernel of some homomorphism.

Suppose $\mathbf{A}$ is a algebra, that θ is a congruence on $\mathbf{A}$, and that $a \in A$. We denote the equivalence class of a with respect to θ by $a/\theta := \{a' : a \;\theta\; a'\}$. We call such equivalence classes **congruence classes**. By A/θ we mean $\{a/\theta : a \in A\}$, which is the partition of A into congruences classes with respect to θ. Operations can be imposed on A/θ in a natural way, leading to the **quotient algebra** denoted by $\mathbf{A}/\theta$. The universe of this algebra is A/θ and for any operation symbol Q and any $a_0, a_1, \ldots, a_{r-1} \in A$ (where r is the rank of Q) we put

$$Q^{\mathbf{A}/\theta}(a_0/\theta, a_1/\theta, \ldots, a_{r-1}/\theta) = Q^{\mathbf{A}}(a_0, a_1, \ldots, a_{r-1})/\theta.$$

It is necessary to check that this definition is independent of the choice of the representatives a_i within each congruence class. But this independence is immediate from the definition of congruence relation. The map $\eta : A \to A/\theta$ defined by

$$\eta(a) = a/\theta,$$

is called the **quotient map**.

The next few theorems are familiar in the context of groups and rings.

THEOREM H (The Homomorphism Theorem). *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be similar algebras, let* h *be a homomorphism from* $\mathbf{A}$ *onto* $\mathbf{B}$*, let* θ *be a congruence on* $\mathbf{A}$*, and let eta be the quotient map from* A *onto* A/θ*. Then*

i. *The kernel of* h *is a congruence on* $\mathbf{A}$*.*
ii. *The quotient map* η *is a homomorphism from* $\mathbf{A}$ *onto* $\mathbf{A}/\theta$*.*
iii. *If* $\theta = \ker h$*, then the unique function* f *from* A/θ *onto* B *satisfying* $f \circ \eta = h$ *is an isomorphism from* $\mathbf{A}/\theta$ *onto* $\mathbf{B}$*.*

THEOREM I (The Second Isomorphism Theorem). i. *Let* $f : \mathbf{A} \to \mathbf{B}$ *and* $g : \mathbf{A} \to \mathbf{C}$ *be homomorphisms such that* $\ker f \subseteq \ker g$ *and* f *is onto* $\mathbf{B}$*. Then there is a unique homomorphism* $h : \mathbf{B} \to \mathbf{C}$ *satisfying* $g = h \circ f$*. Moreover,* h *is an embedding if and only if* $\ker f = \ker g$*.*

ii. *Let* θ *and* ϕ *be congruences of* $\mathbf{A}$ *with* $\theta \subseteq \phi$*. Then*

$$\phi/\theta := \{\langle x/\theta, y/\theta : \langle x, y\rangle \in \phi\}$$

is a congruence of $\mathbf{A}/\theta$*, and the formula*

$$h\left((a/\theta)/(\phi/\theta)\right) = a/\phi$$

defines an isomorphism from $(A/\theta)/(\phi/\theta)$ *onto* $\mathbf{A}/\phi$*.*

We use $\operatorname{Con}\mathbf{A}$ to denote the set of all congruence relations on $\mathbf{A}$. $\operatorname{Con}\mathbf{A}$ is an (algebraic) lattice with respect to the ordering by set inclusion. The largest element of this lattice, denoted by 1_A, is $A \times A$, while the smallest congruence relation, denoted by 0_A, is the identity relation on A. The meet in this lattice is just intersection. The join of a set W of congruence relations is the intersection of all congruence relations which include each of the congruences in W. Another way to say this is that $\bigvee W$ is the congruence relation of $\mathbf{A}$ generated by $\bigcup W$.

THEOREM J (The Correspondence Theorem). *Let* $\mathbf{A}$ *be an algebra and let* θ *be a congruence of* $\mathbf{A}$*. Let* $\mathbf{L}$ *denote the sublattice of* $\operatorname{Con}\mathbf{A}$ *with universe* $\{\phi : \theta \subseteq \phi \in \operatorname{Con}\mathbf{A}\}$*. Define* $F : L \to \operatorname{Con}\mathbf{A}/\theta$ *via*

$$F(\phi) = \phi/\theta$$

for all $\phi \in L$*. Then the map* F *is an isomorphism from* $\mathbf{L}$ *onto* $\operatorname{Con}\mathbf{A}/\theta$*.*

Let $\mathbf{A}$ be an algebra and let $X \subseteq A \times A$. The **congruence relation generated** by X, which we denote by $\operatorname{Cg}^{\mathbf{A}} X$, is just the intersection of all congruence relations on $\mathbf{A}$ which include X. The following characterization of $\operatorname{Cg}^{\mathbf{A}} X$ is very useful.

Let $\mathbf{A}$ be an algebra. By a **basic translation** on $\mathbf{A}$ we mean a function of the form

$$F^{\mathbf{A}}(a_0, a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_{r-1})$$

where F is a basic operation symbol of $\mathbf{A}$, r is the rank of F, $0 \leq i < r$, and $a_0, \ldots, a_{r-1} \in A$. For each natural number k, by a **k-translation** on $\mathbf{A}$ we mean a composition of k or fewer basic translations. A **translation** on $\mathbf{A}$ is simply a k-translation on $\mathbf{A}$ for some natural number k. The identity map on A is the only 0-translation. Every basic translation is a 1-translation.

For any set X we use $\binom{X}{2}$ to denote the collection of all two-element subsets of X. Let $\mathbf{A}$ be an algebra and let $\{a, b\}, \{c, d\} \in \binom{A}{2}$. We write $\{a, b\} \looparrowright_k \{c, d\}$ to denote that there is a k-translation

$\lambda(x)$ on $\mathbf{A}$ such that $\{c,d\} = \{\lambda(a), \lambda(b)\}$. $\{a,b\} \looparrowright \{c,d\}$ means that $\{a,b\} \looparrowright_k \{c,d\}$ for some natural number k.

THEOREM K (The Congruence Generation Theorem). *Let* $\mathbf{A}$ *be an algebra and let* $X \subseteq A^2$. *For all* $a, b \in A$, $\langle a,b\rangle \in \mathrm{Cg}^{\mathbf{A}}(X)$ *if and only if there is some finite sequence* $c_0, \dots, c_n \in A$ *such that*

i. $a = c_0$ *and* $c_n = b$, *and*
ii. *for all* $i < n$ *there is* $\langle x,y\rangle \in X$ *with* $x \neq y$ *so that* $\{x,y\} \looparrowright \{c_i, c_{i+1}\}$.

Let $\mathbf{A}$ be an algebra and let I be any set. For each $i \in I$ let h_i be a homomorphism with domain A. The system $\langle h_i : i \in I\rangle$ **separates points** provided for all $a, b \in A$ with $a \neq b$, there is $i \in I$ such that $h_i(a) \neq h_i(b)$. Likewise, a system $\langle \theta_i : i \in I\rangle$ of congruences of $\mathbf{A}$ **separates points** provided for all $a, b \in A$ with $a \neq b$, there is $i \in I$ such that $\langle a,b\rangle \notin \theta_i$.

THEOREM L. *Let* $\mathbf{A}$ *be an algebra and let* I *be any set. For each* $i \in I$ *let* $\mathbf{B}_i$ *be an algebra,* h_i *be a homomorphism from* $\mathbf{A}$ *onto* $\mathbf{B}_i$ *with kernel* θ_i. *Let* $\mathbf{B} = \prod_I \mathbf{B}_i$ *and let* ρ_i *be the projection of* $\mathbf{B}$ *onto* $\mathbf{B}_i$. *Let* $h : A \to B$ *be defined via*

$$h(a) = \langle h_i(a) : i \in I\rangle$$

for all $a \in A$. *Further, let* $g_i : A/\ker h \to B_i$ *be defined via*

$$g_i\,(a/\ker h) = h_i(a)$$

for all $a \in A$ *and all* $i \in I$. *Then*

i. h *is a homomorphism,*
ii. $h_i = \rho_i \circ h$,
iii. $\ker h = \bigcap_I \theta_i$,
iv. g_i *is a homomorphism from* $\mathbf{A}/\ker h$ *onto* $\mathbf{B}_i$ *for every* $i \in I$, *and*
v. $\langle g_i : i \in I\rangle$ *separates points.*

Let $\mathbf{A}$ be any algebra and let I be any set. For each $i \in I$ let $\mathbf{B}_i$ be an algebra similar to $\mathbf{A}$ and let h_i be a homomorphism from $\mathbf{A}$ onto $\mathbf{B}_i$. The system $\langle h_i : i \in I\rangle$ of homomorphisms is said to be a **subdirect representation** of $\mathbf{A}$ provided it separates points. For such a subdirect representation, the algebras $\mathbf{B}_i$ are referred to as **subdirect factors** of $\mathbf{A}$. The subdirect representation $\langle h_i : i \in I\rangle$ is **trivial** if h_i is an isomorphism for some $i \in I$. An algebra $\mathbf{A}$ is **subdirectly irreducible** if each of its subdirect representations is trivial.

The next theorem is of fundamental importance.

THEOREM M (Birkhoff's Subdirect Representation Theorem). *Every algebra has a subdirect representation with all the factors subdirectly irreducible.*

COROLLARY N. *If* $\mathcal{V}$ *and* $\mathcal{W}$ *are varieties which have the same subdirectly irreducible members, then* $\mathcal{V} = \mathcal{W}$.

The next theorem characterizes subdirectly irreducible algebras.

THEOREM O. *Let* $\mathbf{A}$ *be an algebra. The following are equivalent:*

i. $\mathbf{A}$ *is subdirectly irreducible.*
ii. *There are elements* $a, b \in A$ *with* $a \neq b$ *such that* $a\ \theta\ b$ *for every* $\theta \in \mathrm{Con}\,\mathbf{A}$ *such that* $0_A < \theta$.
iii. *There is* $\mu \in \mathrm{Con}\,\mathbf{A}$ *such that* $0_A < \mu$ *and mu* $\leq \theta$ *for every* $\theta \in \mathrm{Con}\,\mathbf{A}$ *such that* $0_A < \theta$.

iv. 0_A *is strictly meeting irreducible in* $\operatorname{Con}\mathbf{A}$.

Pairs of elements $\langle a, b\rangle$ which fulfill condition (ii) are called **critical for A** and $\mathbf{A}$ is said to be **$\langle a, b\rangle$-critical**. The congruence μ in condition (iii) is the **monolith** of $\mathbf{A}$.

Let $\mathbf{A}$ be an algebra and let κ be a cardinal number. We say that $\mathbf{A}$ is **residually less than** κ if and only if for each pair of distinct elements a and b of A, there is an algebra $\mathbf{B}$ with $|B| < \kappa$ and a homomorphism h from $\mathbf{A}$ onto $\mathbf{B}$ such that $h(a) \neq h(b)$.

Let $\mathcal{K}$ be a class of algebras. The **residual bound** or the **residual character** of $\mathcal{K}$ is the least cardinal κ (if such exists) so that every algebra in $\mathcal{K}$ is residually less than κ. In the event that no such κ exists we say, following Walter Taylor, that $\mathcal{K}$ is **residually large**. On the other hand, if such a κ exists we say that $\mathcal{K}$ is **residually small**. If $\kappa = \omega_1$ we say that $\mathcal{K}$ is **residually countable**. If $\kappa = \omega$ we say that $\mathcal{K}$ is **residually finite**. If κ is finite we say that $\mathcal{K}$ is **residually very finite**.

Theorem P. *Suppose that $\mathcal{K}$ is a class of algebras closed under the formation of homomorphic images. Then $\mathcal{K}$ has residual character κ if and only if κ is the smallest cardinal which is larger than the cardinality of every subdirectly irreducible algebra in $\mathcal{K}$.*

Theorem Q (Quackenbush and Dziobiak). *Let $\mathcal{V}$ be a locally finite variety, and let $\mathbf{S} \in \mathcal{V}$ be a subdirectly irreducible algebra. Every finite subalgebra of $\mathbf{S}$ is embeddable into a finite subdirectly irreducible algebra in $\mathcal{V}$.*

Corollary R. *Every locally finite variety with an infinite subdirectly irreducible algebra must have arbitrarily large finite subdirectly irreducible algebras.*

Theorem S (Taylor). *Fix a similarity type with no more than λ operation symbols, where λ is an infinite cardinal. Let $\mathcal{V}$ be any variety. $\mathcal{V}$ is residually small if and only if every algebra in $\mathcal{V}$ is residually less than $(2^\lambda)^+$.*

Theorem T (McKenzie and Shelah). *Fix a similarity type with only countably many operation symbols. Every variety that has an uncountable subdirectly irreducible algebra, must have a subdirectly irreducible algebra of cardinality at least 2^ω.*

So, for a similarity type with only countably many operation symbols, only the following cardinals can be residual characters of varieties: $1, 3, 4, 5, ..., \omega, \omega_1$, and 2^ω. It is also possible for a variety to be residually large. All these possibilities occur, even among varieties generated by algebras with no more than four elements, as shown by McKenzie.

By constraining the residual character of a finitely generated variety and insisting that the congruence lattice of each algebra in the variety is sufficiently nice, it is sometimes possible to draw the conclusion that the variety is finitely based. Here is one broad condition on (congruence) lattices that serves this purpose:

$\mathrm{SD}_\wedge$ If $x \wedge y = x \wedge z$, then $(x \wedge y) \vee (x \wedge z) = x \wedge (y \vee z)$

A lattice satisfying this condition is said to be **meet-semidistributive**. A variety $\mathcal{V}$ is **congruence meet-semidistributive** provided the congruence lattice of each algebra in $\mathcal{V}$ is meet-semidistributive.

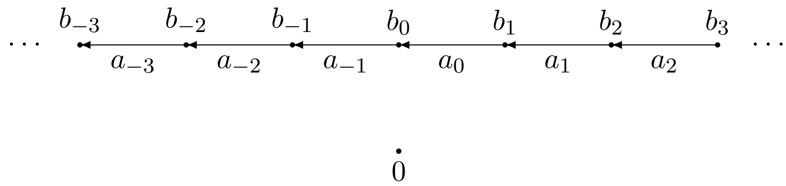
Theorem U (Willard). *Let $\mathcal{V}$ be a congruence meet-semidistributive variety of finite type. If $\mathcal{V}$ is residually finite, then $\mathcal{V}$ is finitely based.*

LECTURE 0

An Interesting Locally Finite Algebra

During the course of developing this material we will deal with algebras that have many basic operations. Among them will always be three denoted by $\cdot, \wedge$, and 0. $\wedge$ will always be a meet-semilattice operation and 0 will always denote the least element with respect to $\leq$, the underlying semilattice order. Meet semilattices of height one are referred to as *flat semilattices.* 0 is an *absorbing element* for the product $\cdot$ in the sense that $0 \cdot x \approx x \cdot 0 \approx 0$ always holds. The product also satisfies $(x \cdot y) \cdot z \approx 0 \approx x \cdot x$. Therefore, only right associated products can produce results other than 0. For this lecture we assume that the remaining operations are term operations built up from $\wedge, \cdot$, and 0.

Let $Q_{\mathbb{Z}} = \{0\} \cup \{a_p : p \in \mathbb{Z}\} \cup \{b_p : p \in \mathbb{Z}\}$, where all the a_p's and b_q's are distinct and different from 0. The algebra $\langle Q_{\mathbb{Z}}, \wedge, 0\rangle$ is a height 1 semilattice with least element 0. The product in $\mathbf{Q}_{\mathbb{Z}}$ is defined so that $a_p \cdot b_{p+1} = b_p$ for all $p \in \mathbb{Z}$, with all other products 0. Here is a picture that might help:



The operation $\cdot$ could be referred to as an *edge operation* and the algebra $\langle Q_{\mathbb{Z}}, \cdot, 0\rangle$ as an *edge algebra.* Any directed labelled graph gives rise to an *edge algebra* provided no two edges directed away from the same vertex have the same label. Specifically, for such a directed graph, $\cdot$ is an *edge operation on A* provided the elements of A fall into three disjoint sets—vertex elements, edge labels, and a default element 0—and $b = a \cdot d$ holds when a labels an edge directed from vertex d to vertex b, with all other $\cdot$ products producing the default element 0. Such edge algebras are related to Shallon's graph algebras. They might also be called "automatic algebras" since they are clearly related to finite automata. However, the term "automatic group" is already in use. In later lectures more complicated ternary operations rooted in digraphs with doubly labelled edges will be used to encode Turing machines and their computations.

$\mathbf{Q}_\omega$ denotes the subalgebra of $\mathbf{Q}_{\mathbb{Z}}$ with universe $\{0\} \cup \{a_p : p \in \omega\} \cup \{b_p : p \in \omega\}$. Likewise, for each natural number n, $\mathbf{Q}_n$ denotes the subalgebra with universe $\{0\} \cup \{a_p : 0 \leq p < n\} \cup \{b_p : 0 \leq p \leq n\}$. This algebra has $2n + 2$ elements.

Theorem 0. *Let $\mathbf{Q}_{\mathbb{Z}}$ be an algebra as described above (in particular with the given basic operations $0, \wedge$, and $\cdot$ and with all other basic operations being term operations built from these). The following hold:*

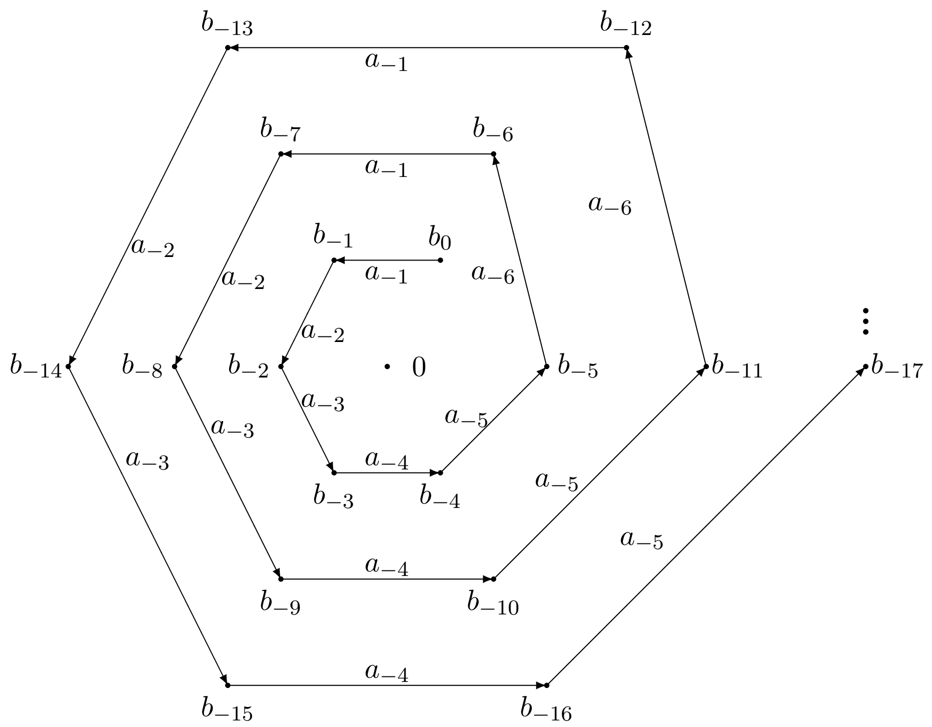
i. $\mathbf{Q}_\omega$ *is subdirectly irreducible, as is each* $\mathbf{Q}_n$*;*
ii. $\mathbf{Q}_{\mathbb{Z}}$ *generates a locally finite variety;*

iii. $\mathbf{Q}_{\mathbb{Z}}$ *is inherently nonfinitely based.*

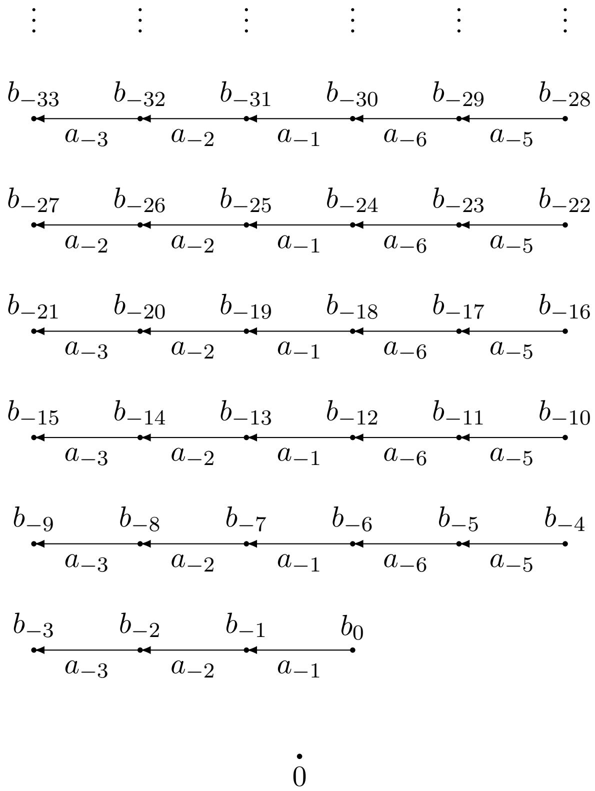
PROOF: Indeed, for (i) we have that $(0, b_0)$ belongs to every nontrivial congruence. To see this, let θ be a nontrivial congruence. First, suppose that $b_0 \;\theta\; c$ with $b_0 \neq c$. Then $b_0 = b_0 \wedge b_0 \;\theta\; b_0 \wedge c = 0$. Next, suppose that $0 < p$ and that θ collapses b_p to c where $b_p \neq c$. We obtain $b_{p-1} = a_{p-1} \cdot b_p \;\theta\; a_{p-1} \cdot c = 0$. So, inductively we have $b_0 \;\theta\; 0$. Finally, suppose that θ collapses a_p to c where $a_p \neq c$. Then we obtain $b_p = a_p \cdot b_{p+1} \;\theta\; c \cdot b_{p+1} = 0$, and so also $b_0 \;\theta\; 0$.

