Extensions of Abelian Automaton Groups

Chris Grossack
(Advisor: Klaus Sutner)

May 8, 2019

Let's Unpack That

Extensions of Abelian Automaton Groups

Let's Unpack That

Extensions of Abelian Automaton Groups

Let's Unpack That

Extensions of Abelian Automaton Groups

Let's Unpack That

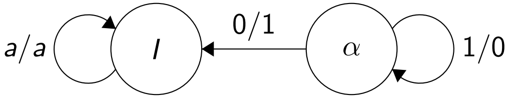
Extensions of Abelian Automaton Groups

Let's Unpack That

Extensions of Abelian Automaton Groups

Finite State Automata

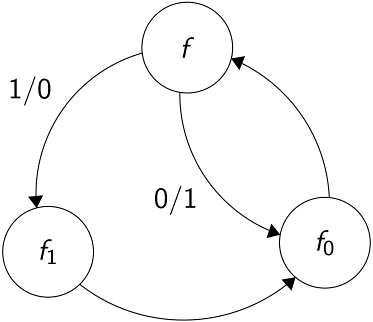
- Combinatorial Objects
- Encode length preserving functions on binary strings
  - states
  - transitions



- One function per state
- Evaluate by following edges

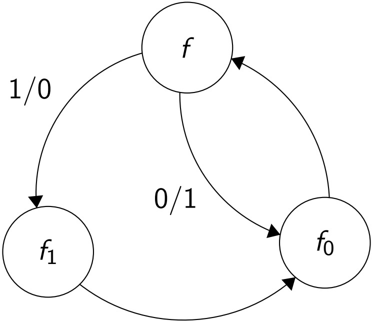
Evaluating Functions

- $\mathcal{A}_2^3$. This automaton will be our friend for the rest of this talk
- Defines three functions:
  - f
  - f_0
  - f_1



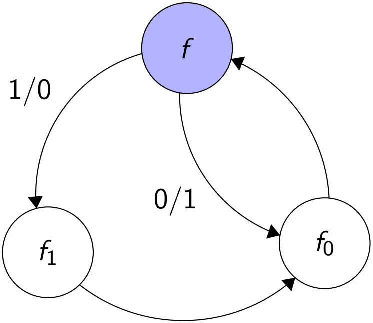
Evaluating Functions

- $\mathcal{A}_2^3$. This automaton will be our friend for the rest of this talk
- Defines three functions:
  - f
  - f_0
  - f_1



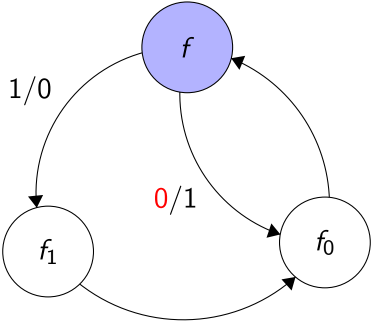
- How do we compute, say, f of a string?

