I. Recognisability by Semigroups and Monoids

Radu Iosif (CNRS/VERIMAG)

Semigroups

$(S, \cdot)$

$a, b \in S \Rightarrow a \cdot b \in S$

$(a \cdot b) \cdot c = a \cdot (b \cdot c)$

$abc \stackrel{\text{def}}{=} a \cdot b \cdot c$

Example 1 $(\{0,1,...,n-1\}, +_n)$ $x +_n y \stackrel{\text{def}}{=} x+y \bmod n$

Example 2 $(2^{A\times A}, \circ)$ $P \circ Q \stackrel{\text{def}}{=} \{(x,y) \mid \exists z \in A \,.\, (x,z) \in P \text{ and } (z,y) \in Q\}$

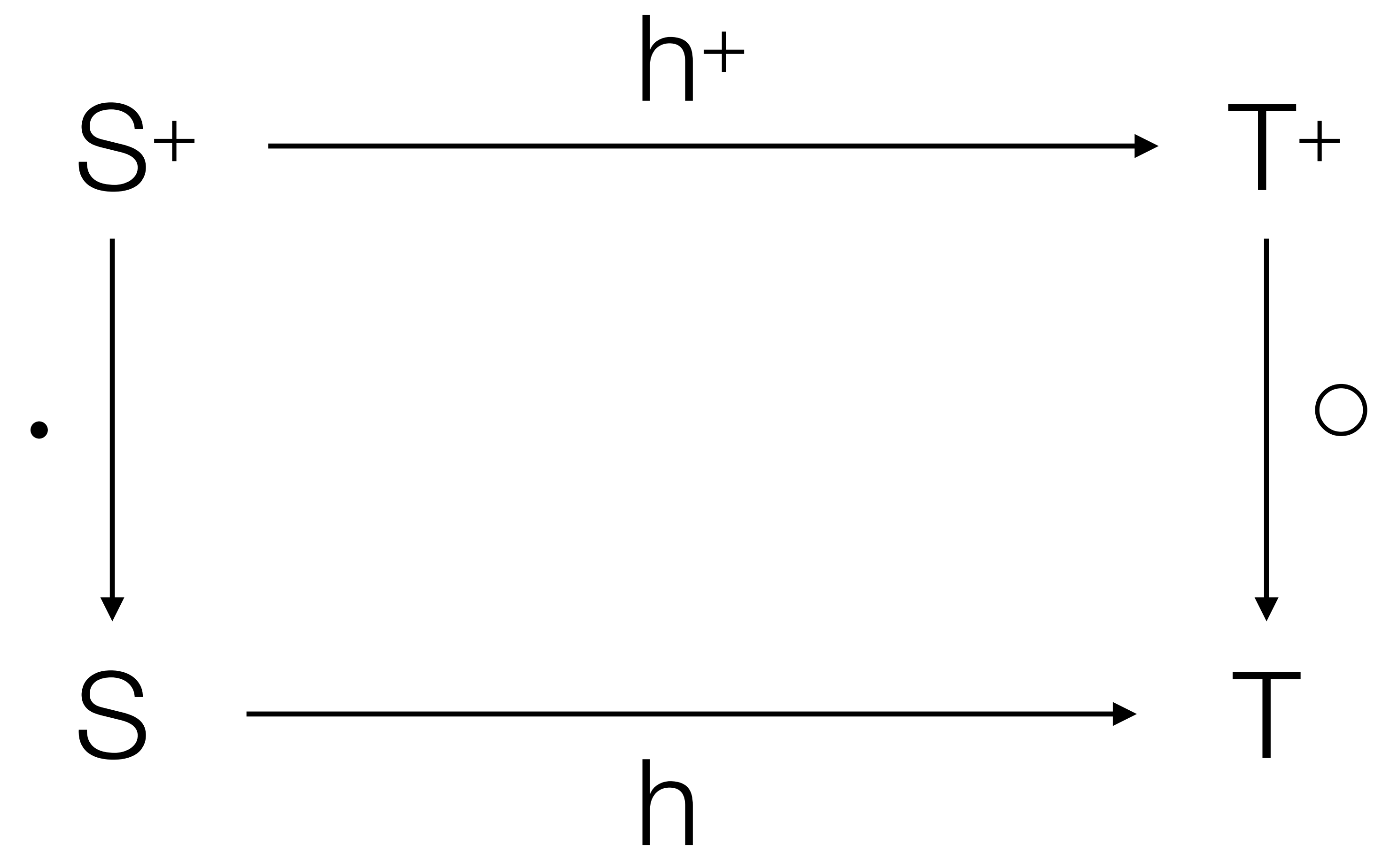
Example 3 $(\Sigma^+, \cdot)$ $u \cdot v \stackrel{\text{def}}{=} uv$