For (ii), note that no a_p results from the product operation and that b_p can only result from the product $a_p \cdot b_{p+1}$. So if S is any subset of $Q_{\mathbb{Z}}$, then $S \cup \{0\} \cup \{b_p : a_p \in S\}$ is a subuniverse of $\mathbf{Q}_{\mathbb{Z}}$. Thus the subuniverse of $\mathbf{Q}_{\mathbb{Z}}$ generated by a set of n elements will have no more than $2n+1$ elements, and usually a lot less. Hence $\mathbf{Q}_{\mathbb{Z}}$ generates a locally finite variety.

To establish (iii) for each large enough natural number N we build an algebra $\mathbf{Q}_{\mathbb{Z}}^{(N)}$. The algebra $\mathbf{Q}_{\mathbb{Z}}^{(6)}$ is pictured below:



The product $\cdot$ in $\mathbf{Q}_{\mathbb{Z}}^{(N)}$ is defined differently, while the meet and 0 retain their old meanings (and the remaining operations are still defined by the same terms). The universe $Q_{\mathbb{Z}}^{(N)} = \{0\} \cup \{a_{-1}, \ldots, a_{-N}\} \cup \{b_p : p \leq 0\}$. The algebra $\mathbf{Q}_{\mathbb{Z}}^{(N)}$ is infinite but finitely generated. Let $\mathbf{B}$ be a subalgebra of $\mathbf{Q}_{\mathbb{Z}}^{(N)}$ generated by fewer than N elements. It follows that some a_p is not in B. Let $\mathbf{C}$ be the subalgebra of $\mathbf{Q}_{\mathbb{Z}}^{(N)}$ whose universe consists of all elements except a_p. Below is a picture of C where $N = 6$ and a_{-4} is the omitted element.



Thus $\mathbf{Q}_{\mathbb{Z}}^{(N)}$ is made from $\mathbf{C}$ by a helical wrapping and the addition of a new element.

Now in C the b_p's are arranged in rows. Select a row. Let θ be the equivalence relation that isolates each b_p on the selected row, as well as each a_q, but collapses all the other b_r's to 0. Evidently, θ is a congruence relation of $\mathbf{C}$. It is also clear that $\mathbf{C}/\theta$ is isomorphic to a subalgebra of $\mathbf{Q}_{\mathbb{Z}}$. Since by selecting different rows we arrive at a family of congruences that separates the points of C, we conclude that $\mathbf{C}$ belongs to the variety generated by $\mathbf{Q}_{\mathbb{Z}}$. Hence, every subalgebra of $\mathbf{Q}_{\mathbb{Z}}^{(N)}$ generated by fewer than N elements belongs to the variety generated by $\mathbf{Q}_{\mathbb{Z}}$. Since this variety is locally finite, we have that it is inherently nonfinitely based. $\Box$

THEOREM 1. *Suppose that the only basic operation symbols are $\cdot, \wedge$, and 0. There is a finite algebra $\mathbf{A}_0$ so that $\mathbf{Q}_{\mathbb{Z}}$ belongs to the variety generated by $\mathbf{A}_0$.*

PROOF: Our approach is to let $\mathbf{A}_0$ denote an unknown finite algebra, set up the obvious conditions based on Birkhoff's HSP Theorem, and then try to solve for $\mathbf{A}_0$.

So we want $\mathbf{B} \subseteq \mathbf{A}_0^{\mathbb{Z}}$ and also $\theta \in \mathrm{Con}\mathbf{B}$ so that $\mathbf{Q}_{\mathbb{Z}} \cong \mathbf{B}/\theta$. To make this as easy as possible we would like $\mathbf{B}$ to be very much like $\mathbf{Q}_{\mathbb{Z}}$. We need to have in $A_0^{\mathbb{Z}}$ elements like a_p and b_p. So we want

$\alpha_p, \beta_p \in A_0^{\mathbb{Z}}$ so that $\beta_p = \alpha_p \beta_{p+1}$ for all $p \in \mathbb{Z}$. Writing this coordinatewise, we obtain conditions on $\mathbf{A}_0$:

$$\alpha_p(i)\beta_{p+1}(i) = \beta_p(i) \quad \text{for all } i \in \mathbb{Z}$$

An easy way to satisfy this is to provide A_0 with five elements $1, H, 2, C$, and D. Then we can let

$$\begin{array}{rccccccccc} \alpha_p := & \dots & 1 & 1 & 1 & H & 2 & 2 & 2 & \dots \\ \beta_p := & \dots & C & C & C & D & D & D & D & \dots \end{array}$$

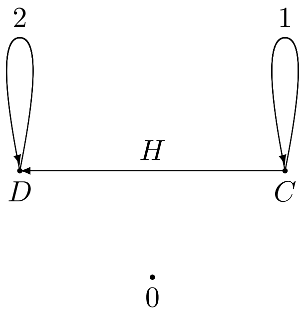
where the change is taking place at the p^{th} position. Then our condition looks like:

$$\begin{array}{rccccccccc} \alpha_p\colon & \dots & 1 & 1 & 1 & H & 2 & 2 & 2 & \dots \\ \beta_{p+1}\colon & \dots & C & C & C & C & D & D & D & \dots \\ \hline \beta_p\colon & \dots & C & C & C & D & D & D & D & \dots \end{array}$$

This imposes conditions on the product in $\mathbf{A}_0$. Indeed,

$$\begin{aligned} 1 \cdot C &= C \\ H \cdot C &= D \\ 2 \cdot D &= D \end{aligned}$$

To complete the description of $\mathbf{A}_0$ we make all the other products 0, a new element, and insist that $\wedge$ make it into a flat semilattice. $\mathbf{A}_0$ has six elements. It is an edge algebra for the digraph below:



To finish, we take $\mathbf{B}$ to be the subalgebra of $\mathbf{A}_0^{\mathbb{Z}}$ generated by all the α's and the β's. Let θ collapse all the $\mathbb{Z}$-tuples in B except the α's and the β's. Then it is easy to check that $\theta \in \operatorname{Con} \mathbf{B}$ and everything works. $\square$

It would be ideal if we could arrange for a finitely generated variety whose subdirectly irreducibles were exactly $\mathbf{Q}_\omega$ and the $\mathbf{Q}_n$'s. The surprising thing is that this is almost possible. With this in mind, the next lecture developes properties of finite subdirectly irreducible algebras in varieties generated by algebras like $\mathbf{A}_0$ above. However, by Lecture 2, we will see that it is convenient to add more elements and even more operations to $\mathbf{A}_0$. Additional elements do not interfere much with the construction above, but care is needed with the additional operations—we must ensure that in $\mathbf{Q}_{\mathbb{Z}}$ they turn out to be term operations built up from $\cdot, \wedge$, and 0. Thus, in Lecture 3 we will revisit the constructions above.

LECTURE 1

Finite Subdirectly Irreducibles Generated by Finite Flat Algebras

In this lecture we will suppose that $\mathbf{A}$ is a finite *flat* algebra (that is, an algebra among whose operations $\wedge$ and 0 can be found which provide the algebra with the structure of a meet-semilattice of height one with least element 0) and that $\mathbf{S}$ is a finite subdirectly irreducible algebra in the variety generated by $\mathbf{A}$. Of course, we have in mind for $\mathbf{A}$ the flat algebra described at the end of Lecture 0, although we should be willing to modify that algebra if the need arises. And we hope to show that $\mathbf{S}$ is one of the $\mathbf{Q}_n$'s.

Now according to Birkhoff's HSP Theorem, $\mathbf{S}$ will always arise as a quotient of some $\mathbf{B}$, which is in turn a subalgebra of $\mathbf{A}^T$ for some T. Since $\mathbf{S}$ is subdirectly irreducible, we know that there is a strictly meet irreducible $\theta \in \operatorname{Con}\mathbf{B}$ such that $\mathbf{S} \cong \mathbf{B}/\theta$. It is more convenient to work with $\mathbf{B}$ than with $\mathbf{S}$. Since $\mathbf{S}$ is finite, we can choose T to be finite. Indeed, in this lecture we assume the following:

- $\mathbf{B} \subseteq \mathbf{A}^T$
- $\theta \in \operatorname{Con}\mathbf{B}$
- θ is strictly meet-irreducible in $\operatorname{Con}\mathbf{B}$.
- $\mathbf{S} \cong \mathbf{B}/\theta$
- T is as small as possible for representing $\mathbf{S}$ in this way.

In particular this last condition entails that if $t \in T$, then there must be $u, v \in B$ so that $(u, v) \notin \theta$ but $u(s) = v(s)$ for all $s \in T - \{t\}$. Our effort at understanding the finite subdirectly irreducible $\mathbf{S}$ is largely focussed on θ.

First, we locate an element in B which is like the element b_0 in $\mathbf{Q}_n$. Since $\mathbf{B}$ is a semilattice, there are elements $u, v \in B$ with $u < v$ and (u, v) critical over θ. Using the finiteness of B pick p to be minimal among all those $v \in B$ such that (u, v) is critical over θ for some $u < v$.

Fact 0. *If $w < p$, then $(w, p) \notin \theta$.*

Proof: Suppose $w < p$ but $w \;\theta\; p$. Pick $u < p$ with (u, p) critical over θ. Then $w = p \wedge w \phi u \wedge w$, for all $\phi \in \operatorname{Con}\mathbf{B}$ with $\theta < \phi$. But this means that either $(w, u \wedge w) \in \theta$ or that $(w, u \wedge m)$ is critical over θ. So by the minimality of p, we have $u \wedge w \;\theta\; w$. But then $u = u \wedge p \;\theta\; u \wedge w \;\theta\; w \;\theta\; p$, contradicting $(u, p) \notin \theta$. □

Now for each $t \in T$ pick $(x, y) \in B^2 - \theta$ so that $x(t) \neq y(t)$ but $x(s) = y(s)$ for all $s \in T - \{t\}$. Pick $u < p$ so that (u, p) is critical over θ. So $(u, p) \in \theta \vee \operatorname{Cg}^{\mathbf{B}}(x, y)$. Then according to Mal'cev's Congruence Generation Theorem there is a finite sequence $e_0, e_1, \ldots, e_n$ of elements of B, of translations $\lambda_0, \ldots, \lambda_{n-1}$ of $\mathbf{B}$, and of two-element subsets $\{z_0, w_0\}, \ldots, \{z_{n-1}, w_{n-1}\}$ each belonging the $\theta \cup \{x, y\}$ such that

$$u = e_0 \quad \{e_i, e_{i+1}\} = \{\lambda_i(z_i), \lambda_i(w_i)\} \text{ for all } i < n \quad e_n = p.$$

But now, meeting every element in the sequence with p, we have

$$u = u \wedge p = e_0 \wedge p \quad \{e_i \wedge p, e_{i+1} \wedge p\} = \{\lambda_i(z_i) \wedge p, \lambda_i(w_i) \wedge p\} \text{ for all } i < n \quad e_n \wedge p = p \wedge p = p$$

Since $u < p$ there must be some $i < n$ so that $p \in \{\lambda_i(z_i) \wedge p, \lambda_i(w_i) \wedge p\}$ where $\lambda_i(z_i) \wedge p \neq \lambda_i(w_i) \wedge p$. Let χ_t denote the element of $\{\lambda_i(z_i) \wedge p, \lambda_i(w_i) \wedge p\}$ which is different from p. Evidently $\chi_t < p$. By Fact 0 we see that $(\chi_t, p) \notin \theta$. Hence, $(z_i, w_i) = (x, y)$ and $\{p, \chi_t\} = \{\lambda_i(x) \wedge p, \lambda_i(y) \wedge p\}$. From this construction we obtain:

- $\chi_t(s) = p(s)$ for all $s \in T - \{t\}$.
- $\chi_t(t) < p(t)$ for all $t \in T$.
- $\chi_t(t) = 0$ and $0 < p(t)$ for all $t \in T$.

The last item listed above is a consequence of the flatness of $\mathbf{A}$. Thus, χ_t agrees with p at all coordinates with the exception of t, where χ_t is 0 while p is not 0. So χ_t is uniquely determined by p and t (and is independent of the choices of x, y, and λ_i made above). We will eventually see—once enough is specified about $\mathbf{A}$—that p is also uniquely determined.

Fix $t_0 \in T$ so that $u \leq \chi_{t_0}$ for some $u < p$ for which (u, p) is critical over θ. Let $q = \chi_{t_0}$.

Fact 1. *p is a maximal element of $\mathbf{A}^T$. $\chi_t \in B$ and p covers χ_t in $\mathbf{A}^T$ for all $t \in T$. (q, p) is critical over θ. Finally, if $u \in A^T$ and $u < p$, then $u \in B$.*

Proof: Essentially, Fact 1 gathers the conclusions we drew above. To see that (q, p) is critical, notice $(q, p) \notin \theta$ according to Fact 0. Let $u \leq q < p$ with (u, p) critical over θ. Then we have $p \phi u = q \wedge u \phi q \wedge p = q$, for all $\phi > \theta$. The elements of A^T less than or equal to p form a Boolean algebra in which every element is a meet of the coatoms χ_t. □

Fact 2. *If $p \,\theta\, x$, then $p = x$*

Proof: Suppose $p \,\theta\, x$. Meeting both sides with p we also get $p \,\theta\, p \wedge x$. From Fact 0, we conclude that $p \not> p \wedge x$. Thus $p \leq x$. But since p is a maximal element, we arrive at $p = x$. □

Fact 3. *$x \,\theta\, y$ if and only if $\mu(x) = p \Leftrightarrow \mu(y) = p$ for all translations μ of $\mathbf{B}$.*

Proof: In the forward direction the result follows from Fact 2.

Now for the converse direction, suppose $(x, y) \notin \theta$. By Fact 1, we know $(q, p) \in \theta \vee \mathrm{Cg}^{\mathbf{B}}(x, y)$. Now repeating the analysis that led to the χ_t's we obtain a translation $\mu = \lambda \wedge p$ so that $\mu(x) \neq \mu(y)$ but $p \in \{\mu(x), \mu(y)\}$. □

Fact 4. *If $x < p$, then $(x, x \wedge q) \in \theta$.*

Proof: (q, p) is critical over θ by Fact 1, so $x = x \wedge p \phi x \wedge q$, for all $\phi > \theta$. Hence, either $(x, x \wedge q) \in \theta$ or $(x, x \wedge q)$ is critical over θ. Since $p > x \geq x \wedge q$, it follows from the minimality of p that $x \,\theta\, x \wedge q$. □

Suppose that x, y, and $z \in B$. Then $(x \wedge y)$ and $(x \wedge z)$ also belong to B and the element x is a common upper bound. Recalling that $\mathbf{B}$ has the structure of a finite $\wedge$-semilattice, it follows that $(x \wedge y)$ and $(x \wedge z)$ must have a least upper bound—we denote it by $(x \wedge y) \vee (x \wedge z)$.

Fact 5. *$\mathbf{S} \in HS\mathbf{A}$ or $(x \wedge y) \vee (x \wedge z)$ is not a polynomial of $\mathbf{B}$.*

PROOF: Suppose $\mathbf{S} \notin HS\mathbf{A}$. Then T has at least two elements. Let $t_1 \in T$ with $t_0 \neq t_1$. Let $q' = \chi_{t_1}$. Since $q' < p$ we have by Fact 4 that $q' \;\theta\; q' \wedge q$. But then, were $(x \wedge y) \vee (x \wedge z)$ a polynomial of $\mathbf{B}$, we would have $p = (p \wedge q) \vee (p \wedge q') \;\theta\; (p \wedge q) \vee (p \wedge q \wedge q') = q$. Since $(p, q) \notin \theta$, we conclude that $(x \wedge y) \vee (x \wedge z)$ is not a polynomial. □

Fact 5 reveals that our investigation of (finite) subdirectly irreducible algebras can be split in two. Since $\mathbf{A}$ is finite, a complete description of the subdirectly irreducible algebras in $HS\mathbf{A}$ can be devised given a description of $\mathbf{A}$. We only note the obvious upper bound on their cardinality. Most of our effort will concern the alternative case when $(x \wedge y) \vee (x \wedge z)$ is not a polynomial of $\mathbf{B}$. It is th subdirectly irreducible algebras arising from these algebras that we want to show must be isomorphic to our $\mathbf{Q}_n$'s.

Here is a lemma that simply gathers together the most salient of the facts just listed.

LEMMA 0. *Suppose that $\mathbf{A}$ is a finite flat algebra and that $\mathbf{S}$ is a finite subdirectly irreducible algebra in $HSP\mathbf{A}$. Choose T, $\mathbf{B}$, and $\theta \in \operatorname{Con}\mathbf{B}$ so that*

- *$\mathbf{B}$ is a subalgebra of $\mathbf{A}^T$,*
- *θ is (strictly) meet irreducible in $\operatorname{Con}\mathbf{B}$.*
- *$\mathbf{S} \cong \mathbf{B}/\theta$, and*
- *T is as small as possible subject to fulfilling the conditions above.*

Then there is an element $p \in B$ such that

i. *$(0, p)$ is critical over θ,*
ii. *$p/\theta = \{p\}$,*
iii. *p is a maximal element of A^T (so $p(s) > 0$ for all $s \in T$), and*
iv. *for all $x, y \in B$, $x \;\theta\; y$ if and only if $\mu(x) = p \Leftrightarrow \mu(y) = p$ for all translations μ of $\mathbf{B}$.*

Moreover, $\mathbf{S} \in HS\mathbf{A}$ or $(x \wedge y) \vee (x \wedge z)$ is not a polynomial of $\mathbf{B}$.

LECTURE 2

The Eight Element Algebra A

The six element algebra which was constructed at the end of Lecture 0 generates a variety with a lot of finite subdirectly irreducible algebras in addition to the $\mathbf{Q}_n$'s. An example is the flat edge algebra $\mathbf{S}_8$ described next. For each $i < 8$, define the following elements of the 8-fold direct power of the algebra constructed in Lecture 0:

$$\begin{aligned} c &= \langle D,D,\dots,D,D,D,\dots\rangle \\ d_i &= \langle D,D,\dots,D,C,D,\dots\rangle \\ r_i &= \langle\ 2,2,\ \dots,2,\ H,2,\ \dots\rangle \end{aligned}$$

where the sole C in d_i and the sole H in r_i occur at position i. Let B_8 be the subset of the 8-fold direct power consisting of c, all the d_i's, all the r_i's, and all the 8-tuples which have 0 in at least one position. Then

$$r_i \cdot d_i = c$$

for every $i < 8$, but any other product of elements of B_8 results in an 8-tuple with 0 in at least one position. This means that B_8 is a subuniverse. Let θ_8 be the equivalence relation on $\mathbf{B}_8$ which collapses into one big block all the 8-tuples in B_8 which have 0 in at least one position, but which isolates all the other members into singletons. It is easy to see that θ_8 is a congruence relation of $\mathbf{B}_8$. Let $\mathbf{S}_8 = \mathbf{B}_8/\theta_8$. $\mathbf{S}_8$ is displayed below in Figure 1.