Evaluating Functions



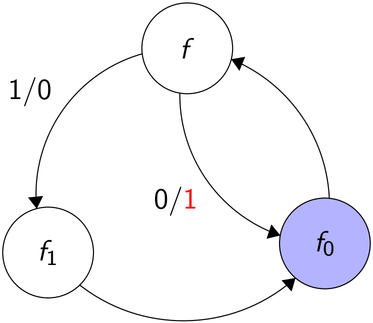
- $f(011010)$

Evaluating Functions



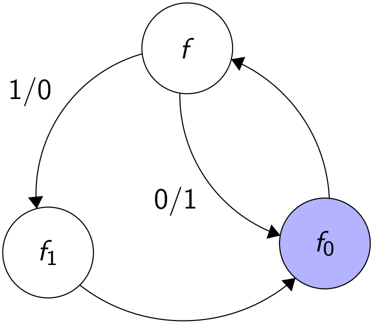
- $f(011010)$

Evaluating Functions



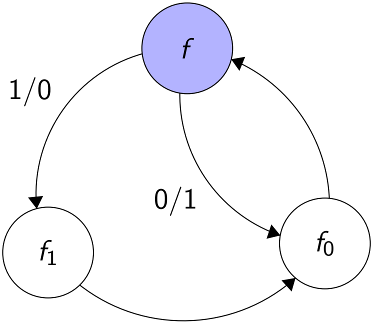
- $1f_0(11010)$

Evaluating Functions



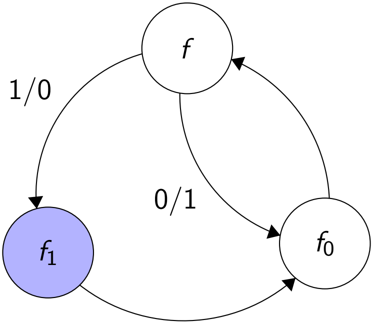
- $1f_0(11010)$

Evaluating Functions



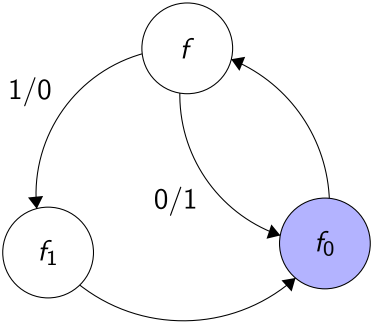
- $11f(1010)$

Evaluating Functions



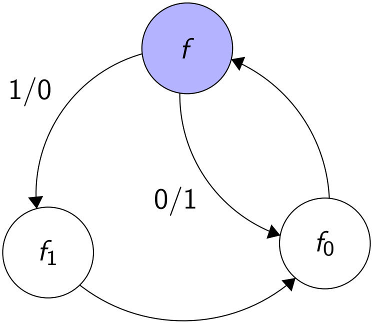
- $110 f_1(010)$

Evaluating Functions



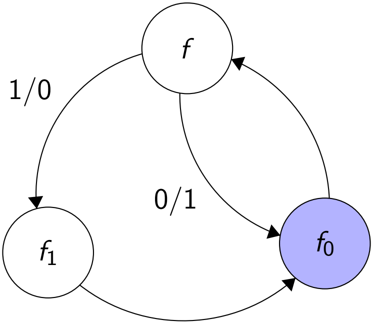
- $1100 f_0(10)$

Evaluating Functions



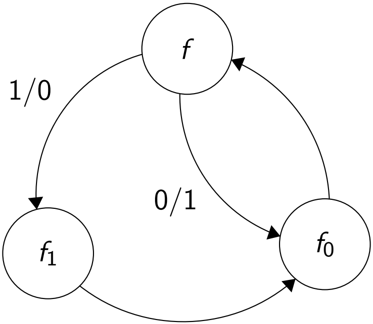
- $11001f(0)$

Evaluating Functions



- $110011f_0(\varepsilon)$

Evaluating Functions



- 110011

Definition

$\partial_0 f$ (resp. $\partial_1 f$) is the unique function such that for every $w \in 2^*$,
$f(0w) = (f(0))(\partial_0 f)(w)$

Definition

$\partial_0 f$ (resp. $\partial_1 f$) is the unique function such that for every $w \in 2^*$, $f(0w) = (f(0))(\partial_0 f)(w)$

Definition

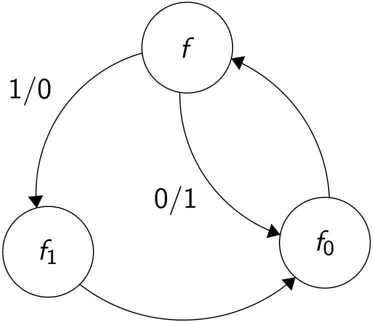
A state f is called Odd if it flips its input bit, and Even otherwise.

