WEEK #3
EQUIVALENCE RELATIONS

REFLEXIVE, SYMMETRIC, TRANSITIVE

LET R BE AN EQUIV RELATION OVER U

R CREATES EQUIV. CLASSES THAT PARTITION U INTO NON-OVERLAPPING SUBSETS. ALL ELEMENTS OF U ARE IN SOME SUBSET AS xRx, $\forall x \in U$

THE **INDEX** OF R IS THE NUMBER OF PARTITIONS

R HAS FINITE INDEX IF R INDUCES A FINITE NUMBER OF PARTITIONS

IF R IS OVER STRINGS, E.G., OVER Σ^*

THEN R IS CALLED RIGHT INVARIANT IFF

$$xRy \Rightarrow \forall z \; xzRyz$$

DFA INDUCES A RIGHT INVARIANT EQUIVALENCE RELATION WHOSE INDEX IS $|Q|$ SO LONG AS ALL STATES ARE REACHABLE I.E., $\forall q \in Q \; \exists x \; \delta^*(q_0, x) = q$ WHERE $\mathcal{Q} = (Q, \Sigma, \delta, q_0, F)$

INDUCED RELATION $R_{\mathcal{Q}}$ IS

$x \, R_{\mathcal{Q}} \, y$ IFF $\delta^*(q_0 x) = \delta^*(q_0, y)$

OBVIOUSLY $\forall z \, \delta^*(q_0, xz) = \delta^*(q_0, yz)$ WHENEVER $\delta^*(q_0, x) = \delta^*(q_0, y)$ AND SO AS $\mathcal{Q}$ IS DETERMINISTIC

$R_{\mathcal{Q}}$ IS RIGHT INVARIANT

STATE COMPATIBILITY (INDISTINGUISHABLE)

$Q = (Q, \Sigma, \delta, q_0, F)$ IS SOME DFA

p, q ARE COMPATIBLE IFF

$\forall z \in \Sigma^* \; \delta^*(p, z) \in F \iff \delta^*(q, z) \in F$

THIS MEANS THEY BOTH ACCEPT OR BOTH REJECT z, AND DO SO FOR ALL z, INCLUDING $z = \lambda$.

IF TWO STATES ARE COMPATIBLE, WE CAN MERGE THEM.

APPROACH TO MINIMIZATION

1. REMOVE ANY STATE THAT IS NOT REACHABLE FROM START STATE (q_0 HERE)
2. DISCOVER AND MERGE COMPATIBLE STATES

STEP 1 IS DONE VIA DFS

STEP 2 IS EASIER IF WE DISCOVER IN COMPATIBLE STATES

p, q ARE INCOMPATIBLE IFF

$$\exists z \in \Sigma^* \; \delta^*(p,z) \in F \;\&\; \delta^*(q,z) \notin F$$
$$\text{OR} \quad \delta^*(p,z) \notin F \;\&\; \delta^*(q,z) \in F$$