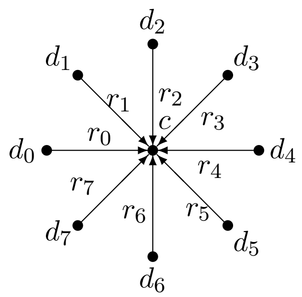


FIGURE 1. The directed graph for $\mathbf{S}_8$

Plainly, $\mathbf{S}_8$ is an algebra in the variety. A routine calculation shows that $(c, 0)$ is a critical pair in $\mathbf{S}_8$. Consequently, $\mathbf{S}_8$ is subdirectly irreducible, as desired. Evidently, there is nothing special in the choice of 8. A similar construction can be carried out for any cardinal in place of 8.

We must modify our little 6-element algebra from Lecture 0 to eliminate subdirectly irreducible algebras like $\mathbf{S}_8$, whose diagrams are not (finite) directed paths. Evidently, for our subdirectly irreducible algebras, we need a kind of unique factorization property:

$$a \cdot b = c \cdot d \neq 0 \Rightarrow a = c \text{ and } b = d.$$

To accomplish this we are going to add some new basic operations and some new elements to our algebra, but we need to have some care since we want $\mathbf{Q}_{\mathbb{Z}}$ to remain essentially unchanged and still to belong to the variety generated by the finite algebra we are trying to devise.

To obtain the unique factorization property we introduce the new basic 4-ary operation U^0:

$$U^0(x, y, z, w) = \begin{cases} xy & \text{if } xy = zw \neq 0 \text{ and } x = z \text{ and } y = w, \\ \overline{xy} & \text{if } xy = zw \neq 0 \text{ and either } x \neq z \text{ or } y \neq w, \\ 0 & \text{otherwise.} \end{cases}$$

At the moment, we should understand that the first case corresponds to the situation when the unique factorization property prevails, the second case corresponds to the failure of the unique factorization property, and the remaining case is just a default. For the moment, $\overline{xy}$ is simply a reminder that the output in this case should depend on xy but differ from xy. Our hope is to obtain the unique factorization property by forcing the first case to happen. In essence, this means preventing the second case. For this purpose we introduce a new basic 5-ary operation S_2:

$$S_2(u, v, x, y, z) = \begin{cases} (x \wedge y) \vee (x \wedge z) & \text{if } u = \overline{v}, \\ 0 & \text{otherwise.} \end{cases}$$

Recall the algebra **B** from Lecture 1. In **B** we know from Fact 5 that $(x \wedge y) \vee (x \wedge z)$ cannot be a polynomial. So S_2 is designed to prevent B from having elements u and v so that $u = \overline{v}$. This in turn will prevent the second case in the definition of U^0 from arising.

To give more sense to this, notice that in six element algebra from Lecture 0, a product xy could have only C, D, or 0 as a value. So we introduce two elements $\overline{C}$ and $\overline{D}$ in addition to the six with which we have been dealing. Further, we stipulate that $\overline{u} = C$ if $u = \overline{C}$ and likewise $\overline{u} = D$ if $u = \overline{D}$. In this way, both U^0 and S_2 have unambiguous definitions, once the product and meet have been extended to operations on the new set with eight elements.

These two additional operations and two additional elements are not quite enough.

$$S_1(u, v, x, y, z) = \begin{cases} (x \wedge y) \vee (x \wedge z) & \text{if } u \in \{1, 2\}, \\ 0 & \text{otherwise.} \end{cases}$$

The role of S_1, as we will see, is ensure that our finite subdirectly irreducible algebra **S** has another property that each $\mathbf{Q}_n$ has—namely, that the labels of the edges are not repeated. Last, here are the operations J and J' which are ternary:

$$J(x, y, z) = \begin{cases} x & \text{if } x = y \neq 0, \\ x \wedge z & \text{if } x = \overline{y}, \\ 0 & \text{otherwise.} \end{cases} \qquad J'(x, y, z) = \begin{cases} x \wedge z & \text{if } x = y \neq 0, \\ x & \text{if } x = \overline{y}, \\ 0 & \text{otherwise.} \end{cases}$$

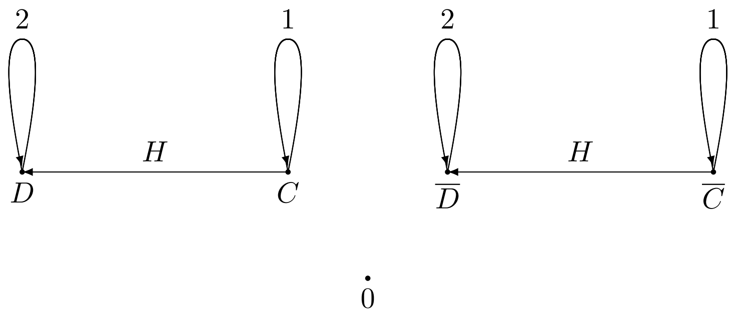
The role of these operations is less forthright. Since we are really working inside a subalgebra of a direct power, we have to contend with coordinate-wise properties. The role of these last two operations is to ensure that we fall into the "good" case at every coordinate.

We are led to an algebra **A** with eight elements and eight basic operations.

The universe is $A = \{0\} \cup \{1, H, 2\} \cup \{C, \overline{C}, D, \overline{D}\}$. We set $U = \{1, H, 2\}$ and $W = \{C, \overline{C}, D, \overline{D}\}$. We regard as an involution on W. The basic operations of **A** are denoted by $0, \wedge, \cdot, J, J', U^0, S_1$, and S_2. $\langle A, \wedge, 0\rangle$ is a flat semilattice with least element 0. The operation $\cdot$ is defined to give the default value 0 except when

$$\begin{aligned} 1 \cdot C &= C & \quad 1 \cdot \overline{C} &= \overline{C} \\ H \cdot C &= D & \quad H \cdot \overline{C} &= \overline{D} \\ 2 \cdot D &= D & \quad 2 \cdot \overline{D} &= \overline{D} \end{aligned}$$

This is an edge operation. Ordinarily, we represent the product $\cdot$ simply by juxtiposition. Here is the diagram of the edge algebra:



The following fact is evident from the definition of the product.

FACT 6. *If λ is a basic translation on* **A** *associated with the product $\cdot$, and $\lambda(a) = \lambda(b) \neq 0$, then $a = b$. The same is true for every translation built using only the product.*

Properties of B based on the Eight Element Algebra A

With the description of our eight element algebra $\mathbf{A}$ in hand, we continue to develop facts about $\mathbf{B}$ and its congruence θ. Denote by B_1 the set consisting of p and all its factors with respect to the product $\cdot$. That is

$$B_1 = \{u : \ \lambda(u) = p \text{ for some nonconstant translation } \lambda \text{ of } \mathbf{B} \text{ built only from the product}\}$$

So $u \in B_1$ if and only if $u = p$ or $u = c_i$ for some factorization $p = c_0 c_1 \dots c_m$ (where this latter product is associated to the right).

Let B_0 denote that set of those tuples in B which contain at least one 0. Plainly $B_0 \subseteq B - B_1$. It is also clear that if $\mathbf{S} \notin HS\mathbf{A}$, then the ranges of the operations S_1 and S_2 are contained in B_0 and hence in $B - B_1$.

The basic operation J of $\mathbf{A}$ is **monotone** in the sense that if $a \leq a', b \leq b'$ and $c \leq c'$ where all these elements belong to A, then $J(a, b, c) \leq J(a', b', c')$.

Fact 7. *Let f be a monotone unary polynomial of $\mathbf{B}$. If $x < p$ and $f(x) = p$, then $f(q) = p$.*

Proof: By Fact 4 we have $x \ \theta \ x \wedge q$. This entails $p = f(x) \ \theta \ f(x \wedge q)$. So by Fact 2 we get $p = f(x \wedge q)$. But then $p \leq f(q)$ by the monotonicity of f. Thus $p = f(q)$ by the maximality of p. □

Proviso: *The facts below are established under the assumption that the ranges of S_1 and S_2 are contained in $B - B_1$.*

Fact 8. *If $u \in B_1$ and $v \in B$ so that for all $s \in T$ either $u(s) = v(s)$ or $u(s) = \overline{v(s)} \in W$, then $u = v$.*

Proof: First suppose $u = p$.

Let $Y = \{s : p(s) = \overline{v(s)}\}$.

Claim: Y is empty.

Proof of the Claim: Since the range of the operation S_2 is disjoint from B_1, it follows that $T \neq Y$. Pick $t' \in T - Y$ and let $q' = \chi_{t'}$. So for each $s \in T$ we have

$$J(p(s), v(s), q'(s)) = \begin{cases} p(s) & \text{if } s \notin Y, \\ p(s) \wedge q'(s) & \text{if } s \in Y. \end{cases}$$

But this entails $J(p, v, q') = p$, since $q'(s) = p(s)$ for all $s \in Y$ because $t' \notin Y$. Therefore, by Fact 7 and the monotonicity of J, we have $J(p, v, q) = p$. But then the definition of J gives us $q(s) = p(s)$ for all $s \in Y$. Since $q(t_0) = 0$, it follows that $t_0 \notin Y$.

Now observe that $J'(p(t_0), v(t_0), q(t_0)) = p(t_0) \wedge 0 = 0$. Hence, $J'(p, v, q) \neq p$. So by Fact 7 and the monotonicity of J', we conclude that $J'(p, v, \chi_t) \neq p$ for all $t \in T$. But for all $s, t \in T$

$$J'(p(s), v(s), \chi_t(s)) = \begin{cases} p(s) \wedge \chi_t(s) & \text{if } s \notin Y, \\ p(s) & \text{if } s \in Y. \end{cases}$$

It follows that $t \notin Y$ for all $t \in Y$. This means Y is empty. So the Claim is established.

Since Y is empty, we also know that $p(s) = v(s)$ for all $s \in T$. Hence $u = v$ as desired.

Now suppose $u \in B_1 - \{p\}$. There are two kinds of elements in $B_1 - \{p\}$—those in U^T and those in W^T. Clearly, we can restrict our attention to the case when $u \in W^T$. Let λ be a translation built from the product such that $\lambda(u) = p$. Set $p' = \lambda(v)$. Since the product respects bars on elements, we see that for each $s \in T$, either $p(s) = p'(s)$ or $p(s) = \overline{p'(s)}$. So by the claim just established, we have $\lambda(u) = p = p' = \lambda(v)$. But then $u = v$ by Fact 6 □

Our basic strategy calls for θ to isolate the members of B_1 and to lump all the elements of $B - B_1$ together. To see that this really does happen, in view of Fact 3 we need the following.

FACT 9. *If $u \in B$ and $\lambda(u) \in B_1$ for some nonconstant translation λ, then $u \in B_1$.*

PROOF: The proof is by induction on the complexity of λ. The initial step of the induction is obvious, since the identity function is the only simplest nonconstant translation. The inductive step breaks down into seven cases, one for each basic operation of positive rank.

CASE $\wedge$: $\lambda(x) = \mu(x) \wedge r$, where $r \in B$.

We have $\lambda(u) \leq \mu(u)$. But every element of B_1 is maximal with respect to the semilattice order. So $\lambda(u) = \mu(u) \in B_1$. Now μ must be nonconstant. Invoking the induction hypothesis, we get $u \in B_1$.

CASE $\cdot$: $\lambda(x) = \mu(x)r$ or $\lambda(x) = r\mu(x)$.

Under the first alternative we have $\mu(u)r = \lambda(u) \in B_1$. So $\mu(u), r \in B_1$. Since μ must be nonconstant, we can invoke the induction hypothesis to conclude that $u \in B_1$. The other alternative is similar.

CASE J: $\lambda(x) = J(\mu(x), r, s)$ or $\lambda(x) = J(r, \mu(x), s)$ or $\lambda(x) = J(r, s, \mu(x))$.

Consider the first alternative. We have $\lambda(u) = J(\mu(u), r, s) \leq \mu(u)$. By the maximality of $\lambda(u)$ we get

$$\lambda(u) = J(\mu(u), r, s) = \mu(u) \in B_1.$$

Now μ cannot be constant. Hence we can invoke the inductive hypothesis to conclude that $u \in B_1$. The second alternative is similar, except that Fact 8 comes into play. Under the last alternative, since $r \geq J(r, s, \mu(u)) = \lambda(u)$ is maximal, we see that r and s fulfill the hypotheses of Fact 8. Consequently, $r = s \in B_1$. But then, $\lambda(x) = J(r, s, \mu(x)) = r$ according to the definition of J. This means the third alternative is impossible, since $\lambda(x)$ is not constant.

CASE J': $\lambda(x) = J'(\mu(x), \nu(x), \rho(x))$.

This case is easier than the last one and is omitted.

CASES S_1 AND S_2: Too easy.

CASE U^0: $\lambda(x) = U^0(\mu(x), s, r', s')$ or $\lambda(x) = U^0(r, \mu(x), r', s')$ or $\lambda(x) = U^0(r, s, \mu(x), s')$ or $\lambda(x) = U^0(r, s, r', \mu(x))$.

Consider the first alternative. We have $\lambda(u) = U^0(\mu(u), s, r', s') \in B_1$. Evidently, $\lambda(u)$ and $\mu(u)s$ satisfy the hypotheses of Fact 8. So $\lambda(u) = \mu(u)s$. Since $\lambda(u) \in B_1$, we know that $\mu(u) \in B_1$

by the definition of B_1. Now μ is nonconstant. So $u \in B_1$ by the inductive hypothesis. The second alternative is similar.

Consider the third alternative. We have $\lambda(u) = U^0(r, s, \mu(u), s')$. Evidently, $\lambda(u)$ and rs satisfy the hyopetheses of Fact 8. So $\lambda(u) = rs$. Then by the definition of T, we have $\lambda(u) = rs = \mu(u)s'$. But then $\mu(u) \in B_1$ and the induction hypotheses applies to yield $u \in B_1$. The fourth alternative is similar. □

FACT 10. *$u/\theta = \{u\}$ for each $u \in B_1$ and $0/\theta = B - B_1$.*

PROOF: Suppose $u \in B_1$ and that $u \;\theta\; v$. Let $\lambda(u) = p$ for some translation λ built just using $\cdot$. It follows that $\lambda(v) = p$ by Fact 3. By Fact 6, we conclude that $u = v$.

Fact 9 says that $B - B_1$ is closed with respect to nonconstant translations. Since $p \in B_1$, we have that $\lambda(u) \neq p$ for all $u \in B - B_1$ and all nonconstant translations λ. Hence, by Fact 3, $B - B_1$ is collapsed by θ. But, as we just saw, B_1 is the union of (singleton) θ-classes. Hence $B - B_1$ is a θ-class. Clearly, $0 \in B - B_1$. □

To establish that $\mathbf{S} \cong \mathbf{Q}_n$ for some natural number n we need to analyze each of our basic operations. We deal with the product first.

Here is the unique factorization property for the product that we require.

FACT 11. *If $ab = cd \in B_1$, then $a = c$ and $b = d$.*

PROOF: Let $u = ab$ and $v = U^0(a, b, c, d)$. From the definition of the operation U^0, we see that u and v satisfy the hypotheses of Fact 8. Hence, $ab = U^0(a, b, c, d)$. But then the definition of U^0 gives $a = c$ and $b = d$. □

In $\mathbf{Q}_{\mathbb{Z}}$ none of the labels of the edges were repeated. We need this property as well. It is the reason why we introduced the operation S_1. The relevant fact is next.

FACT 12. *No factorization of p has repeated factors.*

PROOF: It is clear that if $d_0 d_1 \dots d_{m-1} e = p$ then $e \in W^T$ while $d_0, \dots, d_{m-1} \in U^T$. Suppose that $d_i = d_j$ with $i < j$. Since the range of the operation S_1 is disjoint from B_1, we conclude that B contains no elements from $\{1, 2\}^T$. So pick $s \in T$ so that $d_i(s) = d_j(s) = H$. Now we see

$$p(s) = d_0(s) \dots d_{i-1}(s) H d_{i+1}(s) \dots d_{j-1}(s) H d_{j+1}(s) \dots d_{m-1}(s) e(s)$$

So $p(s) = 0$, violating the maximality of p. □

We are now in a position to describe B_1 more explicitly. Consider the following factorization of p:

$$\begin{aligned} p &= b_0 \\ &= a_0 b_1 \\ &= a_0 a_1 b_2 \\ &\vdots \\ &= a_0 a_1 \dots a_{n-1} b_n \\ &\vdots \end{aligned}$$

Evidently, $a_i \in U^T$ for all i and according to Fact 12 all the a_i's are distinct. But B_1 is finite, so we suppose without loss of generality that b_n cannot be factored. But the unique factorization property Fact 11 entails that the factorization of p displayed above is the only way p can be factored. Consequently,

$$B_1 = \{a_0, a_1, \ldots, a_{n-1}\} \cup \{b_0, b_1, \ldots, b_n\}$$

It is also evident that $b_i \in W^T$ for all i. Were $b_i = b_j$ for some $i \neq j$, it would be easy to construct a factorization of p with repeated factors, in violation of Fact 12. This means that B_1 has $2n+1$ elements, and that $b_i = a_i b_{i+1}$ for each $i < n$. That all the other products of elements chosen from B_1 will belong to B_0, follows easily from the unique factorization property Fact 11. Consequently, at least with respect to the product operation, **S** and $\mathbf{Q}_n$ are isomorphic.

Now consider the operation $\wedge$. Since $\wedge$ is obviously a semilattice operation on **S**, what we need is that **S** is flat.

FACT 13. *If $x, y \in B_1$ and $x \neq y$, then $x \wedge y \in B - B_1$.*

PROOF: Since $x \neq y$ there is $t \in T$ with $x(t) \neq y(t)$. But then $((x \wedge y)(t) = 0$. So $x \wedge y \in B - B_1$. □

Finally, we need to know that the remaining basic operations on **S** can be construed as term operations built up from $\cdot, \wedge$, and 0 in a manner dependent only on the hypotheses that **S** is a finite subdirectly irreducible algebra in $HSP\mathbf{A}$ and that $\mathbf{S} \notin HS\mathbf{A}$. That is the content of the next sequence of facts.

FACT 14. *$U^0(x, y, z, w)\ \theta\ (xy) \wedge (zw)$ for all $x, y, z, w \in B$.*

PROOF: We must show that either $U^0(x, y, z, w)$ and $(xy) \wedge (zw)$ both belong to $B - B_1$ or else $U^0(x, y, z, w) = (xy) \wedge (zw) \in B_1$. Since $B - B_1$ is a θ-class, Fact 8 forces $U^0(x, y, z, w) \in B - B_1$ except in the case that $xy = zw \in B_1$. In that case, $U^0(x, y, z, w) = xy = zw = (xy) \wedge (zw) \in B_1$. But also, $(xy) \wedge (zw) \in B - B_1$ except in the case that $xy = zw \in B_1$. In that case, $U^0(x, y, z, w) = xy = (xy) \wedge (zw) \in B_1$. Therefore, $U^0(x, y, z, w)\ \theta\ (xy) \wedge (zw)$. □

FACT 15. *$J(x, y, z)\ \theta\ x \wedge y$ for all $x, y, z \in B$.*

PROOF: Again, we must show that either $J(x, y, z)$ and $x \wedge y$ both belong to $B - B_1$ or else $J(x, y, z) = x \wedge y \in B_1$. Now again using that $B - B_1$ is a θ-class and Fact 8, $J(x, y, z) \in B - B_1$, except in the case that $x = y \in B_1$. In that case, $J(x, y, z) = x = y = x \wedge y \in B_1$. But also, $x \wedge y \in B - B_1$, except in the case that $x = y \in B_1$. In that case, $x \wedge y = x = J(x, y, z) \in B_1$. Therefore, $J(x, y, z)\ \theta\ x \wedge y$. □

FACT 16. *$J'(x, y, z)\ \theta\ x \wedge y \wedge z$ for all $x, y, z \in B$.*

PROOF: This is too easy. □

FACT 17. *$S_1(u, v, x, y, z)\ \theta\ 0\ \theta\ S_2(u, v, x, y, z)$ for all $u, v, x, y, z \in B$.*

For each natural number n, we take $\mathbf{Q}_n$ to be an algebra on $2n+2$ elements with the basic operations $\cdot, \wedge$, and 0 as described in Lecture 0, and the remaining basic operations determined by the stipulation that the following equations are true in $\mathbf{Q}_n$:

$$U^0(x,y,z,w) \approx (xy) \wedge (zw)$$
$$J(x,y,z) \approx x \wedge y \qquad S_1(u,v,x,y,z) \approx 0$$
$$J'(x,y,z) \approx x \wedge y \wedge z \qquad S_2(u,v,x,y,z) \approx 0$$

Thus we arrive at the desired conclusion.