Definition

$\partial_0 f$ (resp. $\partial_1 f$) is the unique function such that for every $w \in 2^*$,
$f(0w) = (f(0))(\partial_0 f)(w)$

Definition

A state f is called Odd if it flips its input bit, and Even otherwise.



- $\partial_0 f = f_0$ and $\partial_1 f = f_1$
- f is odd, f_0 and f_1 are even

Definition

For f and g in an automaton $\mathcal{A}$, write $f + g$ for the function $(f + g)(x) = f(g(x))$

Definition

For f and g in an automaton $\mathcal{A}$, write $f + g$ for the function $(f + g)(x) = f(g(x))$

- We are interested in the Abelian case.

Definition

For f and g in an automaton $\mathcal{A}$, write $f + g$ for the function $(f + g)(x) = f(g(x))$

- We are interested in the Abelian case.
- For all of our machines, $f + g = g + f$

Definition

For f and g in an automaton $\mathcal{A}$, write $f+g$ for the function $(f+g)(x) = f(g(x))$

- We are interested in the Abelian case.
- For all of our machines, $f+g=g+f$
- Given a machine $\mathcal{A}$, this condition is checkable in polynomial time

Definition

Recall a Group is a set $\mathcal{G}$ *equipped with*

- $0 \in \mathcal{G}$
- $+ : \mathcal{G} \to \mathcal{G} \to \mathcal{G}$ *(associative)*
- $- : \mathcal{G} \to \mathcal{G}$
- *satisfying* $0 + x = x + 0 = x$ *and* $x + (-x) = (-x) + x = 0$

Definition

Recall a Group is a set $\mathcal{G}$ equipped with

- $0 \in \mathcal{G}$
- $+ : \mathcal{G} \to \mathcal{G} \to \mathcal{G}$ *(associative)*
- $- : \mathcal{G} \to \mathcal{G}$
- *satisfying* $0 + x = x + 0 = x$ *and* $x + (-x) = (-x) + x = 0$

- If S is the state set of an automaton $\mathcal{A}$, consider $\mathcal{G}$ to be the closure of S under $+$
- take $0 = id : 2^* \to 2^*$ to be the empty sum

Definition

Recall a Group is a set $\mathcal{G}$ equipped with

- $0 \in \mathcal{G}$
- $+ : \mathcal{G} \to \mathcal{G} \to \mathcal{G}$ *(associative)*
- $- : \mathcal{G} \to \mathcal{G}$
- *satisfying* $0 + x = x + 0 = x$ *and* $x + (-x) = (-x) + x = 0$

- If S is the state set of an automaton $\mathcal{A}$, consider $\mathcal{G}$ to be the closure of S under $+$
- take $0 = id : 2^* \to 2^*$ to be the empty sum
- This does not have $-$ in general
- Our functions don't even need to be *invertible*

Definition

Recall a Group is a set $\mathcal{G}$ equipped with

- $0 \in \mathcal{G}$
- $+ : \mathcal{G} \to \mathcal{G} \to \mathcal{G}$ *(associative)*
- $- : \mathcal{G} \to \mathcal{G}$
- *satisfying* $0 + x = x + 0 = x$ *and* $x + (-x) = (-x) + x = 0$

- If S is the state set of an automaton $\mathcal{A}$, consider $\mathcal{G}$ to be the closure of S under $+$
- take $0 = id : 2^* \to 2^*$ to be the empty sum
- This does not have $-$ in general
- Our functions don't even need to be *invertible*
- Each function is invertible iff each state is invertible in one step

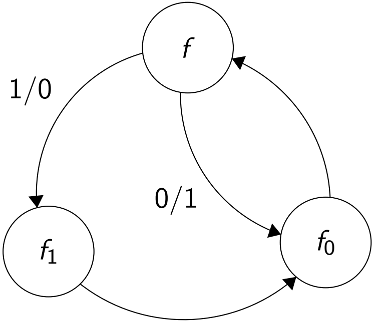
The Inverse Automaton

- Since each state sees an invertible function. . . invert it.