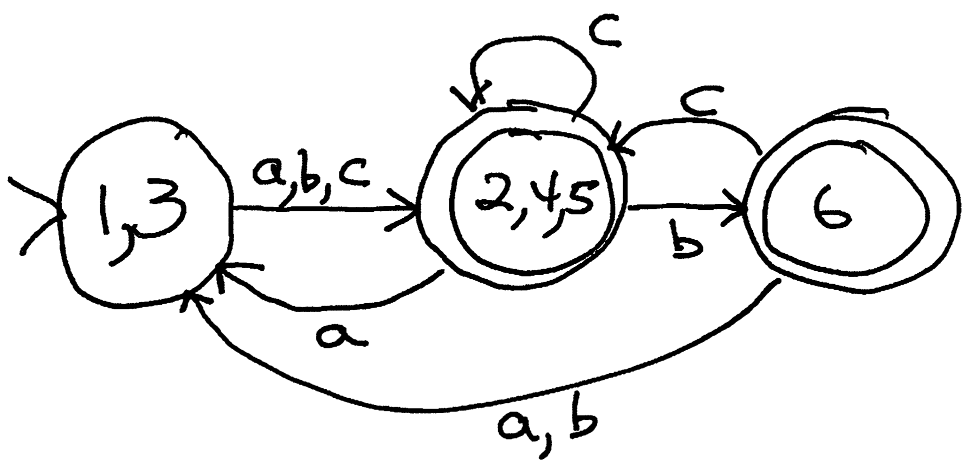
MINIMIZATION BY INCOMPATIBILITY

	a	b	c
>1	5	2	2
2	1	6	2
3	2	4	5
4	3	6	2
5	3	6	5
6	1	3	4

	1	2	3	4	5
2	X				
3	2,5 2,4	X			
4	X	1,3	X		
5	X	1,3	X	2,5	
6	X	~~3,6 2,4~~	X	~~1,3 3,6 2,4~~	~~1,3 3,6 4,5~~

1,3
4,5
2,4
2,5
6

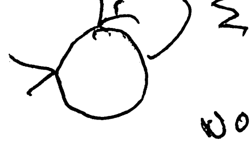
MINIMIZED DFA



NOTE! WE WILL PROVE THERE IS A UNIQUE MINIMUM ~~STATE~~ DFA FOR ANY REGULAR LANGUAGE L, UP TO RENAMING OF STATES

DECISION PROBLEMS FOR REGULAR LANGUAGES

MEMBERSHIP: $w \in \mathcal{L}(Q)$? $Q = (Q, \Sigma, \delta, q_0, F)$



NO FINAL $\emptyset$

EMPTINESS: $\mathcal{L}(Q) = \emptyset$?

OR REMOVE USELESS,

EVERYTHING: $\mathcal{L}(Q) = \Sigma^*$?

FINITENESS: $\mathcal{L}(Q)$ IS FINITE?

NO LOOPS FROM FINAL TO FINAL ON Q HAVING NO USELESS STATES

EQUIVALENCE: $\mathcal{L}(Q_1) = \mathcal{L}(Q_2)$?

$\text{MIN}(Q_1) = \text{MIN}(Q_2)$

WHERE $\text{MIN}(Q)$ IS MINIMUM STATE AUTOMATON

CLOSURE PROPERTIES OF REGULAR LANGUAGES

DID COMPLEMENT AND INTERSECTION USING DFAs

ALSO DID REVERSAL USING DFA BUT COULD HAVE USED REGULAR EXPRESSIONS

GET UNION, CONCATENATION AND $*$ TRIVIALLY USING REGULAR EXPRESSIONS
ALSO EASY TO DO WITH NFAs

SUBSTITUTION IS A MAPPING FROM EACH LETTER IN Σ TO SOME LANGUAGE
THIS IS ~~DENOTED~~ BY A FUNCTION, SAY f,

WHERE
$f(a) = L_a \quad \forall a \in \Sigma$

WHEN EACH L_a IS REGULAR, WE WANT CLOSURE SO THAT IF L IS REGULAR

AND $f(w) = \lambda$ IF $w = \lambda$
$f(ax) = f(a)f(x)$ IF $a \in \Sigma, x \in \Sigma^*$

THEN $f(L) = \{f(w) \mid w \in L\}$
IS ALSO REGULAR

SUBSTITUTION & HOMOMORPHISM

CONTINUING

LET L & L_a $\forall a \in \Sigma$ BE REGULAR
AND LET $f(a) = L_a$ ~~BE~~ A SUBSTITUTION
THEN $f(L)$ IS REGULAR

HOW TO ATTACK WITH REG. EXPRESSIONS

REPLACE EVERY INSTANCE OF EACH $a \in \Sigma$ IN r
WITH (r_a), WHERE r IS REG. EXPR FOR L
AND $\forall a$ r_a IS REGULAR EXPR FOR L_a