Semigroup Homomorphism

$h : (S, \cdot) \rightarrow (T, \circ)$ $\quad\quad$ $h(a \cdot b) = h(a) \circ h(b)$

$S^+ \stackrel{\text{def}}{=}$ set of non-empty sequences of elements from S

$h^+(a_1, a_2, ..., a_n) \stackrel{\text{def}}{=} h(a_1), h(a_2), ..., h(a_n)$



$h(a_1 \cdot a_2 \cdot ... \cdot a_n) = h(a_1) \circ h(a_2) \circ ... \circ h(a_n)$

Monoids

$$a, b \in S \Rightarrow a \cdot b \in S$$

$(S, 1, \cdot)$ $\quad$ $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ $\quad$ $abc \stackrel{def}{=} a \cdot b \cdot c$

$$1 \cdot a = a \cdot 1 = a$$

Example 1 $\quad$ $(\{0,1,\ldots,n-1\}, 0, +_n)$ $\quad$ $x +_n y \stackrel{def}{=} x+y \bmod n$

Example 2 $\quad$ $(2^{A \times A}, Id_A, \circ)$ $\quad$ $P \circ Q \stackrel{def}{=} \{(x,y) \mid \exists z \in A \,.\, (x,z) \in P \text{ and } (z,y) \in Q\}$

$Id_A \stackrel{def}{=} \{(x,x) \mid x \in A\}$

Example 3 $\quad$ $(\Sigma^*, \varepsilon, \cdot)$ $\quad$ $u \cdot v \stackrel{def}{=} uv$