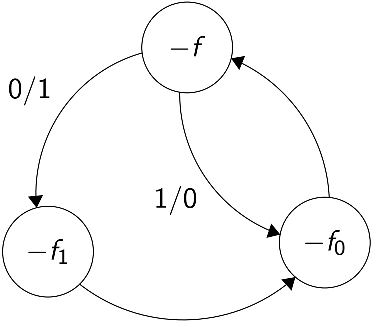
The Inverse Automaton

- Since each state sees an invertible function... invert it.

$\mathcal{A}_2^3$



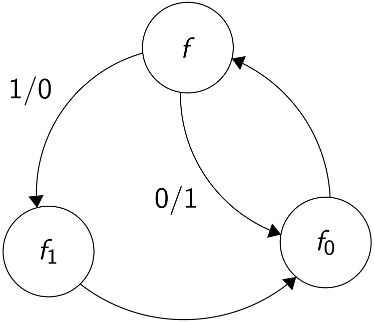
$-\mathcal{A}_2^3$



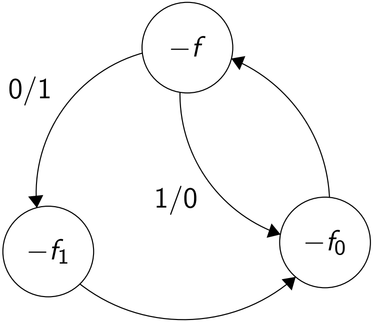
The Inverse Automaton

- Since each state sees an invertible function... invert it.

$\mathcal{A}_2^3$



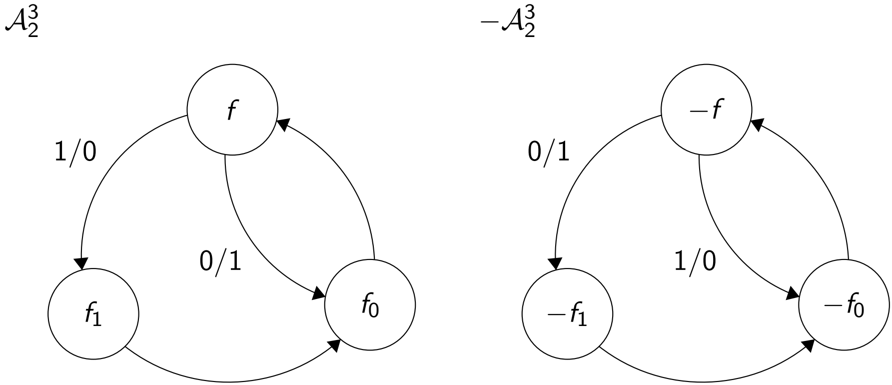
$-\mathcal{A}_2^3$



- Take Care: $\partial_0(-f) = -\partial_1 f$

The Inverse Automaton

- Since each state sees an invertible function... invert it.

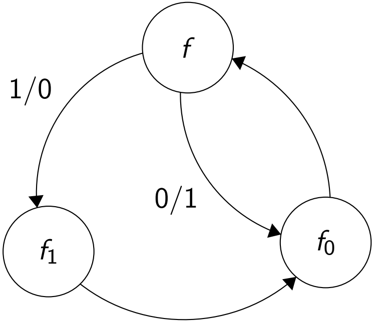


- Take Care: $\partial_0(-f) = -\partial_1 f$
- But: $(f + (-f))(01) = f((-f)(01)) = f(11) = 01$

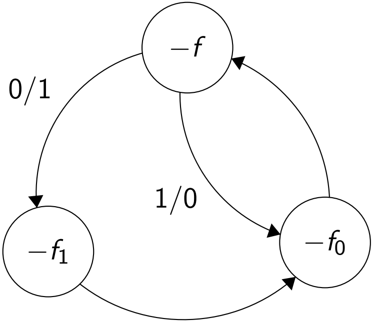
The Inverse Automaton

- Since each state sees an invertible function. . . invert it.