HOW TO ATTACK WITH NFAs. LET Q_L BE
DFA FOR L AND Q_{La}, $a \in \Sigma$, BE DFA FOR L_a
IN Q_L AS FOLLOWS:
CHANGE EACH a-TRANSITION

$$q \xrightarrow{a} p \quad \text{TO} \quad q \xrightarrow{\lambda} Q_{La} \xrightarrow{\lambda} p$$

NOTE: A HOMOMORPHISM IS A SUBSTITUTION
WHERE EACH L_a IS A SINGLE STRING

EXAMPLE SUBST

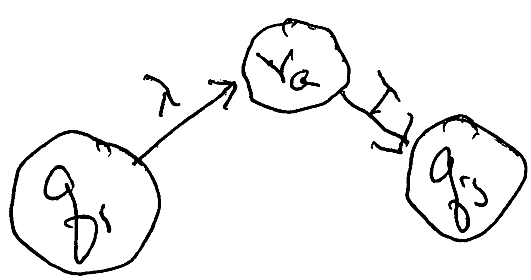
$$a+(b\cdot c)^*$$

$f(a) = L_a$ $\quad r_a$

$f(b) = L_b$ $\quad r_b$

$f(c) = L_c$ $\quad r_c$

$$(r_a)+((r_b)\cdot(r_c))^*$$



QUOTIENT I

$$R_1 / R_2 = \{ x \mid \exists y \in R_2 \text{ s.t. } xy \in R_1 \}$$

KEEPS PREFIXES FROM R_1 THAT HAVE REMAINING POSTFIX IN R_2

$$R_1 / R_2 = h(f(R_1) \cap \Sigma^* \cdot g(R_2))$$

WHERE

$$f(a) = \{a, a'\} \quad a \in \Sigma$$

$$g(a) = a' \quad a \in \Sigma$$

$$\left.\begin{array}{l} h(a) = a \\ h(a') = \lambda \end{array}\right\} a \in \Sigma$$

f IS A SUBSTITUTION

g, h ARE HOMOMORPHISMS

WHY DOES IT WORK?

QUOTIENT 2

$f(R_1)$ GIVES US EVERY COMBINATION OF EVERY STRING IN R_1 WITH SOME OF THE LETTER PRIMED

E.G.

$f(aba) = \{aba, aba', ab'a, ab'a', a'ba, a'ba', a'b'a, a'b'a'\}$

IF $|w| = k$

$|f(w)| = 2^k$

NOTE! $f(R_1)$ IS REGULAR WHENEVER R_1 IS REG.

$g(R_2)$ GIVES US THE ORIGINAL LANGUAGE BUT WITH ALL WORDS ~~HAVING~~ PRIMED LETTERS

$g(aba) = a'b'a'$

CLEARLY, $g(R_2)$ IS REGULAR

$\Sigma^* \cdot g(R_2)$ IS REG. (CLOSURE UNDER ~~INTERSECTION~~ CONCATENATION)

$f(R_1) \cap \Sigma^* \cdot g(R_2)$ IS REG. (CLOSURE UNDER INTERSECTION)

NOTE: $f(R_1) \cap \Sigma^* \cdot g(R_2) =$

$\{xy' \mid xy \in R_1, y \in R_2\}$

QUOTIENT 3

$$h(f(R_1) \cap \Sigma^* \cdot g(R_2))$$
$$= h(\{xy' \mid xy \in R_1, y \in R_2\})$$
$$= \{x \mid \exists y \in R_2 \ni xy \in R_1\}$$
$$= R_1 / R_2$$