Monoid Homomorphism

$h : (S, 1_S, \cdot\,) \rightarrow (T, 1_T, \circ)$

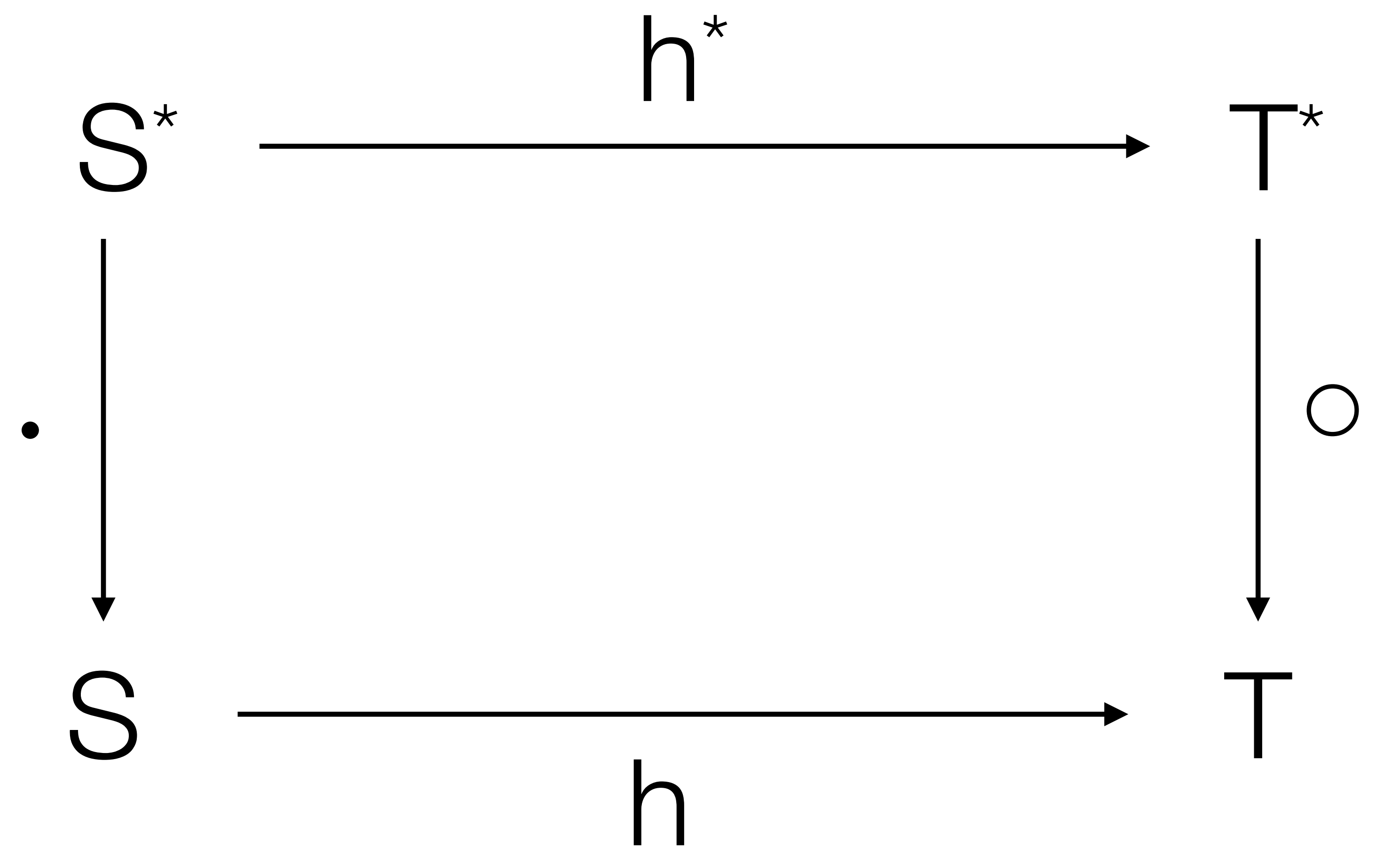
$h(a \cdot b) = h(a) \circ h(b)$

$h(1_S) = 1_T$

$S^* \overset{\text{def}}{=}$ set of sequences of elements from S

$h^*(a_1, a_2, ..., a_n) \overset{\text{def}}{=} h(a_1), h(a_2), ..., h(a_n)$

$h^*(\varepsilon) \overset{\text{def}}{=} \varepsilon$



$h(a_1 \cdot a_2 \cdot ... \cdot a_n) \;=\; h(a_1) \circ h(a_2) \circ ... \circ h(a_n)$

Recognisability

(S, ・) semigroup (same definition for S monoid)

$$S$$
$$\cup\!\!\text{I}$$
$$L$$

Recognisability

(S, ·) semigroup (same definition for S monoid)

$$S \xrightarrow{\quad h \quad} T \text{ finite semigroup}$$

$$\begin{array}{c} S \\ \cup\!\!| \\ L \end{array}$$

Recognisability

(S, ・) semigroup (same definition for S monoid)

$$
\begin{array}{ccc}
S & \xrightarrow{\quad h \quad} & T \text{ finite semigroup} \\
\cup\!| & & \cup\!| \\
L & & U \text{ accepting subset}
\end{array}
$$

Recognisability

(S, ・) semigroup (same definition for S monoid)

$$\begin{array}{ccc} S & \xrightarrow{\quad h \quad} & T \text{ finite semigroup} \\ \cup\!\!\text{I} & & \cup\!\!\text{I} \\ L & = & h^{-1}(U) \text{ accepting subset} \end{array}$$