$\mathcal{A}_2^3$



$-\mathcal{A}_2^3$



- Take Care: $\partial_0(-f) = -\partial_1 f$
- But: $(f + (-f))(01) = f((-f)(01)) = f(11) = 01$
- An easy induction shows these are actually inverses.

- So $\mathcal{G}(\mathcal{A})$ is a group whenever each state sees an invertible function

- So $\mathcal{G}(\mathcal{A})$ is a group whenever each state sees an invertible function
- If $\mathcal{A}$ is abelian, so is $\mathcal{G}(\mathcal{A})$.

- So $\mathcal{G}(\mathcal{A})$ is a group whenever each state sees an invertible function
- If $\mathcal{A}$ is abelian, so is $\mathcal{G}(\mathcal{A})$.
- What groups can we get?

- So $\mathcal{G}(\mathcal{A})$ is a group whenever each state sees an invertible function
- If $\mathcal{A}$ is abelian, so is $\mathcal{G}(\mathcal{A})$.
- What groups can we get?

Theorem (Nekrashevych and Sidki)

The only abelian automaton groups are $(\mathbb{Z}/2\mathbb{Z})^m$ *and* $\mathbb{Z}^m$.

- So $\mathcal{G}(\mathcal{A})$ is a group whenever each state sees an invertible function
- If $\mathcal{A}$ is abelian, so is $\mathcal{G}(\mathcal{A})$.
- What groups can we get?

Theorem (Nekrashevych and Sidki)

The only abelian automaton groups are $(\mathbb{Z}/2\mathbb{Z})^m$ *and* $\mathbb{Z}^m$.

- Given an abelian automaton $\mathcal{A}$, one can check in polynomial time which group it generates.

- So $\mathcal{G}(\mathcal{A})$ is a group whenever each state sees an invertible function
- If $\mathcal{A}$ is abelian, so is $\mathcal{G}(\mathcal{A})$.
- What groups can we get?

Theorem (Nekrashevych and Sidki)

The only abelian automaton groups are $(\mathbb{Z}/2\mathbb{Z})^m$ and $\mathbb{Z}^m$.

- Given an abelian automaton $\mathcal{A}$, one can check in polynomial time which group it generates.
- We will focus on the $\mathbb{Z}^m$ case here

Theorem

$\mathbb{Z}^m$ equipped with a matrix $\mathbf{A}$ and $\bar{e}$ forms a (infinite state) abelian automaton (called $\mathfrak{C}(\mathbf{A}, \bar{e})$) with residuation as shown below. Further, for every abelian automaton $\mathcal{A}$ whose group is $\mathbb{Z}^m$, there exists an $\mathbf{A}$ and $\bar{e}$ such that $\mathcal{A}$ is a finite subautomaton. The odd states are exactly the states with odd first component.

$$\mathbf{A} = \begin{pmatrix} \frac{a_1}{2} & 1 & 0 & \cdots & 0 \\ \frac{a_2}{2} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{a_{m-1}}{2} & 0 & 0 & \cdots & 1 \\ \frac{a_m}{2} & 0 & 0 & \cdots & 0 \end{pmatrix}$$

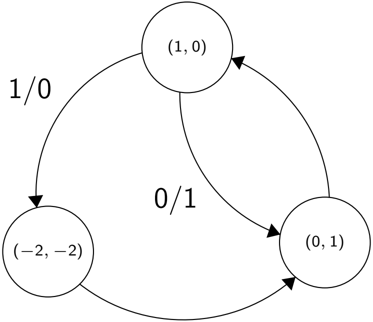
$$\partial_0 \bar{v} = \begin{cases} A(\bar{v}) & \bar{v} \text{ is even} \\ A(\bar{v} - \bar{e}) & \bar{v} \text{ is odd} \end{cases}$$

$$\partial_1 \bar{v} = \begin{cases} A(\bar{v}) & \bar{v} \text{ is even} \\ A(\bar{v} + \bar{e}) & \bar{v} \text{ is odd} \end{cases}$$