LEMMA 1. *Let* $\mathbf{S}$ *be a finite subdirectly irreducible algebra in* $HSP\mathbf{A}$*. Either* $\mathbf{S} \in HS\mathbf{A}$ *or else there is a natural number* n *such that* $\mathbf{S} \cong \mathbf{Q}_n$*.*

What we haven't done in this lecture is prove that any of these expanded $\mathbf{Q}_n$'s belong to the variety generated by our 8-element algebra $\mathbf{A}$.

LECTURE 4

A is Inherently Nonfinitely Based and Has Residual Character ω_1

The algebra $\mathbf{Q}_{\mathbb{Z}}$ and its subalgebras $\mathbf{Q}_\omega$, and $\mathbf{Q}_n$ for each $n \in \omega$, were introduced in Lecture 0. The operations $0, \wedge$, and $\cdot$ were examined in detail, but the only stipulation about any remaining operations was that they must be defined as term operations of these first three. In Lecture 2, five more operation symbols were introduced: U^0, J, J', S_1, and S_2. In the algebras $\mathbf{Q}_{\mathbb{Z}}, \mathbf{Q}_\omega$, and $\mathbf{Q}_n$ these five further basic operations are defined so that the following equations are true:

$$U^0(x,y,z,w) \approx (xy) \wedge (zw)$$
$$J(x,y,z) \approx x \wedge y \qquad S_1(u,v,x,y,z) \approx 0$$
$$J'(x,y,z) \approx x \wedge y \wedge z \qquad S_2(u,v,x,y,z) \approx 0$$

The whole discussion of these algebras in Lecture 0 goes through in this expanded setting, with the exception of the last phase. The five new operations were not defined on the six element algebra $\mathbf{A}_0$ in Lecture 0. We now want to replace that algebra with the eight element algebra $\mathbf{A}$ introduced in Lecture 2. What we need is the following theorem to replace Theorem 1 of Lecture 0.

THEOREM 2. *$\mathbf{Q}_{\mathbb{Z}}$ belongs to the variety generated by $\mathbf{A}$.*

PROOF: We retrace the proof of Theorem 1. First, for each $p \in \mathbb{Z}$ we designate elements α_p and β_p of $A^{\mathbb{Z}}$ as before:

$$\begin{array}{lcccccccc} \alpha_p := & \dots & 1 & 1 & 1 & H & 2 & 2 & 2 & \dots \\ \beta_p := & \dots & C & C & C & D & D & D & D & \dots \end{array}$$

where the change is taking place at the p^{th} position. Next we let $B_1 = \{\alpha_p : p \in \mathbb{Z}\} \cup \{\beta_p : p \in \mathbb{Z}\}$ and we let $\mathbf{B}$ be the subalgebra of $\mathbf{A}^{\mathbb{Z}}$ generated by B_1. B_0 be the set of all elements of B in which 0 occurs. Now let Φ be the map defined from B to $Q_{\mathbb{Z}}$ via

$$\Phi(x) = \begin{cases} a_p & \text{if } x = \alpha_p \text{ for some } p \in \mathbb{Z}, \\ b_p & \text{if } x = \beta_p \text{ for some } p \in \mathbb{Z}, \\ 0 & \text{otherwise} \end{cases}$$

We contend that $B_0 \cup B_1$ is a subuniverse of $\mathbf{B}$ (and so $B = B_0 \cup B_1$) and also that Φ is a homomorphism from $\mathbf{B}$ onto $\mathbf{Q}_{\mathbb{Z}}$. Checking either of these contentions can be done by examining the behavior of each baisc operation case by case. We will do this simultaneously.

CASE 0. Plainly, $0 \in B_0$ and $\Phi(0) = 0$. So this case is secure.

CASE $\wedge$. Suppose that $u, v \in B_0 \cup B_1$. Then either $u = v$ and $u \wedge v = u \in B_0 \cup B_1$ or else $u \neq v$ and $u \wedge v \in B_0$. Hence, $B_0 \cup B_1$ is closed under $\wedge$. But also, $\Phi(u \wedge v) = \Phi(u) \wedge \Phi(v)$.

Case $\cdot$. Suppose that $u, v \in B_0 \cup B_1$. Then either $uv \in B_0$ or for some p we have $u = \alpha_p, v = \beta_{p+1}$ and $uv = \beta_p \in B_1$. It follows that $B_0 \cup B_1$ is closed under $\cdot$ and that Φ preserves $\cdot$.

To handle the remaining cases, the following property of $x, y \in B_0 \cup B_1$ proves useful:

$$(\star) \qquad \text{If for each } s \in \mathbb{Z}, \text{ either } x(s) = y(s) \neq 0 \text{ or } x(s) = \overline{y(s)}, \text{ then } x = y.$$

Since any $x, y \in B_0 \cup B_1$ for which the hypothesis of $\star$ holds must both belong to B_1, and since bars occur in no member of B_1, $\star$ is true.

Case J. For $J(x, y, z) \notin B_0$, observe that the inputs x and y must satisfy the hypothesis of $\star$. Hence, either $J(x, y, z) \in B_0$ or $x = y \in B_1$ and $J(x, y, z) = x \in B_1$. So $B_0 \cup B_1$ is closed under J and Φ preserves J.

Case J'. This case is very similar to the last case.

Case U^0. Let $x, y, z, w \in B_0 \cup B_1$. Let $u = U^0(x, y, z, w)$ and $v = xy$. Then $U^0(x, y, z, w) \in B_0$ unless u and v fulfill the hypothesis of $\star$. In that case, we must have $U^0(x, y, z, w) = u = v \in B_1$. Consequently, $B_0 \cup B_1$ is closed under U^0 and Φ preserves U^0.

Case S_1. Let $u, v, x, y, z \in B_0 \cup B_1$. Then $S_1(u, v, x, y, z) \in B_0$ unless $u \in \{1, 2\}^{\mathbb{Z}}$. But $\{1, 2\}^{\mathbb{Z}}$ and $B_0 \cup B_1$ are disjoint. Consequently, $B_0 \cup B_1$ is closed under S_1 and Φ preserves S_1.

Case S_2. Let $u, v, x, y, z \in B_0 \cup B_1$. Then $S_2(u, v, x, y, z) \in B_0$ unless $u(s) = \overline{v(s)}$ for all $s \in \mathbb{Z}$. But no element of B_1 has a bar at any of its entries. Consequently, $B_0 \cup B_1$ is closed under S_2 and Φ preserves S_2.

□

At this point we know that the eight element algebra **A**, which has eight basic operations, is inherently nonfinitely based, that the finite subdirectly irreducible algebras in the variety generated by **A** are the subdirectly irreducible algebras in $HS\mathbf{A}$ and the algebras $\mathbf{Q}_n$ for each $n \in \omega$, and that $\mathbf{Q}_\omega$ is a countably infinite subdirectly irreducible member of the variety.

We will demonstrate that our variety has no other infinite subdirectly irreducible algebras.

Let **S** be any infinite subdirectly irreducible algebra in the variety generated by **A**. According to the Theorem of Dziodiak and Quackenbush (see the Toolbox), any finite subalgebra of **S** can be embedded into arbitrarily large finite subdirectly irreducible algebras in the variety generated by **A**, i.e. into $\mathbf{Q}_n$ for all large enough n. This means that every finitely generated (= finite) subalgebra of **S** is embeddable into $\mathbf{Q}_\omega$. Consequently, every universal sentence true in $\mathbf{Q}_\omega$ must be true in **S**.

Here are some interesting properties of $\mathbf{Q}_\omega$ which can be expressed with universal sentences:

- Any equation true in $\mathbf{Q}_{\mathbb{Z}}$. For example: $U^0(x, y, z, w) \approx (xy) \wedge (zw)$.
- The height is no bigger than 1: $x \not\approx y \to x \wedge y \approx 0$.
- $xy \approx zw \not\approx 0 \to (x \approx z \;\&\; y \approx w)$.
- $xy \not\approx 0 \not\approx xz \to y \approx z$.
- $xy \not\approx 0 \not\approx zy \to x \approx z$.
- $xy \not\approx 0 \to zx \approx 0 \approx yw$.

Consequently, in **S**, the operations U^0, J, J', S_1, and S_2 are term functions (using the same terms as in $\mathbf{Q}_\omega$) in $0, \wedge$, and $\cdot$. We ignore them from now on. With respect to $\wedge$ and 0, **S** is a height 1 meet-semilattice with least element 0. So the balance of our analysis depends primarily on the product $\cdot$. Since $(xy)z \approx 0$ is true in $\mathbf{Q}_\omega$, we see that in $\mathbf{S}$, just as in $\mathbf{Q}_\omega$, only right-associated products can differ from 0. The last four properties itemized above put further and severe restrictions on the product in $\mathbf{S}$.

We make $S - \{0\}$ into a labelled directed graph as follows. We take as the vertex set those elements which are right factors, outputs or do not occur in nonzero products. We take as the set of labels those elements which are left factors in nonzero products. Our itemized properties entail that the set of vertices and the set of labels are disjoint. We put an edge from b to c and label it with a provided $ab = c$ in $\mathbf{S}$. Our itemized assertions ensure that a vertex can have outdegree at most 1, indegree at most 1, and that every edge has a uniquely determined label which occurs as a label of exactly one edge in the whole graph.

Let C be a connected component of our graph. Let θ_C be the equivalence relation that collapses all the vertices and labels in C to 0, but which isolates every other point. θ_C is a congruence of $\mathbf{S}$. Since $\mathbf{S}$ is subdirectly irreducible, it follows that our graph has only one component. This already implies that $\mathbf{S}$ is countably infinite. But more is true. There are only three possible countable connected graphs of this kind: the one associated with $\mathbb{Z}$ (and then we would have $\mathbf{S} \cong \mathbf{Q}_\mathbb{Z}$), the one associated with ω (and then we would have $\mathbf{S} \cong \mathbf{Q}_\omega$), and the one associated with the set of nonnegative integers (and then $\mathbf{S}$ would be isomorphic to an algebra we might as well call $\mathbf{Q}_{-\omega}$). But neither $\mathbf{Q}_\mathbb{Z}$ nor $\mathbf{Q}_{-\omega}$ is subdirectly irreducible. So $\mathbf{S}$ must be isomorphic to $\mathbf{Q}_\omega$.

We summarize the results in the following theorem.

Theorem 3. *The eight element algebra* $\mathbf{A}$*, which has only eight basic operations, is inherently nonfinitely based. The subdirectly irreducible algebras in the variety generated by* $\mathbf{A}$ *are, up to isomorphism, exactly the subdirectly irreducible homomorphic images of subalgebras of* $\mathbf{A}$*, the algebra* $\mathbf{Q}_\omega$*, and the algebra* $\mathbf{Q}_n$ *for each* $n \in \omega$.

This theorem settles in the negative some outstanding problems. We will say that a variety is *finitely generated* provided it is generated by a finite algebra with only finitely many fundamental operations. It is *residually small* if there is an upper bound on the cardinalities of its subdirectly irreducible algebras. It is *residually finite* if all its subdirectly irreducible algebras are finite. It is *residually very finite* if there is a finite upper bound on the cardinalities of its subdirectly irreducible algebras.

The R-S Conjecture: *Every finitely generated residually small variety is residually very finite.*

The Broader Finite Basis Speculation: *Every finitely generated residually small variety is finitely based.*

Theorem 3 is a counterexample to both of these. However, the two problems below are closely related and still open.

The Quackenbush Conjecture: *Every finitely generated residually finite variety is residually very finite.*

Park's Conjecture: *Every finitely generated residually finite variety is finitely based.*

LECTURE 5

How $\mathbf{A}(\mathcal{T})$ Encodes The Computations of $\mathcal{T}$

In this lecture we describe, in part, McKenzie's machine algebras and show how they capture the computations of Turing machines. Turing machines are finite objects, but the computations that they produce can be endless. So it is reasonable to expect to use a finite algebra to convey the information of any particular Turing machine. However, finite algebras are too small to hold arbitrary computations. The algebra $\mathbf{Q}_{\mathbb{Z}}$, however, suggests a way to grapple with arbitrary computations. The idea is to designate certain elements of the algebra as configurations of a Turing machine and draw labeled directed edges between configurations to represent the transitions of the machine computation. Then we try to realize these directed edges by new operations applied to certain elements. Next we try to find a finite algebra so that the whole thing is happening coordinatewise inside a big direct power. Finally, we will have to add further operations to control all the finite subdirectly irreducible algebras.

For a Turing machine $\mathcal{T}$, we devise a finite algebra $\mathbf{A}(\mathcal{T})$ which enlarges A (in order to have enough distinct elements to code configurations) by adding finitely many elements and which expands $\mathbf{A}$ by adjoining operations to emulate the transitions between configurations, as well as to keep control of the finite subdirectly irreducible algebras. But the analysis of computation itself will go on in $\mathbf{A}(\mathcal{T})^X$ for some large set X [think of $X = \mathbb{Z}$].

We conceive of a Turing machine $\mathcal{T}$ as having finitely many internal states $0, 1, \dots, m$. The machine is always launched in state 1 and we take 0 to be the unique halting state. The Turing machine $\mathcal{T}$ has a tape alphabet consisting of the symbols 0 and 1. The Turing machine itself is a finite collection of 5-tuples each of the form:

$$[i, \gamma, \delta, M, j]$$

This 5-tuple is the instruction, "If you are in state i and you are examining a tape square containing the symbol γ, then write the symbol δ on that square, move one square in the direction M (M must be either L for left or R for right), and pass into internal state j". We insist that no 5-tuple begin with 0 and that otherwise the machine must have exactly one instruction which begins $[i, \gamma, \dots]$ for each state i other than the halting state 0 and each tape symbol γ.

We say Q is a *configuration* for a Turing machine $\mathcal{T}$ provided $Q = \langle t, n, i\rangle$ where $t \in \{0,1\}^{\mathbb{Z}}$, $n \in \mathbb{Z}$, and i is one of the states of $\mathcal{T}$. The idea is that at some stage of a computation, the tape of the machine looks like t, the machine is focussed on square n and is itself in state i.

A significant problem we have to resolve comes from the fact that machine computations, at any given stage, happen at a particular location on the tape, and that these locations are arranged in a sequence with only the adjacent locations available for the next step in the computation. Thus some elements of our "computation algebra" which are used to label those directed edges must also fall into a sequence of “tape locations”. To make short work of this point we take the elements a_p of $\mathbf{Q}_{\mathbb{Z}}$ as a model of how elements fall into sequence. Looking at what we had to have in A to get these a_p's we recall:

$$\begin{array}{rcccccccc}
\alpha_p\colon & \dots & 1 & 1 & 1 & H & 2 & 2 & 2 & \dots \\
\alpha_{p+1}\colon & \dots & 1 & 1 & 1 & 1 & H & 2 & 2 & \dots \\
\alpha_{p+2}\colon & \dots & 1 & 1 & 1 & 1 & 1 & H & 2 & \dots
\end{array}$$

So in all our machine algebras we want a subset $U = \{1, H, 2\}$ making elements like the ones above available in direct powers. To impose the precedence above in the direct power, we impose $2 \prec 2 \prec H \prec 1 \prec 1$ on U. We also use $\prec$ to denote the coordinatewise relation in any direct power of a machine algebra. Suppose $\mathbf{B} = \mathbf{A}(\mathcal{T})^X$. A subset $F \subseteq B$ is *sequentiable* provided

- $F \subseteq U^X$,
- H occurs at least once in f, for each $f \in F$, and
- $\prec$ gives F a structure isomorphic to some convex substructure of the ordered set of integers.

Since H may occur at several places in such an f, sequentiable sets can be more complex than $\{\alpha_p : p \in \mathbb{Z}\}$. For a fixed sequentiable set F the index set X falls into natural pieces that help us see the structure. Look at the following display of the four element sequentiable set $F = \{f_0, f_1, f_2, f_3\}$.

$$\begin{array}{rccccccccccccc}
f_0\colon & 1 & 1 & H & 2 & 2 & 2 & H & 2 & 2 & 2 & 1 & H & 1 \\
f_1\colon & 1 & 1 & 1 & H & 2 & 2 & 1 & 2 & 2 & 2 & 1 & 1 & 1 \\
f_2\colon & 1 & 1 & 1 & 1 & 2 & 2 & 1 & H & 2 & H & 1 & 1 & 1 \\
f_3\colon & 1 & 1 & 1 & 1 & 2 & H & 1 & 1 & 2 & 1 & 1 & 1 & 1
\end{array}$$

Examining the 13 columns, we see that several are exactly the same. In this example the set X has 13 elements and some unspecified arrangement of these thirteen elements underlies the display above. But the particular arrangement of X is immaterial from the point of view of the algebra $\mathbf{A}(\mathcal{T})^X$. Thus we are free to rearrange X to make the precedence on F more transparent. Below is the result of such a rearrangement:

$$\begin{array}{rccccccccccccc}
f_0\colon & 1 & 1 & 1 & 1 & H & H & H & 2 & 2 & 2 & 2 & 2 & 2 \\
f_1\colon & 1 & 1 & 1 & 1 & 1 & 1 & 1 & H & 2 & 2 & 2 & 2 & 2 \\
f_2\colon & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & H & H & 2 & 2 & 2 \\
f_3\colon & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & H & 2 & 2
\end{array}$$

We have put all the columns consisting entirely of 1's to the left. Next we put all the columns beginning with H in position 0, then all columns with H in position 1, and so on. At the right we have placed all columns consisting entirely of 2's. Doing this, we see that there are only $6 = 4 + 2$ different kinds of columns possible:

$$\begin{array}{cccccc}
1 & H & 2 & 2 & 2 & 2 \\
1 & 1 & H & 2 & 2 & 2 \\
1 & 1 & 1 & H & 2 & 2 \\
1 & 1 & 1 & 1 & H & 2
\end{array}$$

This means our sequentiable set F partitions the index set X into 6 blocks. The blocks can be labeled X_L for the set of all indices of columns that are constantly 1, X_R for the set of all indices of columns that are constantly 2, and X_n for the set of all indices where the necessarily unique H occurs at the n^{th} position.

To simplify the presentation a bit and make the pictures understandable, once a sequentiable set F has been specified, we will assume that X is arranged in such a line so that the set X_L is an initial (or left) segment, X_R is a final segment (or right) segment, and the each X_n is placed at the obvious position on the line. Since at its biggest, F can be indexed only by $\mathbb{Z}$, we can accommodate such a line like picture if we are willing to place X_L at $-\infty$ and X_R at $+\infty$.