THINK OF $\Sigma^* \cdot g(R_2)$
AS A FILTER THAT KEEPS
ONLY DESIRABLE STRINGS FROM
$f(R_1)$ WHOSE UNPRIMED PARTS ARE TO BE KEPT

h IS THE CLEANUP THAT RIDS
US OF PRIMED LETTERS

METATECHNIQUE

$h(f(L) \cap$ SOME REG EXPR. AS FILTER$)$

PREFIX (INIT) (FIRST)

$$\text{PREFIX}(R) = h(f(R) \cap \Sigma^* g(\Sigma^*))$$
$$= R/\Sigma^*$$

POSTFIX (FINAL) AKA SUFFIX

$$\text{SUFFIX}(R) = h(f(R) \cap g(\Sigma^*)\Sigma^*)$$
$$= (R^R/\Sigma^*)^R$$

SUBSTRING

$$\text{SUBSTRING}(R) = h(f(R) \cap g(\Sigma^*)\Sigma^* g(\Sigma^*))$$

SUFFIX (PREFIX(R))

PREFIX (SUFFIX(R))

HOW TO ATTACK

IN EACH CASE, ASSUME DFA, $\mathcal{Q}$, ACCEPTS L
ALSO ASSUME $\mathcal{Q}$ HAS NO UNREACHABLE STATES

PREFIX DIRECTLY

CHANGE $\mathcal{Q} = (Q, \Sigma, \delta, q_0, F)$
BY MAKING NEW FINAL SET
$F' = \{ q \mid \exists x \ \delta^*(q, x) \in F \}$

POSTFIX DIRECTLY

CHANGE $\mathcal{Q} = (Q, \Sigma, \delta, q_0, F)$
BY MAKING NEW INITIAL SET OF START STATES
$q_0' = \{ q \mid \exists x \ \delta^*(q, x) \in F \}$

SUBSTRING FROM ABOVE

THINK ABOUT IT

CHALLENGE 1

$\text{EVERYOTHER}(L) = \{ x_1 x_3 \cdots x_{2k-1} \mid k \geq 1$

EITHER $x_1 x_2 x_3 x_4 \cdots x_{2k-1} x_{2k} \in L$

OR $x_1 x x_3 x_4 \cdots x_{2k-1} \in L$

WHERE $\forall i \; x_i \in \Sigma \}$

L REGULAR IMPLIES EVERYOTHER (L) REGULAR

$$h\left(f(L) \cap \left((\Sigma \cdot g(\Sigma))^{+} \cup (\Sigma \cdot g(\Sigma))^{*} \Sigma \right)\right)$$

CHALLENGE 2

$\text{PROPERPREFIX}(L) = \{ x \mid \exists y \in \Sigma^{+}, \; xy \in L \;\&\; x \neq \lambda \}$

L REGULAR IMPLIES PROPERPREFIX (L) REGULAR

$$h\left(f(L) \cap \Sigma^{+} g(\Sigma^{+})\right)$$

CHALLENGE 3

ARE REGULAR LANGUAGES CLOSED UNDER

$\frac{1}{2}(L) = \{ w \mid \exists y \in \Sigma^*, |y| = |w| \,\&\, wy \in L \}$?