- $a_i \in \mathbb{Z}$
- $\mathbf{A}$ has irreducible characteristic polynomial
- $\bar{e}$ (the Residuation Vector) is odd

Example:

Take $\mathbf{A} = \begin{pmatrix} -1 & 1 \\ -\frac{1}{2} & 0 \end{pmatrix}$, and $\bar{e} = (3, 2)$. Then:



- It is natural to ask for what matrices $\mathbf{A}$ and vectors $\bar{e}$ can we find a given $\mathcal{A}$ in $\mathfrak{C}(\mathbf{A}, \bar{e})$, and at what vectors $\bar{v}$ are its states?

- It is natural to ask for what matrices $\mathbf{A}$ and vectors $\bar{e}$ can we find a given $\mathcal{A}$ in $\mathfrak{C}(\mathbf{A}, \bar{e})$, and at what vectors $\bar{v}$ are its states?
- It can be shown that only one $\mathbf{A}$ works

- It is natural to ask for what matrices $\mathbf{A}$ and vectors $\bar{e}$ can we find a given $\mathcal{A}$ in $\mathfrak{C}(\mathbf{A}, \bar{e})$, and at what vectors $\bar{v}$ are its states?
- It can be shown that only one $\mathbf{A}$ works
- Becker even found a way of computing $\mathbf{A}$ given the automaton

- It is natural to ask for what matrices $\mathbf{A}$ and vectors $\bar{e}$ can we find a given $\mathcal{A}$ in $\mathfrak{C}(\mathbf{A}, \bar{e})$, and at what vectors $\bar{v}$ are its states?
- It can be shown that only one $\mathbf{A}$ works
- Becker even found a way of computing $\mathbf{A}$ given the automaton
- It can also be shown that for any $\bar{e}$, if $\mathcal{A}$ is a subautomaton, its location in the structure is unique

- It is natural to ask for what matrices $\mathbf{A}$ and vectors $\bar{e}$ can we find a given $\mathcal{A}$ in $\mathfrak{C}(\mathbf{A}, \bar{e})$, and at what vectors $\bar{v}$ are its states?
- It can be shown that only one $\mathbf{A}$ works
- Becker even found a way of computing $\mathbf{A}$ given the automaton
- It can also be shown that for any $\bar{e}$, if $\mathcal{A}$ is a subautomaton, its location in the structure is unique
- There are infinitely many choices of $\bar{e}$ though, and the goal is to understand them.

$\mathbb{Z}[x]$-module

- We can multiply vectors $\bar{v} \in \mathbb{Z}^m$ by scalars in $\mathbb{Z}$

$\mathbb{Z}[x]$-module

- We can multiply vectors $\bar{v} \in \mathbb{Z}^m$ by scalars in $\mathbb{Z}$
- Seemingly weird idea:

$\mathbb{Z}[x]$-module

- We can multiply vectors $\bar{v} \in \mathbb{Z}^m$ by scalars in $\mathbb{Z}$
- Seemingly weird idea:
  - Can we take *polynomials* as our scalars instead?

$\mathbb{Z}[x]$-module

- We can multiply vectors $\bar{v} \in \mathbb{Z}^m$ by scalars in $\mathbb{Z}$
- Seemingly weird idea:
  - Can we take *polynomials* as our scalars instead?
- Typically this is done by setting $x\bar{v} = \mathbf{A}\bar{v}$ for some linear transformation.

$\mathbb{Z}[x]$-module

- We can multiply vectors $\bar{v} \in \mathbb{Z}^m$ by scalars in $\mathbb{Z}$
- Seemingly weird idea:
  - Can we take *polynomials* as our scalars instead?
- Typically this is done by setting $x\bar{v} = \mathbf{A}\bar{v}$ for some linear transformation.
- If only we had a obvious linear transformation floating around our structure that one might try...