Now let F be the four element sequentiable set above but with the columns collapsed to 6 and arranged as in the last display, and let $Q = \langle t, 2, i\rangle$ be a configuration. We code Q by

$$\begin{array}{rcccccc} \beta\colon & C^0_{i,t(2)} & C^{t(0)}_{i,t(2)} & C^{t(1)}_{i,t(2)} & M^{t(2)}_i & D^{t(3)}_{i,t(2)} & D^0_{i,t(2)} \\ \text{block:} & X_L & X_0 & X_1 & X_2 & X_3 & X_R \end{array}$$

This gives a real forest of superscripts and subscripts and the truth is that we will need a few more to get to full generality. However, we can decode it a bit. The C's mean "left of the reading head". The D's mean "to the right of the reading head". M locates where the machine reading head is. The index i specifies the state of the machine. The subscript $t(2)$ tells what symbol is written on the tape square scanned by the reading head. Finally, the indices $t(j)$ tell us what is printed on the corresponding square of the tape, unless it is too far off to the left (in X_L) or too far off to the right (in X_R), in which case we have used 0 as a default value (other choices would be okay). So reading across the superscripts is like reading across the tape. In this way, each component of β carries a lot of information about the configuration.

Now X in this example had 13 elements rather than 6, so the β above is too short. However, by duplicating the entries in β the correct number of times (e.g. the first entry $C^0_{i,t(2)}$ should occur 4 times while the last entry $D^0_{i,t(2)}$ should occur twice) we would get a β of the correct length. That $|X| = 13$ is immaterial. But our particular sequentiable set had only four elements, it was indexed with the convex set $\{0, 1, 2, 3\}$, and we took $n = 2$ in our configuration. To get the general case, let I be any convex subset of $\mathbb{Z}$ and suppose that F is a sequentiable set indexed by I. Let $n \in I$ and let $Q = \langle t, n, i\rangle$ be a configuration. Then we use the β below as a code for Q and we say that *β codes Q over F.*

$$\beta(x) = \begin{cases} C^0_{i,t(n)} & \text{if } x \in X_L. \\ C^{t(j)}_{i,t(n)} & \text{if } x \in X_j \text{ and } j < n \text{ and } j \in I. \\ M^{t(n)}_i & \text{if } x \in X_j \text{ and } j = n \in I. \\ D^{t(j)}_{i,t(n)} & \text{if } x \in X_j \text{ and } n < j \in I. \\ D^0_{i,t(n)} & \text{if } x \in X_R. \end{cases}$$

Capturing the Transitions Between Configurations

To get a grip on how to handle the transition between configurations let $\mathbf{B} = \mathbf{A}(\mathcal{T})^{\mathbb{Z}}$ and let $F = \{\alpha_p : p \in \mathbb{Z}\}$. Then F is a sequentiable set indexed by $\mathbb{Z}$, and the partition imposed on $\mathbb{Z}$ by F consists of singleton sets $\{p\}$. Let $Q = \langle t, n, i\rangle$ be a configuration of $\mathcal{T}$, let $t(n) = \gamma$, and suppose that $[i, \gamma, \delta, L, j]$ is an instruction in $\mathcal{T}$. It also proves convenient to let $t(n-1) = \varepsilon$. Then $\mathcal{T}(Q) = \langle s, n-1, j\rangle$ is the configuration following Q in the computation of $\mathcal{T}$, where

$$s(k) = \begin{cases} \delta & \text{if } k = n, \\ t(k) & \text{otherwise.} \end{cases}$$

The configuration Q is coded over F by

$$\beta = \quad \dots \quad C^{t(n-3)}_{i,\gamma} \quad C^{t(n-2)}_{i,\gamma} \quad C^{\varepsilon}_{i,\gamma} \quad M^{\gamma}_{i} \quad D^{t(n+1)}_{i,\gamma} \quad D^{t(n+2)}_{i,\gamma} \quad D^{t(n+3)}_{i,\gamma} \quad \dots$$

whereas the configuration $\mathcal{T}(Q)$ is coded over F by

$$\mathcal{T}(\beta) = \quad \dots \quad C^{t(n-3)}_{j,\varepsilon} \quad C^{t(n-2)}_{j,\varepsilon} \quad M^{\varepsilon}_{j} \quad D^{\delta}_{j,\varepsilon} \quad D^{t(n+1)}_{j,\varepsilon} \quad D^{t(n+2)}_{j,\varepsilon} \quad D^{t(n+3)}_{j,\varepsilon} \quad \dots$$

$\mathcal{T}(\beta)$ differs from β in several ways. First, the two positions indexed by $n-1$ and n undergo a change of character from C to M and from M to D. Second, the remaining changes amount to changing γ to ε and i to j in various subscripts and superscripts. The idea is to effect this transition with a new operation for the machine instruction $[i,\gamma,\delta,L,j]$. Changes of the first kind have to do with two tape locations. Our new operation must combine the two location elements, α_{n-1} and α_n, with the configuration element β to produce the new configuration element $\mathcal{T}(\beta)$—our "instruction" operation should be ternary. To see what is needed to accomplish this, look at

$$\begin{array}{rcccccccccc}
\alpha_{n-1} = & \dots & 1 & 1 & H & 2 & 2 & 2 & 2 & \dots \\
\alpha_{n} = & \dots & 1 & 1 & 1 & H & 2 & 2 & 2 & \dots \\
\beta = & \dots & C^{t(n-3)}_{i,\gamma} & C^{t(n-2)}_{i,\gamma} & C^{\varepsilon}_{i,\gamma} & M^{\gamma}_{i} & D^{t(n+1)}_{i,\gamma} & D^{t(n+2)}_{i,\gamma} & D^{t(n+3)}_{i,\gamma} & \dots \\
\hline
\mathcal{T}(\beta) = & \dots & C^{t(n-3)}_{j,\varepsilon} & C^{t(n-2)}_{j,\varepsilon} & M^{\varepsilon}_{j} & D^{\delta}_{j,\varepsilon} & D^{t(n+1)}_{j,\varepsilon} & D^{t(n+2)}_{j,\varepsilon} & D^{t(n+3)}_{j,\varepsilon} & \dots
\end{array}$$

The instruction $[i,\gamma,\delta,L,j]$ makes no reference to ε (the symbol written on square $n-1$ of the tape). Since our operation must act coordinatewise, we will build ε into the operation itself. So to each machine instruction we will associate two ternary operations, one for each of the two possible values of ε. Since the machine instructions for a fixed Turing machine $\mathcal{T}$ are determined by their first two components we will denote the operations corresponding to the machine instruction above by $F_{i\gamma\varepsilon}$. What must happen in $\mathbf{A}(\mathcal{T})$ to accomplish the transition above is

$$\begin{aligned}
F_{i\gamma\varepsilon}(1,1,C^{\nu}_{i,\gamma}) &= C^{\nu}_{j,\varepsilon} \\
F_{i\gamma\varepsilon}(2,2,D^{\nu}_{i,\gamma}) &= D^{\nu}_{j,\varepsilon} \\
F_{i\gamma\varepsilon}(H,1,C^{\varepsilon}_{i,\gamma}) &= M^{\varepsilon}_{j} \\
F_{i\gamma\varepsilon}(2,H,M^{\gamma}_{i}) &= D^{\delta}_{j,\varepsilon}
\end{aligned}$$

We would like to declare that in $\mathbf{A}(\mathcal{T})$ the operation $F_{i\gamma\varepsilon}$ results in the default value 0 except in the cases above. Ultimately, this won't do since we will find it necessary to introduce barred versions of all those C's, D's, and M's with all the attached subscripts and superscripts in order to control the finite subdirectly irreducible algebras. So we will have to revisit the definition of $F_{i\gamma\varepsilon}$. For the present, it is no great distortion to think that all the other values are 0.

A similar analysis of right-moving instructions leads the ternary operations $F_{i\gamma\varepsilon}$ being defined (with caveats about barred elements) in $\mathbf{A}(\mathcal{T})$ via

$$\begin{aligned}
F_{i\gamma\varepsilon}(1,1,C^{\nu}_{i,\gamma}) &= C^{\nu}_{j,\varepsilon} \\
F_{i\gamma\varepsilon}(2,2,D^{\nu}_{i,\gamma}) &= D^{\nu}_{j,\varepsilon} \\
F_{i\gamma\varepsilon}(H,1,M^{\gamma}_{i}) &= C^{\delta}_{j,\varepsilon} \\
F_{i\gamma\varepsilon}(2,H,D^{\varepsilon}_{i,\gamma}) &= M^{\varepsilon}_{j}
\end{aligned}$$

With this definition, in $\mathbf{A}(\mathcal{T})^{\mathbb{Z}}$

$$F_{i\gamma\varepsilon}(\alpha_n,\alpha_{n+1},\beta) = \mathcal{T}(\beta)$$

provided β is as above, ε is the symbol on tape square $n+1$, and $[i, \gamma, \delta, R, j]$ is an instruction of $\mathcal{T}$. For a given Turing machine $\mathcal{T}$, the definition of $F_{i\gamma\varepsilon}$ is unambiguous, since whether $F_{i\gamma\varepsilon}$ should be left or right moving can be determined from $\mathcal{T}, i$, and γ.

These operations can be envisioned as edge operations, where, however, the edges representing a particular operation now have two labels.

$$C^{\nu}_{j,\varepsilon} \xleftarrow{\quad 1 \qquad\qquad 1 \quad} C^{\nu}_{i,\gamma}$$

$$D^{\nu}_{j,\varepsilon} \xleftarrow{\quad 2 \qquad\qquad 2 \quad} D^{\nu}_{i,\gamma}$$

$$M^{\varepsilon}_{j} \xleftarrow{\quad H \qquad\qquad 1 \quad} C^{\varepsilon}_{i,\gamma} \qquad\qquad C^{\delta}_{j,\varepsilon} \xleftarrow{\quad H \qquad\qquad 1 \quad} M^{\gamma}_{i}$$

$$D^{\delta}_{j,\varepsilon} \xleftarrow{\quad 2 \qquad\qquad H \quad} M^{\gamma}_{i} \qquad\qquad M^{\varepsilon}_{j} \xleftarrow{\quad 2 \qquad\qquad H \quad} D^{\varepsilon}_{i,\gamma}$$

Left-Moving Case Right-Moving Case

Here is a useful fact, apparent in the diagrams above.

FACT 18. *If λ basic translation on $\mathbf{A}(\mathcal{T})$ associated with one of the operations $F_{i\gamma\varepsilon}$, and $\lambda(a) = \lambda(b) \neq 0$, then $a = b$. The same is true for every translation built only using the basic operations $F_{i\gamma\varepsilon}$, various choices of i, γ, and ε allowed.*

On the basis of these definitions, we obtain the following very useful conclusion.

The Key Coding Lemma: *Let $\mathcal{T}$ be a Turing machine, and let X be a set. Let F be a sequentiable set for $\mathbf{A}(\mathcal{T})^X$ and let i be a nonhalting state of $\mathcal{T}$. Finally, let $\gamma, \varepsilon \in \{0, 1\}$ and let f, g, and β be any elements of $A(\mathcal{T})^X$.*

Then $F_{i\gamma\varepsilon}(f, g, \beta) = \mathcal{T}(\beta)$ if

- *β codes a configuration Q over F,*
- *i and γ are the first two components of the $\mathcal{T}$ instruction determined by Q,*
- *$f, g \in F$ with $f \prec g$ and these two elements refer to the two adjacent tape squares involved in the motion called for in the instruction,*
- *ε is the symbol in the square to which the reading head is being moved, and*
- *$\mathcal{T}(\beta)$ codes the configuration $\mathcal{T}(Q)$ over F;*

Otherwise 0 occurs in $F_{i\gamma\varepsilon}(f, g, \beta)$. □

LECTURE 6

$\mathbf{A}(\mathcal{T})$ and What Happens If $\mathcal{T}$ Doesn't Halt

The basic plan is to do for $\mathbf{A}(\mathcal{T})$ what we did for $\mathbf{A}$. We were able to prove for $\mathbf{A}$ three crucial things:

(1) $\mathbf{Q}_{\mathbb{Z}}$ is in the variety generated by $\mathbf{A}$ (and hence that variety was inherently nonfinitely based and had a countably infinite subdirectly irreducible member).
(2) Any finite subdirectly irreducible in the variety, except possibly a few very small ones, had a very well determined structure (in fact they were all embeddable into $\mathbf{Q}_{\mathbb{Z}}$).
(3) There were no other infinite subdirectly irreducible algebras in the variety.

It was the second point that compelled us to adjoin additional elements and operations to our original 6-element algebra. Having done that, we had to revisit the first point to assure ourselves that the new elements and operations were innocuous. The third point depended on the first two and the Dziobiak-Quackenbush Theorem.

Proceeding along the same lines with $\mathbf{A}(\mathcal{T})$ we are able to do the following:

(1) $\mathbf{Q}_{\mathbb{Z}}$ is in the variety generated by $\mathbf{A}(\mathcal{T})$, provided $\mathcal{T}$ does not halt.
(2) In the event that $\mathcal{T}$ halts, the cardinality of any finite subdirectly irreducible can be bounded by a function of the size of $\mathcal{T}$ and the number of tape squares it visits before halting.
(3) In the event that $\mathcal{T}$ halts, the variety generated by $\mathbf{A}(\mathcal{T})$ has no infinite subdirectly irreducible algebras.
(4) In the event that $\mathcal{T}$ halts, the variety generated by $\mathbf{A}(\mathcal{T})$ is finitely based.

In the second point, at the cost of adding more elements and more operations to our 8-element algebra $\mathbf{A}$, we can ensure that any sequentiable set arising in the construction of a finite subdirectly irreducible cannot be large enough to accommodate the full halting computation. (The idea is that being able to reach a "halting configuration" would force the forbidden $(x \wedge y) \vee (x \wedge z)$ to be a polynomial.) Then we need to argue that bounding the size of sequentiable sets entails a bound on the subdirectly irreducible algebra itself. In the first point, after making an inessential modification to $\mathbf{Q}_{\mathbb{Z}}$ to make it into an algebra of the correct similarity type, it is the inaccessibility of the codes of halting configurations that ensures that the extra operations we had to add to accomplish the second point are innocuous. The third point is an immediate consequence of Quackenbush's Theorem. The fourth point requires a tough proof due to Ross Willard.

The Algebra $\mathbf{A}(\mathcal{T})$

Let $\mathcal{T}$ be a Turing machine with states $0, 1, \ldots, m$. The universe of the algebra $\mathbf{A}(\mathcal{T})$ is easiest to describe in pieces. For each of the $4m + 4$ choices of $i = 0, 1, \ldots, m$ and $\gamma, \delta \in \{0, 1\}$, we need four distinct elements denoted by $C^{\delta}_{i,\gamma}, \overline{C^{\delta}_{i,\gamma}}, D^{\delta}_{i,\gamma}$, and $\overline{D^{\delta}_{i,\gamma}}$. For each of the $2m + 2$ choices of $i = 0, 1, \ldots, m$ and $\gamma \in \{0, 1\}$, we need two elements denoted by M^{γ}_i and $\overline{M^{\gamma}_i}$. The unbarred versions were needed to code configurations. The barred versions help us control the finite subdirectly irreducible algebras. Let V be the set comprised of all $20m + 20$ of these elements. We also let V_i denote the set of 20 elements of V whose first lower index is i. In particular, V_0 contains all the elements used in coding halting configurations. The universe of $\mathbf{A}(\mathcal{T})$ is just

$$A(\mathcal{T}) = \{0\} \cup U \cup W \cup V$$

where $U = \{1, H, 2\}$ and $W = \{C, \bar{C}, D, \bar{D}\}$. Thus the size of $\mathbf{A}(\mathcal{T})$ is $20m + 28$ where m is the number of nonhalting states of $\mathcal{T}$.

The old algebra $\mathbf{A}$ will be a subreduct of $\mathbf{A}(\mathcal{T})$. Indeed, we insist that $\wedge$ make $\mathbf{A}(\mathcal{T})$ into a height 1 meet-semilattice with least element 0, and that any product involving a new element results in 0. The definitions of the remaining old operations are changed little or not at all. Here are the J's:

$$J(x, y, z) = \begin{cases} x & \text{if } x = y \neq 0 \\ x \wedge z & \text{if } x = \bar{y} \in V \cup W \\ 0 & \text{otherwise.} \end{cases}$$

$$J'(x, y, z) = \begin{cases} x \wedge z & \text{if } x = y \neq 0 \\ x & \text{if } x = \bar{y} \in V \cup W \\ 0 & \text{otherwise.} \end{cases}$$

Along with the old S's we insert one more:

$$S_0(u, v, x, y, z) = \begin{cases} (x \wedge y) \vee (x \wedge z) & \text{if } u \in V_0, \\ 0 & \text{otherwise.} \end{cases}$$

$$S_1(u, v, x, y, z) = \begin{cases} (x \wedge y) \vee (x \wedge z) & \text{if } u \in \{1, 2\}, \\ 0 & \text{otherwise.} \end{cases}$$

$$S_2(u, v, x, y, z) = \begin{cases} (x \wedge y) \vee (x \wedge z) & \text{if } u = \bar{v} \in V \cup W, \\ 0 & \text{otherwise.} \end{cases}$$

Along with the old U^0 we insert two new operations $U^1_{i\gamma\varepsilon}$ and $U^2_{i\gamma\varepsilon}$ for each of the $4m$ choices of i, γ, and ε, where i is a nohalting state:

$$U^0(x,y,z,w) = \begin{cases} xy & \text{if } xy = zw \neq 0 \text{ and } x = z \text{ and } y = w \\ \overline{xy} & \text{if } xy = zw \neq 0 \text{ and } x \neq z \text{ or } y \neq w \\ 0 & \text{otherwise.} \end{cases}$$

$$U^1_{i\gamma\varepsilon}(x,y,z,w) = \begin{cases} F_{i\gamma\varepsilon}(x,y,w) & \text{if } x \prec z \text{ and } F_{i\gamma\varepsilon}(x,y,w) \neq 0 \text{ and } y = z \\ \overline{F_{i\gamma\varepsilon}(x,y,w)} & \text{if } x \prec z \text{ and } F_{i\gamma\varepsilon}(x,y,w) \neq 0 \text{ and } y \neq z \\ 0 & \text{otherwise.} \end{cases}$$

$$U^2_{i\gamma\varepsilon}(x,y,z,w) = \begin{cases} F_{i\gamma\varepsilon}(y,z,w) & \text{if } x \prec z \text{ and } F_{i\gamma\varepsilon}(y,z,w) \neq 0 \text{ and } x = y \\ \overline{F_{i\gamma\varepsilon}(y,z,w)} & \text{if } x \prec z \text{ and } F_{i\gamma\varepsilon}(y,z,w) \neq 0 \text{ and } x \neq y \\ 0 & \text{otherwise.} \end{cases}$$

Finally, we need the $4m$ ternary operations $F_{i\gamma\varepsilon}$ introduced in Lecture 4 (but extended to accomodate the barred elements of V) and one further unary operation which serves to set up initial configurations:

$$I(x) = \begin{cases} C^0_{1,0} & \text{if } x = 1, \\ M^0_1 & \text{if } x = H, \\ D^0_{1,0} & \text{if } x = 2, \\ 0 & \text{otherwise.} \end{cases}$$

Notice that for outputs other than 0, the operation I is one-to-one. In this way, the next fact is an extension of Fact 18

FACT 19. *If λ is any translation of $\mathbf{A}(\mathcal{T})$ build only from the basic operations I and $F_{i\gamma\varepsilon}$, various choices of i, γ, and ε allowed, and $\lambda(a) = \lambda(b) \neq 0$, then $a = b$.*

While all this is relatively intricate, the F's and the I plainly help us emulate the computations of the Turing machine. The role of the S's is to prevent certain kinds of elements from getting into the picture during the construction of finite subdirectly irreducible algebras. U^0 was crucial to get a kind of unique decomposition result for $\cdot$ in the finite subdirectly irreducible algebras. The U^1 and U^2 operations play a similar role in connection with the F operations.

What Happens If $\mathcal{T}$ Does Not Halt

Now we expand $\mathbf{Q}_{\mathbb{Z}}$ to the similarity type appropriate to $\mathcal{T}$ by insisting that all the following equations hold in the expansion:

$$\begin{aligned} U^0(x,y,z,w) &\approx (xy)\wedge(zw) & S_0(u,v,x,y,z) &\approx 0\\ J(x,y,z) &\approx x\wedge y & S_1(u,v,x,y,z) &\approx 0\\ J'(x,y,z) &\approx x\wedge y\wedge z & S_2(u,v,x,y,z) &\approx 0\\ F_{i\gamma\varepsilon}(x,y,w) &\approx 0 & I(x) &\approx 0\\ U^1_{i\gamma\varepsilon}(x,y,z,w) &\approx 0 & U^2_{i\gamma\varepsilon}(x,y,z,w) &\approx 0 \end{aligned}$$

for all choices of i,γ and ε.