HINT: IT IS TRUE

HINT: CANNOT USE REGEXPR. TO SHOW IT

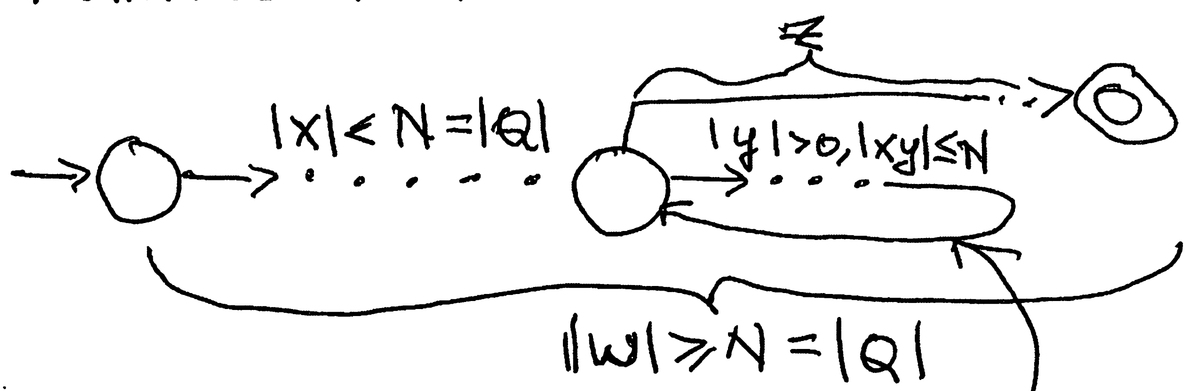
HINT: CANNOT USE OUR META APPROACH

HINT: CANNOT DO DIRECTLY WITH DFA

HINT: KEY IS NFA

WEEK #3

PUMPING LEMMA



$$w = \underbrace{xyz}_{\ge N = |Q|} \in L$$

CAN REPEAT y-PART 0 OR MORE TIMES

CONSEQUENCE

$$xy^i z \in L, \forall i \ge 0$$

P.L. IS BASED ON PIGEONHOLE PRINCIPLE

IF HAVE N PLACES AND VISIT THESE > N TIMES THEN MUST REUSE AT LEAST ONE

STRING OF LENGTH N MAKES N+1 VISITS

$\lambda, a_1, a_2, \ldots, a_N$

PUMPING LEMMA

PIDGEON HOLE PRICIPLE

$Q = (Q, \Sigma, \delta, q_0, F)$

$Q = \{q_1, q_2, \ldots, q_n\}$ $n = |Q|$

$w \in L$

$|w| \geq n$ $n = |Q|$

THEN

$\delta(q_1, a_1 a_2 \ldots a_{|w|})$

FIRST REPEAT IS i_k

$$w = \underbrace{q_{i_1}\, q_{i_2} \cdots q_{i_k}}_{x}\ \underbrace{\cdots}_{y}\ z$$

$|xy| \leq N$ $|y| > 0$

STATEMENT OF PUMPING LEMMA FOR REGULAR LANGUAGES

LET L BE REGULAR, THEN $\exists N > 0$, DEPENDENT ONLY ON L, SUCH THAT, IF $w \in L$ AND $|w| \geq N$, THEN w CAN BE WRITTEN IN THE FORM $w = xyz$, WHERE $|xy| \leq N$, $|y| > 0$ AND $\forall i \geq 0$ $xy^iz \in L$

PUMPING LEMMA GIVES RISE TO AN ADVERSARIAL PROCESS

EXAMPLE: $L = \{a^n b^n \mid n > 0\}$

ME: ASSUME L IS REGULAR

P.L.: GIVES ME $N > 0$

I CANNOT CHOOSE N.

I CANNOT MAKE ANY ASSUMPTION ABOUT N EXCEPT THAT $N > 0$

ME: I CHOOSE A STRING IN L.

CHOOSE $a^N b^N \in L$

P.L.: CHOOSES AN x, y, z SUCH THAT $a^N b^N = xyz$, $|xy| \le N$, $|y| > 0$

AND P.L. STATES $\forall i \geq 0 \; xy^i z \in L$

ME: I CHOOSE VALUE OF i

I WILL CHOOSE $i = 0$

P.L.: IT SAYS $xz \in L$ BUT

$xz = a^{N-|y|} b^N$, $|y| > 0$

ME: GOTCHA

$xz = a^{N-|y|} b^N$ HAS FEWER a's THAN b's SO $xz \notin L$

CONFLICT: WE MUST CONCLUDE L IS NOT REGULAR

APPLY P.L. TO

$L = \{ xx \mid x \in \{a,b\}^+ \}$

ASSUME L IS REGULAR

P.L.: GIVES ME $N > 0$

ME: CHOOSE $s = a^N b a^N b \in L$

P.L.: $s = xyz$ WHERE $|xy| \leq N, |y| > 0$

AND $\forall i \geq 0$ $xy^i z \in L$

ME: CHOOSE $i = 0$

P.L.: $xz \in L$ SO

$a^{N-|y|} b a^N b \in L$ WHERE $|y| > 0$

ME: $N - |y| < N$ SO FEWER a'S PRECEDE FIRST b THAN PRECEDE SECOND b (CONSECUTIVE a'S IN EACH CASE)