$\mathbb{Z}[x]$-module

- We can multiply vectors $\bar{v} \in \mathbb{Z}^m$ by scalars in $\mathbb{Z}$
- Seemingly weird idea:
  - Can we take *polynomials* as our scalars instead?
- Typically this is done by setting $x\bar{v} = \mathbf{A}\bar{v}$ for some linear transformation.
- If only we had a obvious linear transformation floating around our structure that one might try...
- For technical reasons, we'll use $\mathbf{A}^{-1}$ instead of $\mathbf{A}$.

$\mathbb{Z}[x]$-module

- We can multiply vectors $\bar{v} \in \mathbb{Z}^m$ by scalars in $\mathbb{Z}$
- Seemingly weird idea:
  - Can we take *polynomials* as our scalars instead?
- Typically this is done by setting $x\bar{v} = \mathbf{A}\bar{v}$ for some linear transformation.
- If only we had a obvious linear transformation floating around our structure that one might try...
- For technical reasons, we'll use $\mathbf{A}^{-1}$ instead of $\mathbf{A}$.

Definition

For $p \in \mathbb{Z}[x]$ and $\bar{v} \in \mathfrak{C}(\mathbf{A}, \bar{e})$, put $p \cdot \bar{v} = (p(\mathbf{A}^{-1}))\bar{v}$

- Why should we care?

- Why should we care?

Theorem

for each $\bar{v} \in \mathbb{Z}^m$, there is $p_{\bar{v}} \in \mathbb{Z}[x]$ such that $p_{\bar{v}} \cdot e_1 = \bar{v}$

- Why should we care?

Theorem

for each $\bar{v} \in \mathbb{Z}^m$, there is $p_{\bar{v}} \in \mathbb{Z}[x]$ such that $p_{\bar{v}} \cdot e_1 = \bar{v}$

Theorem

If $\bar{e}$ is an odd vector, then $\varphi_{\bar{e}} : \mathfrak{C}(\mathbf{A}, \bar{e}_1) \hookrightarrow \mathfrak{C}(\mathbf{A}, \bar{e})$ by $\varphi_{\bar{e}}(\bar{v}) = p_{\bar{e}} \cdot \bar{v}$ is an embedding, and preserves the group structure and the residuation structure.

- Why should we care?

Theorem

for each $\bar{v} \in \mathbb{Z}^m$, there is $p_{\bar{v}} \in \mathbb{Z}[x]$ such that $p_{\bar{v}} \cdot e_1 = \bar{v}$

Theorem

If $\bar{e}$ is an odd vector, then $\varphi_{\bar{e}} : \mathfrak{C}(\mathbf{A}, \bar{e}_1) \hookrightarrow \mathfrak{C}(\mathbf{A}, \bar{e})$ by $\varphi_{\bar{e}}(\bar{v}) = p_{\bar{e}} \cdot \bar{v}$ is an embedding, and preserves the group structure and the residuation structure.

- This embedding is surjective if and only if $p_{\bar{e}}$ is a unit in $\mathbb{Z}[x]/\chi$, where χ is the characteristic polynomial of $\mathbf{A}^{-1}$

- Why should we care?

Theorem

for each $\bar{v} \in \mathbb{Z}^m$, there is $p_{\bar{v}} \in \mathbb{Z}[x]$ such that $p_{\bar{v}} \cdot e_1 = \bar{v}$

Theorem

If $\bar{e}$ is an odd vector, then $\varphi_{\bar{e}} : \mathfrak{C}(\mathbf{A}, \bar{e}_1) \hookrightarrow \mathfrak{C}(\mathbf{A}, \bar{e})$ by $\varphi_{\bar{e}}(\bar{v}) = p_{\bar{e}} \cdot \bar{v}$ is an embedding, and preserves the group structure and the residuation structure.