This sort of inessential expansion leaves its key properties intact: any locally finite variety to which (this expanded) $\mathbf{Q}_{\mathbb{Z}}$ belongs will be inherently nonfinitely based, and $\mathbf{Q}_{\mathbb{Z}}$ has a countably infinite subalgebra $\mathbf{Q}_\omega$ which is subdirectly irreducible.

Theorem 4. *If $\mathcal{T}$ does not halt, then $\mathbf{Q}_{\mathbb{Z}}$ belongs to the variety generated by $\mathbf{A}(\mathcal{T})$. In particular, if $\mathcal{T}$ does not halt, then $\mathbf{A}(\mathcal{T})$ is inherently nonfinitely based and the variety it generates is not residually finite.*

Proof: We follow the pattern set in the proofs of Theorems 1 and 2. For each $p\in\mathbb{Z}$ we take $\alpha_p,\beta_p\in A(\mathcal{T})^{\mathbb{Z}}$ to be the same elements we used before:

$$\begin{array}{lcccccccc} \alpha_p:= & \dots & 1 & 1 & 1 & H & 2 & 2 & 2 & \dots\\ \beta_p:= & \dots & C & C & C & D & D & D & D & \dots \end{array}$$

where the change is taking place at the p^{th} position. Next we let $B_1=\{\alpha_p : p\in\mathbb{Z}\}\cup\{\beta_p : p\in\mathbb{Z}\}$ and we take $\mathbf{B}$ to be the subalgebra of $\mathbf{A}(\mathcal{T})^{\mathbb{Z}}$ generated by B_1. Let B_0 denote the subset of B consisting of all those $\mathbb{Z}$-tuples in B which contain at least one 0. The set $\{\alpha_p : p\in\mathbb{Z}\}$ is sequentiable and consists of all the tuples in B belonging to $U^{\mathbb{Z}}$, since none of the operations of $\mathbf{A}(\mathcal{T})$ ever produces an element of U. Now for every $p\in\mathbb{Z}$

$$\begin{array}{lcccccccc} I(\alpha_p):= & \dots & C^0_{1,0} & C^0_{1,0} & C^0_{1,0} & M^0_1 & D^0_{1,0} & D^0_{1,0} & D^0_{1,0} & \dots \end{array}$$

which gives the code of a configuration (the all-0 tape with the machine in state 1 reading square p). The $F_{i\gamma\varepsilon}$'s may now be applied, step by step, to produce the codes of further configurations reached as the computation of $\mathcal{T}$ proceeds. Plainly, all these codes of configurations belong to B. Let C denote the set of all these configuration codes. We will prove that $C\cup B_0\cup B_1$ is a subuniverse of $\mathbf{A}(\mathcal{T})^{\mathbb{Z}}$, and therefore $B=C\cup B_0\cup B_1$.

Now let Φ be the map defined from B to $Q_{\mathbb{Z}}$ via

$$\Phi(x)=\begin{cases} a_p & \text{if } x=\alpha_p \text{ for some } p\in\mathbb{Z},\\ b_p & \text{if } x=\beta_p \text{ for some } p\in\mathbb{Z},\\ 0 & \text{otherwise}\end{cases}$$

We contend that Φ is a homomorphism from $\mathbf{B}$ onto $\mathbf{Q}_{\mathbb{Z}}$. To verify this, as well as that $C \cup B_0 \cup B_1$ is a subuniverse, requires us to examine the behavior of each of our operations on $C \cup B_0 \cup B_1$. For each operation in turn, we show that this set in closed and that Φ preserves the operation.

Case 0: Evidently $0 = \ldots, 0, 0, 0, 0, \cdots \in B_0$ and so $\Phi(0) = 0$.

Case $\wedge$: Evidently, $u \wedge v = u$ if $u = v$ and $u \wedge v \in B_0$ if $u \neq v$, for all $u, v \in C \cup B_0 \cup B_1$. Hence, our set is closed under $\wedge$ and $\Phi(u \wedge v) = \Phi(u) \wedge \Phi(v)$.

Case $\cdot$: Clearly, $\alpha_p \cdot \beta_{p+1} = \beta_p$ for all $p \in \mathbb{Z}$, with all other $\cdot$-products resulting in elements of B_0. So our set is closed under $\cdot$ and Φ preserves $\cdot$.

Case $F_{i\gamma\varepsilon}$: According to the Key Coding Lemma, the results of applying $F_{i\gamma\varepsilon}$ to members of $C \cup B_0 \cup B_1$ lie in $C \cup B_0$. Hence, $C \cup B_0 \cup B_1$ is closed under this operation and Φ preserves the operation.

Case I: Applied to elements of $C \cup B_0 \cup B_1$, I produces only elements of $C \cup B_0$. Hence, $C \cup B_0 \cup B_1$ is closed with respect to I, and Φ preserves I.

Observe that no barred elements occur in any of the members of $C \cup B_1$. It follows that

$(\star)$ if $u, v \in C \cup B_0 \cup B_1$ with $u(p) = v(p) \neq 0$ or $u(p) = \overline{v(p)} \in V \cup W$ for all $p \in \mathbb{Z}$, then $u = v$.

Case J: Evidently, $J(x, y, z) \in B_0$ if $x \in B_0$ or $y \in B_0$ or $x \neq y$, according to $(\star)$. Otherwise, $J(x, y, z) = x$. This entails that $C \cup B_0 \cup B_1$ is closed under J and Φ preserves J.

Case J': Likewise, $J'(x, y, z) \in B_0$ if $x \in B_0$ or $y \in B_0$ or $x \neq y$, according to $(\star)$. Otherwise, $J'(x, y, z) = x \wedge z$. This entails that $C \cup B_0 \cup B_1$ is closed under J and Φ preserves J.

Case U^0: If $xy \in B_0$ or $zw \in B_0$ or $xy \neq zw$, then we have $U^0(x, y, z, w) \in B_0$ and $(xy) \wedge (zw) \in B_0$, for any elements $x, y, z, w \in C \cup B_0 \cup B_1$. On the other hand, if $xy = zw \notin B_0$ it must be that $x = z = \alpha_p$ and $y = w = \beta_{p+1}$ for some $p \in \mathbb{Z}$. In that case, $U^0(x, y, z, w) = xy = zx = (xy) \wedge (zw)$. Thus, $C \cup B_0 \cup B_1$ is closed under U^0 and Φ preserves U^0.

Observe that for $u, v \in C \cup B_0 \cup B_1$, we have $u \prec v$ only when $u = \alpha_p$ and $v = \alpha_{p+1}$ for some $p \in \mathbb{Z}$. In particular,

$(*)$ With respect to $\prec$, every element of $C \cup B_0 \cup B_1$ has at most one predecessor and at most one successor.

Case $U^1_{i\gamma\varepsilon}$: In case $F_{i\gamma\varepsilon}(x, y, w) \in B_0$ or $x \not\prec z$, we have $U^1_{i\gamma\varepsilon}(x, y, z, w) \in B_0$. In the alternative case, it follows from the definition of $F_{i\gamma\varepsilon}$ that $x \prec y$. In view of $(*)$ it must be that $y = z$. So $U^1_{i\gamma\varepsilon}(x, y, z, w) = F_{i\gamma\varepsilon}(x, y, w) \in C$. Therefore, the application of $U^1_{i\gamma\varepsilon}$ always results in an element of $C \cup B_0$. Consequently, $C \cup B_0 \cup B_1$ is closed with respect to $U^1_{i\gamma\varepsilon}$ and Φ preserves this operation.

Case $U^2_{i\gamma\varepsilon}$: This case is like the one above, but it exploits the uniqueness of predecessors instead of successors.

Case S_0: Since $\mathcal{T}$ does not halt, the set $V_0^{\mathbb{Z}}$ is disjoint from $C \cup B_0 \cup B_1$. It follows that the application of S_0 always results in an element of B_0. Thus S_0 is preserved by Φ and $C \cup B_0 \cup B_1$ is closed with respect to S_0. **It should be noted that this is the sole place in the argument where the fact the $\mathcal{T}$ does not halt comes into play.**

CASE S_1: The set $\{1,2\}^{\mathbb{Z}}$ is disjoint from $C \cup B_0 \cup B_1$. It follows that the application of S_1 always results in an element of B_0. Thus S_1 is preserved by Φ and $C \cup B_0 \cup B_1$ is closed with respect to S_1.

CASE S_2: It follows from $\star$ that the application of S_2 always results in an element of B_0. Thus S_2 is preserved by Φ and $C \cup B_0 \cup B_1$ is closed with respect to S_2.

So $\mathbf{Q}_{\mathbb{Z}}$ belongs to the variety generated by $\mathbf{A}(\mathcal{T})$. $\square$

LECTURE 7

When $\mathcal{T}$ Halts: Finite Subdirectly Irreducible Algebras of Sequentiable Type

Throughout this lecture we assume that $\mathcal{T}$ is a Turing machine that eventually halts when started on the all-0 tape. We denote by $\pi(\mathcal{T})$ the number of squares examined by $\mathcal{T}$ in the course of its computation. Thus $\pi(\mathcal{T})$ is the length of the stretch of tape which comes into use for this computation. Our ambition is to describe all the finite subdirectly irreducible algebras in the variety generated by $\mathbf{A}(\mathcal{T})$, or at any rate to bound their size. From the facts developed in Lectures 1 and 2 we already have a lot of information at our disposal. Once again we take $\mathbf{S}$ to be a finite subdirectly irreducible algebra in the variety and we fix a finite set T, $\mathbf{B}$, and θ, so that

- $\mathbf{B} \subseteq \mathbf{A}(\mathcal{T})^T$
- θ is strictly meet-irreducible in $\operatorname{Con}\mathbf{B}$.
- $\mathbf{S}$ is isomorphic to $\mathbf{B}/\theta$.
- T is as small as possible for representing $\mathbf{S}$ in this way.
- $|T| > 1$ (i.e. $\mathbf{S} \notin HS\mathbf{A}(\mathcal{T})$).

Among other things, we know that $(x \wedge y) \vee (x \wedge z)$ is not a polynomial of $\mathbf{B}$ (Fact 5). We also have an element $p \in B$ so that $(p, 0)$ is critical over θ. In Lecture 2 the analysis revealed that all the elements of S, except 0, arose from a unique longest factorization of p using the product $\cdot$. We want, loosely speaking, to do the same thing now; but the machine operations I and $F_{i\gamma\varepsilon}$ have to be considered along with $\cdot$. We will change the definition of B_1. Thus, the facts that grew out of our analysis of the old version of B_1 must be re-examined. Also, Fact 9 was proved using an analysis by cases, with one case for each basic operation. Now we have more operations. Finally, we have modified all the old operations by extending their domains, (in the case of J, J', and S_2, we have done this by treating the new elements in V like the elements in W). However, in all its essential features the old analysis can be carried forward.

We take B_0 to be the collection of all elements of B which contain at least one 0. In $\mathbf{B}$ the ranges of S_0, S_1, and S_2 lie entirely in B_0. Moreover, V_0^T and $\{1,2\}^T$ are disjoint from B and there are no elements $u, v \in B$ so that $u = \bar{v} \in (V \cup W)^T$. This is just a direct consequence of Fact 5.

Fact 20. *Every sequentiable subset of B has fewer than $\pi(\mathcal{T})$ members.*

Proof: By the Key Coding Lemma any large enough sequentiable set would allow us, using I and the $F_{i\gamma\varepsilon}$'s, to emulate in $\mathbf{B}$ the entire halting computation of $\mathcal{T}$, producing an element of V_0^T in B. Then, via S_2, $(x \wedge y) \vee (x \wedge z)$ would be a polynomial of $\mathbf{B}$. □

Next we restate a part of Fact 8 in our expanded setting. The only difference is the insertion of V in the statement and the proof.

Fact 21. *If $v \in B$ and $p(s) = v(s)$ or $p(s) = \overline{v(s)} \in V \cup W$ for all $s \in T$, then $p = v$.* □

The next fact splits our analysis into two cases.

FACT 22. *Either $p \in V^T$ or $p \in W^T$.*

PROOF: First notice that there must be a nonconstant unary polynomial f and $u \in B$ with $f(u) = p$ but $u \neq p$. Otherwise, it follows from Fact 3 that $B - \{p\}$ is a θ-class. This means that our subdirectly irreducible algebra $\mathbf{S}$ has only two elements, and indeed is isomorphic to a subalgebra of $\mathbf{A}(\mathcal{T})$. This contradicts our assumption that T has at least two elements.

Let λ be a nonconstant unary polynomial of least complexity so that for some $u \in B$ with $u \neq p$ we have $\lambda(u) = p$. Also fix such a u. Now the rest of the argument falls into cases according to the leading operation symbol of λ.

CASE $\wedge$: $\lambda(x) = \mu(x) \wedge r$. Then $p = \mu(u) \wedge r$ Since p is maximal, we conclude that $p = \mu(u)$. This leads to a violation of the minimality of λ.

CASE $\cdot$: The range of λ is included in $B_0 \cup W^T$. This means $p \in W^T$.

CASE I: The range of λ is included in $B_0 \cup V^T$. This means $p \in V^T$.

CASES $F_{i\gamma\varepsilon}$: The range of λ is included in $B_0 \cup V^T$. So $p \in V^T$.

CASES S_i: Impossible: the range of each S_i is included in B_0.

CASES $U^0, U^j_{i\gamma\varepsilon}$: These cases put $p \in W^T$ (for U^0) or $p \in V^T$ (for $U^j_{i\gamma\varepsilon}$'s).

CASE J: $\lambda(x) = J(\mu(x), r, s)$, or $\lambda(x) = J(r, \mu(x), s)$, or $\lambda(x) = J(r, s, \mu(x))$. Under the first alternative, $p = \lambda(u) = J(\mu(u), r, s) \leq \mu(u)$. Then $p = \mu(u) = r$ by Fact 21 and the maximality of p. This violates the minimality of λ. The same reasoning applies to the second alternative. So consider the last alternative. Then $p = J(r, s, \mu(u)) \leq r$. Then $p = r$, and so Fact 21 implies that $p = r = s$. But this means that $\lambda(x) = J(p, p, \mu(x)) = p$, and so λ is constant. This case is impossible.

CASE J': This is like the last case, but easier. □

$\mathbf{S}$ is of *sequentiable* type if $p \in W^T$ and of *machine* type otherwise.

FACT 23. *Finite subdirectly irreducibles of sequentiable type have fewer than $2\pi(\mathcal{T})$ members.*

PROOF: We can just follow the old analysis for $\mathbf{A}$, paying a modest amount of attention to the additional operations, and observing that a sequentiable set arises in a natural way.

Now $p \in W^T$. Let B_1 be the set of all factors of p with respect to $\cdot$. Now *all* our previously established facts hold, as is evident in all cases except for Fact 9. This fact asserts that, *if $u \in B$ and $\lambda(u) \in B_1$ for some nonconstant translation λ, then $u \in B_1$*. The proof of Fact 9 relied on a case-by-case analysis according to the leading operation symbol. To get a proof for Fact 9 in our expanded similarity type, we have to consider the operations $I, F_{i\gamma\varepsilon}, U^1_{i\gamma\varepsilon}, U^2_{i\gamma\varepsilon}$, and S_0. (Actually, there are also minor changes in the definitions of J, J', and S_2, which merit a small amount of attention not provided here.) All these cases are trivial because $\lambda(u) \notin B_1$ for any u if the leading operation is any of these, since $B_1 \subseteq U^T \cup W^T$.

As in our analysis for $\mathbf{A}$, we have $B_1 = \{a_0, a_1, \ldots, a_{n-1}\} \cup \{b_0, b_1, \ldots, b_n\}$ where $b_k = a_k b_{k+1}$ for all $k < n$ and $b_0 = p$. Also $B - B_1$ is the θ-class of 0, B_1 splits into singletons modulo θ, and $a_k \in U^T$ and $b_k \in W^T$ for all k. It remains to see that $\{a_k : k < n\}$ is a sequentiable set. Since $\pi(\mathcal{T})$ bounds the size of sequentiable sets, we would be finished. We need $a_k \prec a_{k+1}$ for all $k < n - 1$. Let $t \in T$, and suppose first that $a_{k+1}(t) = 1$. Then $b_{k+1}(t) \in \{C, \overline{C}\}$, so $a_k(t) \in \{1, H\}$. Hence $a_k(t) \prec a_{k+1}(t)$. Next, suppose that $a_{k+1}(t) = H$. Then $b_{k+1}(t) \in \{D, \overline{D}\}$, so $a_k(t) = 2$. Hence,

$a_k(t) \prec a_{k+1}(t)$. Finally, suppose $a_{k+1}(t) = 2$. Then $b_{k+1}(t) \in \{D, \overline{D}\}$, so $a_k(t) = 2 \prec 2 = a_{k+1}(t)$. thus, $a_k \prec a_{k+1}$ and $\{a_k : k < n\}$ is sequentiable. $\square$

LECTURE 8

When $\mathfrak{T}$ Halts: Finite Subdirectly Irreducible Algebras of Machine Type

We now consider the case when the finite subdirectly irreducible algebra **S** introduced in Lecture 6 is of machine type. So we have $p \in V^T$. In this case, we let B_1 be the smallest subset of B which includes p and which is closed under the inverses of all the machine operations I and $F_{i\gamma\varepsilon}$. Hence,

$$\begin{aligned} B_1 = \{u : \lambda(u) = p \text{ for some nonconstant translation } \lambda \text{ of } \mathbf{A}(\mathfrak{T}) \\ \text{built only from the machine operations}\} \end{aligned}$$

It is easy to see that since $p \in V^T$, then $B_1 \subseteq U^T \cup V^T$. It also follows that if λ is a translation built up from the machine operations, and $\lambda(u) = p$, then all the coefficients of λ also belong to B_1.

Since we have now substantially altered the definition of B_1, we will need to re-examine Facts 8 and 9. Here is the new version of Fact 8. It is an immediate consequence of Fact 21 and Fact 19.

FACT 24. *If $u \in B_1$ and $v \in B$ so that for all $s \in T$ either $u(s) = v(s)$ or $u(s) = \overline{v(s)} \in V \cup W$, then $u = v$.*

Here is the new version of Fact 9. The statement has not changed, but the proof is different, accommodating the change in the definition of B_1.

FACT 25. *If $u \in B$ and $\lambda(u) \in B_1$ for some nonconstant translation λ, then $u \in B_1$.*

PROOF: The proof is by induction on the complexity of λ. The initial step of the induction is obvious, since the identity function is the only simplest nonconstant translation. For the inductive step we take $\lambda(x) = \nu(\mu(x))$, where $\nu(x)$ is a basic translation and $\mu(x)$ is a translation with smaller complexity than λ. The work breaks down into cases according to the basic operation associated with ν.

CASE $\wedge$: $\lambda(x) = \mu(x) \wedge r$. But every element of B_1 is maximal with respect to the semilattice order. So $\lambda(u) = \mu(u) \in B_1$. Invoking the induction hypothesis for $\mu(x)$, we get $u \in B_1$.

CASE $\cdot$: This cannot happen since then the range λ would be included in $B_0 \cup W^T$, which is disjoint from B_1.

CASES $F_{i\gamma\varepsilon}$: Since $\nu(\mu(u)) = \lambda(u) \in B_1$, it follows from the definition of B_1, that $\mu(u) \in B_1$. Now the induction hypothesis applies.

CASE I: $\lambda(x) = I(\mu(x))$. By the definition of B_1, $\mu(u) \in B_1$. So the induction hypothesis applied.

CASE J: $\nu(x) = J(v, y, z)$, where x is one of v, y, and z, while the remaining two are coefficients. First, suppose x is either v or y. From Fact 24 and the maximality of the members of B_1 it follows that $\mu(u) = \lambda(u) \in B_1$. So the induction hypothesis applies. Now suppose x is z and son v and y are coefficients. In this case, it follows from Fact 24 that $v = y = \lambda(u) \in B_1$. But this means that $\nu(x) = v$ and so λ is constant. That cannot happen.