THUS, $a^{N-|y|} b a^N b \notin L$ $\times$

CONFLICT: WE MUST CONCLUDE L IS **NOT** REGULAR

APPLY P.L. TO
$L = \{ x w x \mid x, w \in \{a,b\}^+ \}$

ASSUME L REGULAR

P.L.: GIVES ME $N > 0$
ME: CHOOSE $s = a^N b\, a\, a^N b \in L$
P.L.: $s = xyz$ WHERE $|xy| \le N$, $|y| > 0$
AND $\forall i \ge 0$ $xy^i z \in L$

ME: CHOOSE $i = 2$
P.L.: $xy^2z = xyyz \in L$ SO
$a^{N+|y|} b\, a\, a^N b \in L$ WHERE $|y| > 0$

ME: $N + |y| > N$ SO MORE THAN N a's PRECEDE b
BUT $a a^N$ PRECEDING SECOND b
CANNOT CARVE OUT A PART FOR
w, $|w| > 0$, AND HAVE $> N$ a's
IN THIS SUBWORD
THUS, $a^{N+|y|} b\, a\, a^N b \notin L$ ✗

CONFLICT: WE MUST CONCLUDE L IS <u>NOT</u>
REGULAR

NOTE: ON PREVIOUS PAGE

COULD NOT USE $i=0$

WHY?

$a^{N-|y|}baa^{N}b \in L \quad |y|>0$

BY FACTORING AS FOLLOWS

$\underbrace{a^{N-|y|}}ba^{|y|+1}\underbrace{a^{N-|y|}}b.$

CAN VIEW AS

$x = a^{N-|y|}b$

$w = a^{|y|+1}$

SO THIS IS IN L

AND WE DO NOT HAVE A CONTRADICTION

APPLY P.L. TO

$L = \{ x x^R \mid x \in \{a,b\}^+ \}$

ASSUME L IS REGULAR

P.L.: GIVES ME $N > 0$

ME: CHOOSE $S = a^N bba^N$

P.L.: $S = xyz$ WHERE $|xy| \le N$, $|y| > 0$

AND $\forall i \ge 0$ $xy^iz \in L$

ME: CHOOSE $i = 0$

P.L.: $xz \in L$ SO

$a^{N-|y|} bba^N \in L$ WHERE $|y| > 0$

ME: $N - |y| < N$ SO FEWER a's

PRECEDE ~~FIRST~~ bb THAN FOLLOW

THUS, $a^{N-|y|} bba^N \notin L$ ✗

CONFLICT! WE MUST CONCLUDE

L IS <u>NOT</u> REGULAR

CHALLENGE

$L_1 = \{ x x^R \mid x \in \{a,b\}^+ \}$ IS NOT REGULAR

$L_2 = \{ x x \mid x \in \{a,b\}^+ \}$ IS NOT REGULAR

$L_3 = \{ x w x \mid w, x \in \{a,b\}^+ \}$ " " "