- This embedding is surjective if and only if $p_{\bar{e}}$ is a unit in $\mathbb{Z}[x]/\chi$, where χ is the characteristic polynomial of $\mathbf{A}^{-1}$
- So different residuation vectors give groups which *extend* the group $\mathfrak{C}(\mathbf{A}, \bar{e}_1)$

- Why should we care?

Theorem

for each $\bar{v} \in \mathbb{Z}^m$, there is $p_{\bar{v}} \in \mathbb{Z}[x]$ such that $p_{\bar{v}} \cdot e_1 = \bar{v}$

Theorem

If $\bar{e}$ is an odd vector, then $\varphi_{\bar{e}} : \mathfrak{C}(\mathbf{A}, \bar{e}_1) \hookrightarrow \mathfrak{C}(\mathbf{A}, \bar{e})$ by $\varphi_{\bar{e}}(\bar{v}) = p_{\bar{e}} \cdot \bar{v}$ is an embedding, and preserves the group structure and the residuation structure.

- This embedding is surjective if and only if $p_{\bar{e}}$ is a unit in $\mathbb{Z}[x]/\chi$, where χ is the characteristic polynomial of $\mathbf{A}^{-1}$
- So different residuation vectors give groups which *extend* the group $\mathfrak{C}(\mathbf{A}, \bar{e}_1)$
- Also, if $p_{\bar{e}}$ divides $p_{\bar{r}}$, then $\mathfrak{C}(\mathbf{A}, \bar{r})$ extends $\mathfrak{C}(\mathbf{A}, \bar{e})$.

Theorem

If $\mathcal{A}$ is an automaton whose group is $\mathbb{Z}^m$, then for each odd state in $\mathcal{A}$, there is exactly one $\bar{e}$ which locates that state at $\bar{e}_1$ in $\mathfrak{C}(\mathbf{A}, \bar{e})$. Further, if $\bar{e}$ and $\bar{r}$ are two such residuation vectors, they differ by a unit. This procedure is effective.

Theorem

If $\mathcal{A}$ has a state located at $\bar{v} \in \mathfrak{C}(\mathbf{A}, \bar{e})$, then $\bar{v}$ is located at $p \cdot \bar{v} \in \mathfrak{C}(\mathbf{A}, p \cdot \bar{e})$

- Incredibly, we now understand residuation vectors!

Theorem

If $\mathcal{A}$ is an automaton whose group is $\mathbb{Z}^m$, then for each odd state in $\mathcal{A}$, there is exactly one $\bar{e}$ which locates that state at $\bar{e}_1$ in $\mathfrak{C}(\mathbf{A}, \bar{e})$. Further, if $\bar{e}$ and $\bar{r}$ are two such residuation vectors, they differ by a unit. This procedure is effective.

Theorem

If $\mathcal{A}$ has a state located at $\bar{v} \in \mathfrak{C}(\mathbf{A}, \bar{e})$, then $\bar{v}$ is located at $p \cdot \bar{v} \in \mathfrak{C}(\mathbf{A}, p \cdot \bar{e})$

- Incredibly, we now understand residuation vectors!
- First find $\bar{r}$ such that $\mathcal{A}$ has a state at $\bar{e}_1$.
- $\mathcal{A}$ is a subautomaton of $\mathfrak{C}(\mathbf{A}, \bar{e})$ if and only if $p_{\bar{r}}$ divides $p_{\bar{e}}$
- Also, if $\mathcal{A}$ *is* a subautomaton, then $qp_{\bar{r}} = p_{\bar{e}}$, and $\mathcal{A}$ is located at $q \cdot \bar{e}_1$

- It turns out we can "scale by an infinite polynomial" to get a universal structure which contains *every* automaton (with the correct matrix $\mathbf{A}$) at exactly one location.
- The construction is a bit involved, so we don't have time to discuss it, but it is computable, and removes the need for the extra parameter $\bar{e}$.

Questions?