CASE J': This case is easier than the last one and its discussion is omitted.
CASES S_0, S_1 AND S_2: Too easy—the range of λ would be included in B_0.
CASE U^0: This cannot happen since the range of λ would be included in $B_0 \cup W^T$, which is disjoint from B_1.
CASES $U^j_{i\gamma\varepsilon}$: $\nu(x) = U^j_{i\gamma\varepsilon}(v, y, z, w)$, where exactly one of v, y, z, and w is x and the reamining ones are coefficients, which we will regard as constant functions.

The other case being similar, we suppose that $j = 1$. Evidently, $\lambda(u)$ and $F_{i\gamma\varepsilon}(v(u), y(u), w(u))$ satisfy the hypotheses of Fact 24. So $\lambda(u) = F_{i\gamma\varepsilon}(v(u), y(u), w(u)) = F_{i\gamma\varepsilon}(v(u), z(u), w(u))$ (since also $y(u) = w(u)$) follows from the definition of $U^1_{i\gamma\varepsilon}$. So $v(u), y(u), z(u), w(u) \in B_1$, by the definition of B_1. So $\mu(u) \in B_1$ and the induction hypothesis applies. □

Here is the new version of Fact 10. Again, the statement is the same, but B_1 has a new meaning. The proof it like that for Fact 10, but it uses Fact 25 in place of Fact 9 and Fact 19 in place of Fact 6.

FACT 26. *$u/\theta = \{u\}$ for each $u \in B_1$ and $0/\theta = B - B_1$.*

Thus to bound the cardinality of $\mathbf{S}$ we need to bound $|B_1|$. This will be the focus of our efforts in the next lecture. However, here we can remark that in fact a complete analysis of finite subdirectly irreducible algebras of machine type, as well as those of sequentiable type, is at hand. This further analysis would describe the behavior of all the operations. We will not pursue this more detailed analysis, except to point out that all these subdirectly irreducible algebras are flat.

LECTURE 9

When $\mathcal{T}$ Halts: Bounding the Subdirectly Irreducibles

In this lecture we will complete our analysis of the subdirectly irreducible algebras generated by $\mathbf{A}(\mathcal{T})$ in the case when $\mathcal{T}$ halts. Fact 23 already provides a bound on the size of the finite subdirectly irreducible algebras of sequentiable type. The last lecture provided a description of the finite subdirectly irreducible algebras of machine type. Our next task is to bound the size of these algebras. So we continue to consider the case when $\mathbf{S}$ is of machine type.

We can suppose that no component of $p \in V^T$ is a barred element. (The basic reason is that the operations $F_{i\gamma\varepsilon}$ do not alter whether a symbol is barred. Hence the distribution of bars in any member of $B_1 \cap V^T$ is the same as the distribution of bars in p.) Now $B_1 \subseteq U^T \cup V^T$. Let $\Omega = B_1 \cap V^T$ and $\Sigma = B_1 \cap U^T$. Look first in more detail at Ω. We define Ω_n by the following recursion.

$$\begin{aligned}\Omega_0 &= \{p\}\\ \Omega_{n+1} &= \Omega_n \cup \{u \in B_1 : F_{i\gamma\varepsilon}(f,g,u) \in \Omega_n \text{ for some } f,g \in B \text{ and some } i,\gamma,\varepsilon\}\end{aligned}$$

Evidently, $\Omega = \bigcup_n \Omega_n$. We will say that $f \in U^T$ **matches** $v \in V^T$ provided for all $t \in T$

$$\begin{aligned} f(t) = 1 &\Leftrightarrow v(t) \text{ is a } C^{\nu}_{i\gamma}\\ f(t) = H &\Leftrightarrow v(t) \text{ is an } M^{\gamma}_{i}\\ f(t) = 2 &\Leftrightarrow v(t) \text{ is a } D^{\nu}_{i\gamma}\end{aligned}$$

Observe that every $v \in V^T$ matches exactly one $f \in U^T$. For each natural number n, we let $\Sigma_n = \{f \in \Sigma :\ f \text{ matches } v \text{ for some } v \in \Omega_n\}$. By referring to the definition of $F_{i\gamma\varepsilon}$, we have that the elements of the two element set $\{f, g\}$ match the elements of the two element set $\{u, v\}$ whenever $F_{i\gamma\varepsilon}(f, g, u) = v \in \Omega$ (the order in which this matching occurs depends on whether the underlying Turing machine instruction is right-moving or left-moving). It follows that $\Sigma = \bigcup_n \Sigma_n$.

Fact 27. *Σ is a sequentiable set.*

Proof: We argue by induction that Σ_n is sequentiable.

Initial Step: Observe that Σ_0 has only one element. (Σ_0 cannot be empty, since then our subdirectly irreducible $\mathbf{S}$ would be in $HS\mathbf{A}(\mathcal{T})$.) Since $\Sigma_0 \subseteq B_1 \cap U^T$ and B is disjoint for $\{1, 2\}^T$, we see that its element has to have H in at least one place. Thus, Σ_0 is a sequentiable set.

Inductive Step: Suppose $h \in \Sigma_{n+1} - \Sigma_n$. Pick $u \in \Omega_{n+1} - \Omega_n$ so that h matches u. Further, pick $F_{i\gamma\varepsilon}, f, g$, and v so that $F_{i\gamma\varepsilon}(f, g, u) = v \in \Omega_n$. It does no harm to suppose that we have a left-moving operation. So g matches u and f matches v. It follows that $h = g$, that $f \in \Sigma_n$, and that $f \prec g$. By the inductive hypothesis, we have that Σ_n is sequentiable. Let us display Σ_n as

$$f_a \prec f_{a+1} \prec \dots f_b$$

In the event that $f = f_b$ we have $\Omega_n \cup \{h\}$ sequentiable as desired. On the other hand, if $f = f_c$ for some $c < b$, then, in view of Fact 24, we know $U^1_{i\gamma\varepsilon}(f, h, f_{c+1}, u) = F_{i\gamma\varepsilon}(f, h, u)$. So we would be able to conclude that $h = f_{c+1} \in \Sigma_n$, contrary to our choice of h. Reasoning in the same way, we see that it is not possible that Σ_{n+1} extends Σ_n on the right in any more elaborate way. Indeed, suppose $h' \in \Sigma_{n+1} - \Sigma_n$ and that $F_{i'\gamma'\varepsilon'}(f_b, g', u') = v' \in \Omega_n$, where h' matches u'. We take this operation to be left-moving. Then from $U^1_{i'\gamma'\varepsilon'}(f_b, h', h, u') = F_{i'\gamma'\varepsilon'}(f_b, h', u')$ we are able to conclude that $h = h'$.

Right-moving operations are handled in a way similar to what we just did for left-moving operations, but using $U^2_{i\gamma\varepsilon}$. □

FACT 28. *Σ has fewer than $\pi(\mathcal{T})$ elements.* □

To obtain a bound on the cardinality of Ω we must recall that the sequentiable set Σ partitions T into $T_L, T_a, \dots, T_b, T_R$ where $\Sigma = \{f_a, \dots, f_b\}$.

FACT 29. *$u \restriction T_c$ is constant for each $u \in \Omega$ and each $c \in \{a, \dots, b\}$.*

PROOF: The proof is accomplished in stages, each stage showing that more elements of Ω are constant on more T_c's until everything is accomplished. This proof needs some preliminary observations.

Suppose that $u \in \Omega_{n+1} - \Omega_n$ with $F_{i\gamma\varepsilon}(f_c, f_{c+1}, u) = v \in \Omega_n$. In this case we will say that u, c and $c+1$ become *active* at stage $n+1$. (We regard p as the only element active at stage 0 and no member of $c \in \{a, \dots, b\}$ as active at stage 0.) The definition of $F_{i\gamma\varepsilon}$ entails that $u \restriction T_c, u \restriction T_{c+1}, v \restriction T_c$ and $v \restriction T_{c+1}$ are all constant. Moreover, for all d, $u \restriction T_d$ is constant if and only if $v \restriction T_d$ is constant. In checking this, it helps to notice that the relevant subscripts and superscripts can all be determined from $F_{i\gamma\varepsilon}$ and the related Turing machine instruction $[i, \gamma, \delta, M, j]$. Also, if $I(f) = u \in \Omega$, then $u \restriction T_d$ is constant for all d.

Now we argue by induction on n, that every member of Ω_n is constant on T_c for all c that have become active by stage n and that, for all d and all $v, v' \in \Omega_n$, $v \restriction T_d$ is constant if and only if $v' \restriction T_d$ is constant.

The initial step of the induction holds vacuously.

For the inductive step, suppose $u, u' \in \Omega_{n+1} - \Omega_n$ with

$$F_{i\gamma\varepsilon}(f_c, f_{c+1}, u) = v \in \Omega_n \qquad \text{and} \qquad F_{i'\gamma'\varepsilon'}(f_{c'}, f_{c'+1}, u') = v' \in \Omega_n$$

Now our preliminary observations give the conclusions that u and u' are constant on all the d's active by stage n as well as for $c, c', c+1$, and $c'+1$, some of which may have become active for stage $n+1$. Moreover, we also conclude that, for all d, u is constant on T_d if and only if v is constant on T_d if and only if v' is constant of T_d if and only if u' is constant on T_d. In this way, the inductive step is complete. □

Now we just count things to obtain:

FACT 30. *Ω has no more than $2^s ms$ elements where $s = |\Sigma|$ and m is the number of nonhalting states of $\mathcal{T}$.*

PROOF: For each $u \in \Omega$ there are no more than s possibilities for $c \in \{a, \dots, b\}$ so that $u(t) = M_i^\gamma$, for some i and some γ and all $t \in T_c$. Having fixed one of these possibilities there are m choices for i and two choices for γ. Now for d with $a \leq d < c$ we must have a ν so that $u(t) = C_{i\gamma}^\nu$ for all $t \in T_d$. Thus for each such d there are no more than two possibilities for ν. Likewise, if $c < d \leq b$, then there is some ν so that $u(t) = D_{i\gamma}^\nu$ for all $t \in T_d$. Again, for each such d there are no more than two possibilities for ν. Thus, far we have bounded the number of possibilities for u by $2^s ms$, as desired—but we still have to examine what $u(t)$ is like when $t \in T_L \cup T_R$. Suppose $t \in T_L$. Then $f_c(t) = 1$ for all $c \in \{a, \dots, b\}$. From the definition of the operations $F_{i\gamma\varepsilon}$, it follows that $u(t) = C_{i\gamma}^\nu$, where ν is determined by $p(t) = C_{i'\gamma'}^\nu$, and i and γ are the same subscripts that occur throughout u. So u is determined on T_L by our previous choices and by the structure of p. Likewise, u is determined on T_R. So the desired bound is established. $\square$

THEOREM 5. *If $\mathcal{T}$ halts, then the cardinality of any subdirectly irreducible member of the variety generated by $\mathbf{A}(\mathcal{T})$ is no greater than the maximum of 2π, $2^{(\pi-1)}m(\pi-1)+\pi$ and $20m+28$, where π is the number of tape squares used by $\mathcal{T}$ in its halting computation and m is the number of nonhalting states of $\mathcal{T}$; moreover, every subdirectly irreducible algebra in the variety is flat.* $\square$

The $20m+28$ that occurs above is just the cardinality of $\mathbf{A}(\mathcal{T})$. It bounds the cardinalities of the subdirectly irreducibles that belong to $HS\mathbf{A}(\mathcal{T})$. The 2π bounds the cardinalities of the subdirectly irreducible algebras of sequentiable type. The $2^{(\pi-1)}m(\pi-1)+\pi$ bounds the cardinalities of the subdirectly irreducible algebras of machine type.

It is clear that much more was accomplished than just establishing the bound on subdirectly irreducible algebras given above. Our analysis is very close to a complete description (given a description of the behavior of $\mathcal{T}$) of all the subdirectly irreducible algebras, even in the case that $\mathcal{T}$ does not halt. The only way in which the hypothesis that $\mathcal{T}$ does not halt entered into consideration of the finite subdirectly irreducible algebras was in bounding their size. The analysis of their structure holds regardless. In the case that $\mathcal{T}$ does not halt, McKenzie describes how to carry this description of the finite subdirectly irreducible algebras up to the infinite subdirectly irreducibles, via an argument relying on Quackenbush's Theorem. His conclusion is that such varieties have residual character ω_1: while they have countably infinite subdirectly irreducible algebras, they have none of any larger cardinality.

Finally, we have in hand all the pieces of McKenzie's first undecidability result about finite algebras:

THEOREM 6. *The set of finite algebras of finite type which generate residually very finite varieties is not recursive. Indeed, that set is recursively inseparable from the set of finite algebras of finite type which generate varieties of residual character ω_1.* $\square$

LECTURE 10

$\mathbf{A}(\mathcal{T})$ is Finitely Based When $\mathcal{T}$ Halts

We have already seen in Lecture 6 a proof that if $\mathcal{T}$ does not halt, then $\mathbf{A}(\mathcal{T})$ is inherently nonfinitely based. In the preceding lecture, we found that when $\mathcal{T}$ halts, then $\mathbf{A}(\mathcal{T})$ has a finite residual bound, and indeed all the subdirectly irreducible algebras in the variety generated of $\mathbf{A}(\mathcal{T})$ are flat. Therefore, to complete the resolution of Tarski's Finite Basis Problem, it only remains to prove that finite algebras of this kind must be finitely based. We accomplish that in this lecture.

McKenzie's original solution of Tarski's Finite Basis Problem followed a different approach. That approach used a variant of the algebras $\mathbf{A}(\mathcal{T})$ and involved a detailed analysis of normal forms of terms, rather than an analysis of subdirectly irreducible algebras. Subsequent work by Ross Willard, relying on McKenzie's algebras $\mathbf{A}(\mathcal{T})$ and his analysis of subdirectly irreducible algebras, as well as considerable insight on Willard's part led to a second solution to Tarski's Finite Basis Problem. Still, Willard's solution, while more easily comprehended than McKenzie's, is by no means straightforward. It depends on a detailed understanding of how the many fundamental operations at hand interact.

Fortunately, there is now a direct route to our desired conclusion. The meet-semidistributive law reads

$$(\mathrm{SD}_\wedge) \qquad x \wedge y = x \wedge z \Rightarrow (x \wedge y) \vee (x \wedge z) = x \wedge (y \vee z).$$

We say that a variety $\mathcal{V}$ is **congruence $\wedge$-semidistributive** provided the congruence lattice of each algebra in $\mathcal{V}$ satisfies $\mathrm{SD}_\wedge$. Recently, Willard as established the following extension of Baker's Finite Basis Theorem.

THEOREM 7. *Every residually finite congruence $\wedge$-semidistributive variety with only finitely many basic operation symbols is finitely based.*

We will not offer a proof of this powerful theorem here. Rather, we will first show that it applies to our situation: that every algebra with a semilattice operation is congruence $\wedge$-semidistributive. Then we will provide a proof of a much weaker version of Theorem 7 that is adequate to demonstrate that $\mathbf{A}(\mathcal{T})$ is finitely based when $\mathcal{T}$ halts.

THEOREM 8. *If $\mathbf{A}$ is an algebra with a basic operation $\wedge$ so that $\langle A, \wedge\rangle$ is a semilattice, then* $\operatorname{Con}\mathbf{A}$ *is $\wedge$-semidistributive.*

PROOF. First notice that if $\theta \in \operatorname{Con}\mathbf{A}$ and $a \in A$, then $(a,b) \in \theta$ if and only if $(a \wedge b, a) \in \theta$ and $(a \wedge b, b) \in \theta$. This means that a congruence θ of $\mathbf{A}$ are entirely determined by the pairs $(a,b) \in \theta$ such that $a < b$.

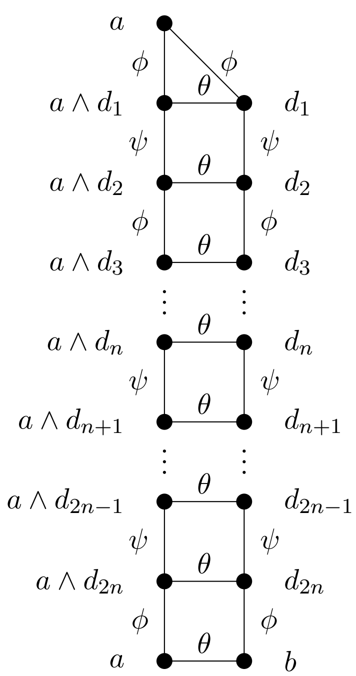
Now suppose θ, ϕ, and $\psi \in \operatorname{Con}\mathbf{A}$ with $\theta \cap \phi = \theta \cap \psi$, and that $a < b$ with $(a,b) \in \theta \cap (\phi \vee \psi)$. To demonstrate $\wedge$-semidistributivity it suffices to show that $(a,b) \in \phi$.

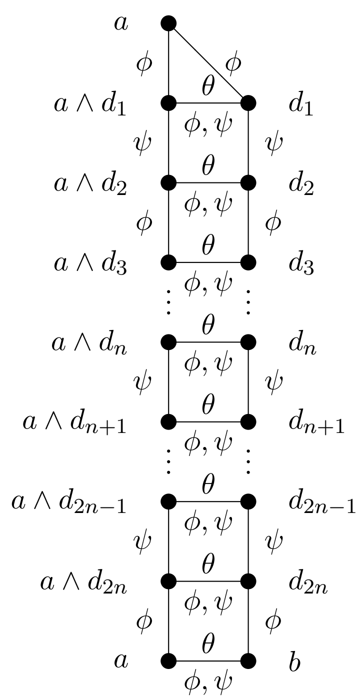
Because $(a,b) \in (\phi \vee \psi)$ for some $n > 0$ we can pick $c_1, c_2, \dots, c_n \in A$ so that

$$a \;\phi\; c_1 \;\psi\; c_2 \;\phi \dots c_{n-1} \;\psi\; c_n \;\phi\; b$$

Without loss of generality, we assume that $c_i \leq b$ for each i. (Otherwise, we can replace each c_i by $c_i \wedge b$.) To make the diagrams below more legible, for $i \leq n$, let $d_i = c_1 \wedge c_2 \wedge \ldots \wedge c_i$, and for $n < i \leq 2n$, let $d_i = c_{i-n} \wedge c_{i+1-n} \wedge \ldots \wedge c_n$. Hence, $d_i \leq b$ for all i. In consequence, $a \wedge d_i \; \theta \; d_i$ for all i. Likewise, either $d_i \; \phi \; d_{i+1}$ or $d_i \; \psi \; d_{i+1}$ depending on the parity of i.

Now just observe the following diagrams:





The diagram on the left displays the congruence relations between these elements as described above. The diagram on the right shows additional relations along the horizontal lines. These are deduced as follows. By transitivity, we see that $(a \wedge d_1) \; \phi \; d_1$, which allows us to label the top horizontal line with ϕ. Hence, $(a \wedge d_1) \; \theta \cap \phi \; d_1$. But $\theta \cap \phi = \theta \cap \psi$, by our hyposthesis. This permits us to label the top horizontal line with ψ as well. But then transitivity allows us to label the next horizontal line with ψ, and so we can continue until we reach the bottom line, which includes the conclusion we desire. $\square$

THEOREM 9. *If $\mathcal{T}$ halts, the* $\mathbf{A}(\mathcal{T})$ *is finitely based.*

THEOREM 10. *There is no algorithm which determines whether a finite algebra is finitely based.*

As it stands, our line of reasoning leading to the last theorem is incomplete, since no proof of Theorem 7 has been provided. To remedy this situation we offer the following weaker version, also due to Ross Willard.

THEOREM 11. *Let* $\mathbf{A}$ *be a finite algebra with only finite many basic operations, among which is a binary operation* $\wedge$ *so that* $\langle A, \wedge \rangle$ *is a semilattice. If the variety generated by* $\mathbf{A}$ *has a finite residual bound, then* $\mathbf{A}$ *is finitely based.*

PROOF. Let $\mathcal{V}$ be the variety generated by $\mathbf{A}$.