WHAT ABOUT

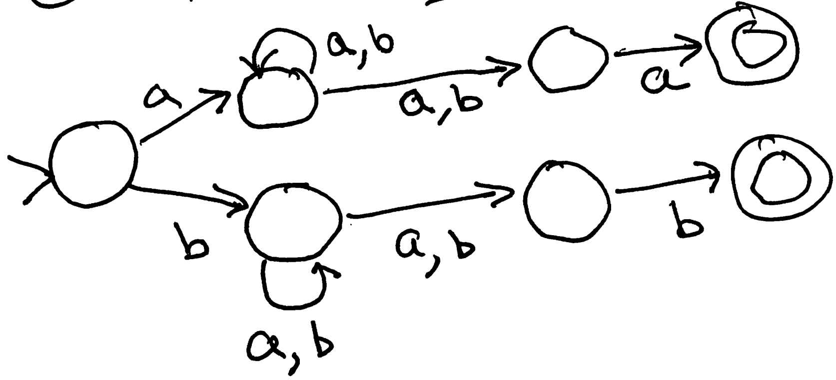
$L = \{ x w x^R \mid x, w \in \{a,b\}^+ \}$?

$a w a \qquad a \in \Sigma$

$b w b \qquad b \in \Sigma$

$a \Sigma^* a$

$\cup \; b \Sigma^* b$ IS REGULAR



MYHILL - NERODE

THE FOLLOWING ARE EQUIVALENT

(1) $L = \mathcal{L}(Q)$ FOR SOME DFA Q

(2) L IS THE UNION OF SOME OF THE CLASSES OF A RIGHT INVARIANT EQUIV. REL., R, OF FINITE INDEX

(3) THE SPECIFIC RIGHT INVARIANT EQUIV. REL ASSOCIATED WITH L

$$x R_L y \Leftrightarrow \forall z\, [xz \in L \Leftrightarrow yz \in L]$$

HAS FINITE INDEC

(1) $\Rightarrow$ (2). JUST USE R_Q FOR R

(2) $\Rightarrow$ (3) JUST NOTE

$$x R y \Rightarrow \forall z\, [xz\, R\, yz]$$
$$\Rightarrow \forall z\, [xz \in L \Leftrightarrow yz \in L]$$

AND SO R IS A REFINEMENT OF R_L AND HAS INDEX $\geq R_L$

THUS, INDEX OF R_L IS FINITE

(3) $\Rightarrow$ (1) JUST CHOOSE EACH PARTITION OF R_L AS A STATE AND

$$\delta([x], a) = [xa]$$
$$F = \{[x] \mid x \in L\}$$

NOTION OF REPRESTATIVE ELEMENT

ON PRIOR PAGE WE DESCRIBED AN EQUIVALANCE CLASS OVER SOME SET OF STRINGS BY JUST CHOOSING A REPRESENTATIVE ELEMENT, SAY X, AND USING $[x]$ TO INDICATE $\{y \mid xRy\}$

NOTION OF UNIQUE MIN STATE DFA

STARTING WITH SOME DFA Q WE CAN USE TECHNIQUE SHOWN LAST TIME TO FIND Q', $Q = (Q, \Sigma, \delta, q_0, F)$, $Q' = (Q', \Sigma, \delta', q_0', F')$

WHERE $|Q'| \leq |Q|$ AND

$$x R_Q y \Rightarrow x R_{Q'} y$$

THIS IS PRECISELY THE AUTATAN YOU GET FROM MYHILL-NERODE'S STAGE 3

MYHILL-NERODE AND WHAT IS NOT REGULAR

FOR ANY LANGUAGE L, DEFINE R_L FROM MYHILL-NERODE, THAT IS,

$$xR_Ly \Leftrightarrow \forall z\ [xz \in L \Leftrightarrow yz \in L]$$

IF R_L HAS FINITE INDEX, L IS REGULAR,
ELSE L IS NOT REGULAR

E.G., LET $L = \{x \mid x \in \{0,1\}^*$ AND x IS ODD PARITY$\}$

TWO CLASSES INDUCED BY R_L ARE

$[\lambda] = \{x \mid x \in \{0,1\}^*$ AND x HAS EVEN PARITY$\}$
$[1] = \{x \mid x \in \{0,1\}^*$ AND x HAS ODD PARITY$\}$
$L = [1]$

INDEX IS 2

SO MINIMAL STATE DFA HAS TWO STATES