This proof depends on finding three schemas of first-order sentences $\Delta_{k,N}$, Φ_N, and $\Theta_{m,N}$ (where k, n, m, and N are natural numbers) which, roughly speaking, assert bounds on the height and the cardinality of subdirectly irreducible algebras. Before describing these sentences we formulate two key lemmas which they fulfill and describe how the lemmas yield a proof of the theorem.

We say an algebra $\mathbf{B}$ has the N**-approximation property** provided for all $a, b, c, d \in B$ with $d < c$ and all translations λ of $\mathbf{B}$, if $\lambda(a) \wedge c \not\leq d$ and $\lambda(b) \wedge c \leq d$ there is a translation λ' of $\mathbf{B}$ with complexity bounded by N so that $\lambda'(a) \wedge c \not\leq d$ and $\lambda'(b) \wedge c \leq d$. A class $\mathcal{K}$ of algebras has the N-approximation property provided every algebra in $\mathcal{K}$ has the property. The algebra $\mathbf{B}$ has the **strong** N**-approximation property** provided for all $a, b \in B$ and all translations λ of $\mathbf{B}$ there is a translation λ' of complexity bounded by N so that $\lambda(a) = \lambda'(a)$ and $\lambda(b) = \lambda'(b)$. The class $\mathcal{K}$ has the strong N-approximation property provided each algebra in $\mathcal{K}$ has the property.

For any class $\mathcal{K}$ of algebras, we use $P_2\mathcal{K}$ to denote the class of all algebras isomorphic to direct products with no more than 2 factors chosen from $\mathcal{K}$.

LEMMA 2. *Let $\mathcal{K}$ be a class of algebras. If $SP_2\mathcal{K}$ has the strong N-approximation property and $SP\mathcal{K}$ is locally finite, then $SP\mathcal{K} \models \Phi_{N+1}$. Conversely, if $\mathcal{K} \models \Phi_N$, then $\mathcal{K}$ has the N-approximation property.*

LEMMA 3. *Suppose that $\mathcal{W}$ is a variety such that $\wedge$ denotes a semilattice operation in $\mathcal{W}$, and $\mathcal{W} \models \Phi_N$.*

1. *If $\mathcal{W}$ has residual bound m, then $\mathcal{W} \models \Theta_{m,N}$.*
2. *If $\mathcal{W} \models \Theta_{m,N}$ and $\mathcal{W}$ is residually of finite height, then $\mathcal{W}$ is residually bounded by m.*
3. *$\mathcal{W} \models \Delta_{k,N}$ if and only if $\mathcal{W}$ has height residually bounded by k.*

Here is how to prove the theorem based on these lemmas.

Let m be the residual bound of $\mathcal{V}$. Let k bound the residual height of $\mathcal{V}$. Let $\mathcal{K}$ be the class of subdirectly irreducible algebras belonging to $\mathcal{V}$.

Now, up to isomorphism, $SP_2\mathcal{K}$ is a finite set of finite algebras, m^2 being an upper bound on their cardinalities. This means that no algebra in $SP_2\mathcal{K}$ can have more than $(m^2)^{(m^2)}$ translations. Consequently, let N bound the complexity of all the translations for all the algebras in $SP_2\mathcal{K}$. Observe that $SP_2\mathcal{K}$ has the strong N-approximation property.

Then according to Lemma 2, we know that $\mathcal{V} = SP\mathcal{K} \models \Phi_{N+1}$. In addition, according to Lemma 3, part (1), we find that $\mathcal{V} \models \Theta_{m,N+1} \& \Delta_{k,N+1}$. By the Compactness Theorem, let Σ be a finite set of equations, each true in $\mathcal{V}$, so that

$$\Sigma \models \Phi_{N+1} \& \Theta_{m,N+1} \& \Delta_{k,N+1}.$$

Now let $\mathcal{W} = \operatorname{Mod}\Sigma$. Notice that $\mathcal{V} \subseteq \mathcal{W}$. Then according the Lemma 3, part (3), we have that $\mathcal{W}$ has height residually bounded by k. Next, according to Lemma 3, part (2), we find that $\mathcal{W}$ is residually bounded by m. This means that $\mathcal{W}$ is a variety that has, up to isomorphism, only finitely many subdirectly irreducible algebras. In particular, there are, up to isomorphism, only finitely many subdirectly irreducible algebras in $\mathcal{W} - \mathcal{V}$. Let $\mathbf{S}_0, \mathbf{S}_1, \ldots, \mathbf{S}_{n-1}$ be a list of respresentatives of these subdirectly irreducibles. For each $i < n$, pick an equation $s_i \approx t_i$ which is true in $\mathcal{V}$ but fails in $\mathbf{S}_i$. Let $\Gamma = \Sigma \cup \{s_i \approx t_i : i < n\}$. Then Γ is the desired finite basis for $\mathcal{V}$, since the subdirectly irreducible models of Γ are precisely the subdirectly irreducible algebras belonging to $\mathcal{V}$. $\square$

Two Lemmata for Three Schemata

Here is the first bunch of those elementary sentences:

$$(\Phi_N)\quad \begin{array}{l} \text{For all } x, y, z, \text{ and } w \text{ with } w < z, \\ \text{if } \lambda(y) \wedge z \leq w \text{ and } \lambda(x) \wedge z \nleq w \text{ for some translation } \lambda \text{ of complexity } N+1, \\ \text{then } \lambda'(y) \wedge z \leq w \text{ and } \lambda'(x) \wedge z \nleq w \text{ for some translation } \lambda' \text{ of complexity } N. \end{array}$$

LEMMA 2. *Let $\mathcal{K}$ be a class of algebras. If $SP_2\mathcal{K}$ has the strong N-approximation property and $SP\mathcal{K}$ is locally finite, then $SP\mathcal{K} \models \Phi_{N+1}$. Conversely, if $\mathcal{K} \models \Phi_N$, then $\mathcal{K}$ has the N-approximation property.*

PROOF. Here we suppose that $SP\mathcal{K}$ is locally finite and $SP_2\mathcal{K}$ has the strong N-approximation property. We wish to establish that $SP\mathcal{K} \models \Phi_{N+1}$. So suppose $\mathbf{B} \in SP\mathcal{K}$, that λ is a translation of $\mathbf{B}$ of complexity less than or equal to $N+2$, and that $a, b, c, d \in B$ with $d < c$ such that

$$\lambda(b) \wedge c \leq d \text{ and } \lambda(a) \wedge c \nleq d.$$

Since $SP\mathcal{K}$ is locally finite, we know that the subalgebra $\mathbf{B}'$ of $\mathbf{B}$ generated by a, b, c, d, and all the elements of B that are involved in λ as coefficients is finite. It follows that $\mathbf{B}'$ is (isomorphic to) a subalgebra of a direct product $\mathbf{B}_0 \times \mathbf{B}_1 \times \cdots \times \mathbf{B}_{n-1}$ of finitely many algebras from $\mathcal{K}$. Without loss of generality we suppose that this direct product has been indexed so that $(\lambda(a))_0 \wedge c_0 \nleq d_0$.

Now for each $i < n$, let $\mathbf{C}_i$ be the projection of $\mathbf{B}'$ into $\mathbf{B}_0 \times \mathbf{B}_i$. (In particular, $\mathbf{C}_0$ is the diagonal subalgebra of $\mathbf{B}_0 \times \mathbf{B}_0$.) We use $\lambda_{0,i}$ to denote the translation of $\mathbf{C}_i$ obtained by applying the projection of $\mathbf{B}'$ onto $\mathbf{C}_i$ to λ. In this way we know

$$\lambda_{0,i}(b_{0,i}) \wedge c_{0,i} \leq d_{0,i} \text{ for all } i < n, \text{ and } \lambda_{0,0}(a_{0,0}) \wedge c_{0,0} \nleq d_{0,0}.$$

Because $SP_2\mathcal{K}$ has the strong N-approximation property, for each $i < n$, we pick a translation λ'_i of $\mathbf{C}_i$ of complexity bounded by N, so that $\lambda'_i(a_{0,i}) = \lambda_{0,i}(a_{0,i})$ and $\lambda'_i(b_{0,i}) = \lambda_{0,i}(b_{0,i})$ for all $i < n$.

For each $i < n$, let μ_i be a translation of $\mathbf{B}'$ obtained by pulling λ'_i back through the projection. Observe that μ_i has complexity bounded by N, for every $i < n$.

Now define

$$\begin{aligned} e_0 &= \mu_0(a) \wedge \mu_1(a) \wedge \ldots \wedge \mu_{n-1}(a) \\ e_1 &= \mu_0(b) \wedge \mu_1(a) \wedge \ldots \wedge \mu_{n-1}(a) \\ e_2 &= \mu_0(b) \wedge \mu_1(b) \wedge \ldots \wedge \mu_{n-1}(a) \\ &\vdots \\ e_n &= \mu_0(b) \wedge \mu_1(b) \wedge \ldots \wedge \mu_{n-1}(b) \end{aligned}$$

Observe that $e_0 \wedge c \not\leq d$, as can be seen by examining the projection of each side of this inequality to $\mathbf{B}_0$. In detail, the projection of $\mu_j(a) \wedge c$ onto $\mathbf{C}_j$ yields $\lambda'(a_{0,j}) \wedge c_{0,j} = \lambda_{0,j}(a_{0,j}) \wedge c_{0,j}$. Projecting this in turn on $\mathbf{B}_0$ gives $(\lambda(a))_0 \wedge c_0$. So the projection of $\mu_j(a) \wedge c$ to $\mathbf{B}_0$ produces $(\lambda(a))_0 \wedge c_0$, a value independent of j. Therefore the projection of projection of $e_0 \wedge c$ to $\mathbf{B}_0$ is $(\lambda(a))_0 \wedge c_0 \not\leq d_0$. Hence $e_0 \wedge c \not\leq d$.

On the other hand $e_n \wedge c \leq d$, since, for each $i < n$, the projection of this inequality to $\mathbf{C}_i$ holds by construction. Let $j < n$ be as small as possible so that $e_j \wedge c \not\leq d$ but $e_{j+1} \wedge c \leq d$. Let ν be the translation of $\mathbf{B}'$ defined by

$$\nu(x) = \mu_0(b) \wedge \ldots \wedge \mu_j(b) \wedge \mu_j(x) \wedge \mu_{j+2}(a) \wedge \ldots \wedge \mu_{n-1}(a).$$

Then ν is a translation with complexity bounded by $N+1$ and $\nu(a) \wedge c \not\leq d$ while $\nu(b) \wedge c \leq d$. Since ν can also be regarded as a translation of $\mathbf{B}$, we have the conclusion that $SP\mathcal{K} \models \Phi_{N+1}$. This completes the first contention of the lemma.

Now suppose that $\mathcal{K} \models \Phi_N$. Our objective is to prove that $\mathcal{K}$ has the N-approximation property. Let $\mathbf{B} \in \mathcal{K}$. We need to show that for every $n > N$ and every $a, b, c, d \in B$ with $d < c$, if there is a translation λ of complexity bounded by n such that

$$\lambda(a) \wedge c \not\leq d \text{ and } \lambda(b) \wedge c \leq d,$$

then there is a translation λ' of complexity bounded by N such that

$$\lambda'(a) \wedge c \not\leq d \text{ and } \lambda'(b) \wedge c \leq d,$$

We do this by induction on n. The initial step is $n = N+1$, and this is ensured by Φ_N. For the inductive step, we take $n = n'+1$. So there must be a basic translation β and a translation μ of complexity bounded by n' so that $\lambda(x) = \mu(\beta(x))$. Let $a' = \beta(a)$ and $b' = \beta(b)$. According to the induction hypothesis (applied to a', b', c, d and with μ in place of λ), there is a translation μ' of complexity bounded by N so that

$$\mu'(\beta(a)) \wedge c \not\leq d \text{ and } \mu'(\beta(b)) \wedge c \leq d.$$

Now $\mu'(\beta(x))$ is a translation of complexity bounded by $N+1$. Invoking Φ_N, we obtain the desired transation λ' of complexity bounded by N. This completes the induction, and the proof of the lemma.

□

Here are the other two bunches of sentences.

($\Theta_{m,N}$) There do not exist $x, y, u_0, u_1, \ldots, u_m$ so that $y < x$ and for each i, j with $i < j \leq m$ there exists $z \leq y$ and a translation λ of complexity less than or equal to $N+1$ such that $\{x, z\} = \{\lambda(u_i), \lambda(u_j)\}$.

and

For all translations $\lambda_0, \dots, \lambda_k$ of complexity less than or equal to N, and all $x_0, x_1, \dots, x_{k+1}$ with $x_0 \leq x_1 \leq \dots \leq x_{k+1}$, and for all choices $\{u_i, v_i\} = \{x_i, x_{i+1}\}$ and for all z, w,
if all of the following inequalities hold

$$(\Delta_{k,N}) \qquad \begin{aligned} z \wedge \lambda_0(u_0) &\leq w \\ z \wedge \lambda_0(v_0) \wedge \lambda_1(u_1) &\leq w \\ &\vdots \\ z \wedge \lambda_0(v_0) \wedge \ldots \wedge \lambda_{k-1}(v_{k-1}) \wedge \lambda_k(u_k) &\leq w, \end{aligned}$$

then

$$z \wedge \lambda_0(v_0) \wedge \ldots \wedge \lambda_{k-1}(v_{k-1}) \wedge \lambda_k(v_k) \leq w.$$

LEMMA 3. *Suppose that $\mathcal{W}$ is a variety such that $\wedge$ denotes a semilattice operation in $\mathcal{W}$, and $\mathcal{W} \models \Phi_N$.*

1. *If $\mathcal{W}$ has residual bound m, then $\mathcal{W} \models \Theta_{m,N}$.*
2. *If $\mathcal{W} \models \Theta_{m,N}$ and $\mathcal{W}$ is residually of finite height, then $\mathcal{W}$ is residually bounded by m.*
3. *$\mathcal{W} \models \Delta_{k,N}$ if and only if $\mathcal{W}$ has height residually bounded by k.*

PROOF. For this proof we take $\mathcal{W}$ to be a variety for which $\wedge$ denotes a semilatice operation. We further suppose that $\mathcal{W} \models \Phi_N$. The lemma has three parts.

PART 1. Here we also suppose $\mathcal{W}$ has residual bound m. We want to establish that $\mathcal{W} \models \Theta_{m,N}$. But an examination of $\Theta_{m,N}$, reveals that if $\mathbf{B}$ is an algebra in which $\Theta_{m,N}$ *fails*, then there are $a, b, c_0, \dots, c_{m-1} \in B$ with $b < a$ so that every congruence of $\mathbf{B}$ which separates a and b, must also separate all the c_i's from each other. Let $\theta \in \operatorname{Con} \mathbf{B}$ be a maximal congruence separating a and b. Then $\mathbf{B}/\theta$ is subdirectly irreducible and has at least $m+1$ members. Hence $\mathbf{B} \notin \mathcal{W}$. This means $\mathcal{W} \models \Theta_{m,N}$.

PART 2. In this part we also suppose that $\mathcal{W} \models \Theta_{m,N}$ and that $\mathcal{W}$ is residually of finite height. Our object is to show that m is a residual bound for $\mathcal{W}$. So let $\mathbf{B} \in \mathcal{W}$ be subdirectly irreducible. Then we can find $p, q \in B$ with $q < p$ so that (q, p) is a critical pair. Let p^* be an element of B minimal among those $r \in B$ such that $r \leq p$ while $r \nleq q$. The existence of such a minimal element is a consequence of the fact that $\mathbf{B}$ has finite height. Put $q^* = p^* \wedge q$. It follows that (q^*, p^*) is also a critical pair for $\mathbf{B}$. Moreover, if $x \wedge p^* \nleq q^*$, then $x \wedge p^* = p^*$ for all $x \in B$.

Now suppose $a, b \in B$ with $a \neq b$. Then $(q^*, p^*) \in \operatorname{Cg}^B(a, b)$. It follows from Mal'cev's Congruence Generation Theorem that there is a translation λ of $\mathbf{B}$ such that, without loss of generality, $\lambda(a) \wedge p^* \leq q^*$ but $\lambda(b) \wedge p^* \nleq q^*$. In view of Lemma 2, we can suppose that the complexity of λ is bounded by N. According to the last sentence of the preceding paragraph, we see that $\lambda(b) \wedge p^* = p^*$. Let $\mu(x) = \lambda(x) \wedge p*$. So μ is a translation whose complexity is bounded by $N+1$. So for any pair $a, b \in B$ of distinct elements, there is a translation μ of complexity bounded by $N+1$ such that $\{\mu(a), \mu(b)\} = \{p^*, r\}$, where $r \leq q^*$. Since $\mathbf{B} \models \Theta_{m,N}$, it follows that B can have no more than m elements. Consequently, m is a residual bound for $\mathcal{W}$.

PART 3. First, let us suppose that k residually bounds the height in $\mathcal{W}$. Let $\mathbf{B} \in \mathcal{W}$. Our objective is to prove that $\models\Delta_{k,N}$. Now an examination of $\Delta_{k,n}$ reveals that it is a conjunction of finitely many sentences, each of which is a universally quantified implication for which the hypothesis is a conjunction of equations and the conclusion in an equation. The truth of such sentences with this syntactic structure (known as conjunctions of quasi-equations) is readily seen to be preserved by both the formation of direct products and of subalgebras. Consequently, it is only necessary to verify that $\Delta_{k,N}$ holds when $\mathbf{B}$ is subdirectly irreducible. But then $\mathbf{B}$ has height bounded by k. Examination of $\Delta_{k,N}$, shows that for any $x_0, \dots, x_{k+1}$ chosen from B to fulfill the hypothesis, we must have $x_i = x_{i+1}$ for some i. But this forces $u_i = v_i$. In consequence, the left side of the conclusion of $\Delta_{k,N}$ is no larger than the left side of the i^{th} hypothesis. Hence, $\mathbf{B} \models \Delta_{k,N}$ as desired.

For the converse, we will suppose that $\mathbf{B} \in \mathcal{W}$ is a subdirectly irreducible algebra of height larger than k, and construct a failure of $\Delta_{k,N}$. By the height of $\mathbf{B}$ pick $a_0 < a_1 < \cdots < a_{k+1}$ in B. Since $\mathbf{B}$ is subdirectly irreducible, pick $q < p$ in B so that (q, p) is a critical pair for $\mathbf{B}$. To construct the failure we will define translations λ_i (of complexity bounded by N) and elements c_i and d_i. We do this recursively as follows. Let $i \leq k$ and suppose that the earlier λ's, c's and d's are in hand, so that for $j < i$

$$\begin{aligned} \{c_j, d_j\} &= \{a_j, a_{j+1}\} \\ \lambda_j(c_j) \wedge p_j &\leq q \\ p_i &\not\leq q \end{aligned}$$

where for convenience we have taken

$$p_j := p \wedge \lambda_0(d_0) \wedge \ldots \wedge \lambda_{j-1}(d_{j-1}).$$

Now let $q_i = p_i \wedge q$. Hence $q_i < p_i$ and it is easy to verify that (q_i, p_i) is also a critical pair. Hence, $(q_i, p_i) \in \mathrm{Cg}^{\mathbf{B}}(a_i, a_{i+1})$. By Mal'cev's Congruence Generation Theorem and Lemma 2, pick a translation λ_i of complexity bounded by N, and $c_i, d_i \in B$ so that $\{c_i, d_i\} = \{a_i, a_{i+1}\}$ and

$$\lambda_i(c_i) \wedge p_i \leq q_i \text{ and } \lambda_i(d_i) \wedge p_i \not\leq q_i.$$

Finally, notice $p_{i+1} = p_i \wedge \lambda_i(d_i)$. Because, $q_i \leq q$ we have

$$\begin{aligned} \{c_i, d_i\} &= \{a_i, a_{i+1}\} \\ \lambda_i(c_i) \wedge p_i &\leq q \\ p_{i+1} &\not\leq q \end{aligned}$$

This completes the recursive construction of a failure of $\Delta_{k,N}$ in $\mathbf{B}$.

Thus all parts of the lemma are established.

□

THAT'S ALL FOLKS!