Recognisability

(S, ·) semigroup (same definition for S monoid)

$$\begin{array}{ccc} S & \xrightarrow{\quad h \quad} & T \text{ finite semigroup} \\ \rotatebox{90}{$\supseteq$} & & \rotatebox{90}{$\supseteq$} \\ L & = & h^{-1}(U) \text{ accepting subset} \end{array}$$

Example 1 The set of odd (even) numbers is recognisable in $(\mathbb{N}, +)$

Recognisable Word Languages

Theorem $L \subseteq \Sigma^*$ is recognisable in the monoid $(\Sigma^*, \varepsilon, \cdot)$

$\Leftrightarrow$

L is the language of a nondeterministic finite automaton $A = (Q, I, F, \rightarrow)$

"⇒"

$$\begin{array}{ccc} \Sigma^* & \xrightarrow{\quad h \quad} & (Q, 1_Q, \circ) \text{ finite monoid} \\ \cup\!\!| & & \cup\!\!| \\ L & = & h^{-1}(F) \end{array}$$

$A = (Q, \{1_Q\}, F, \rightarrow) \qquad q \overset{a}{\rightarrow} q \circ h(a)$

Recognisable Word Languages

Theorem $L \subseteq \Sigma^*$ is recognisable in the monoid $(\Sigma^*, \varepsilon, \cdot)$

$\Leftrightarrow$

L is the language of a nondeterministic finite automaton $A = (Q, I, F, \rightarrow)$

"⇐"

$$\Sigma^* \xrightarrow{h(w)=\{ (q,q') \mid q \overset{w}{\rightarrow} q'\}} (2^{Q\times Q}, Id_Q, \circ)$$

$$\cup\!\!| \qquad\qquad\qquad \cup\!\!|$$

$$L \quad = \quad h^{-1}(\{R \in 2^{Q\times Q} \mid \exists i \in I\ \exists f \in F\ .\ (i,f) \in R\})$$

Syntactic Monoids

$(S, 1_S, \cdot)$ monoid $\sim \subseteq S \times S$ $a \sim b \Rightarrow sat \sim sbt$, for all $s, t \in S$ congruence

Syntactic Monoids

$(S, 1_S, \cdot)$ monoid $\quad \sim \subseteq S \times S \quad a \sim b \Rightarrow sat \sim sbt$, for all $s, t \in S$ congruence

$L \subseteq S \quad a \sim_L b \stackrel{\text{def}}{=} sat \in L \Leftrightarrow sbt \in L$, for all $s, t \in S$ syntactic congruence

Syntactic Monoids

$(S, 1_S, \cdot)$ monoid $\sim \subseteq S \times S$ $a \sim b \Rightarrow sat \sim sbt$, for all $s, t \in S$ congruence

$L \subseteq S$ $a \sim_L b \overset{\text{def}}{=} sat \in L \Leftrightarrow sbt \in L$, for all $s, t \in S$ syntactic congruence

$(S_{/\sim L}, [1_S]_{\sim L}, \cdot)$ $[a]_{\sim L} \cdot [b]_{\sim L} \overset{\text{def}}{=} [ab]_{\sim L}$ syntactic monoid

Syntactic Monoids

$(S, 1_S, \cdot)$ monoid $\sim \subseteq S \times S$ $a \sim b \Rightarrow sat \sim sbt$, for all $s, t \in S$ congruence

$L \subseteq S$ $a \sim_L b \stackrel{\text{def}}{=} sat \in L \Leftrightarrow sbt \in L$, for all $s, t \in S$ syntactic congruence

$(S_{/\sim L}, [1_S]_{\sim L}, \cdot)$ $[a]_{\sim L} \cdot [b]_{\sim L} \stackrel{\text{def}}{=} [ab]_{\sim L}$ syntactic monoid

Theorem [Myhill-Nerode]

$L \subseteq S$ is recognisable $\Leftrightarrow$ $S_{/\sim L}$ is finite

Syntactic Monoids

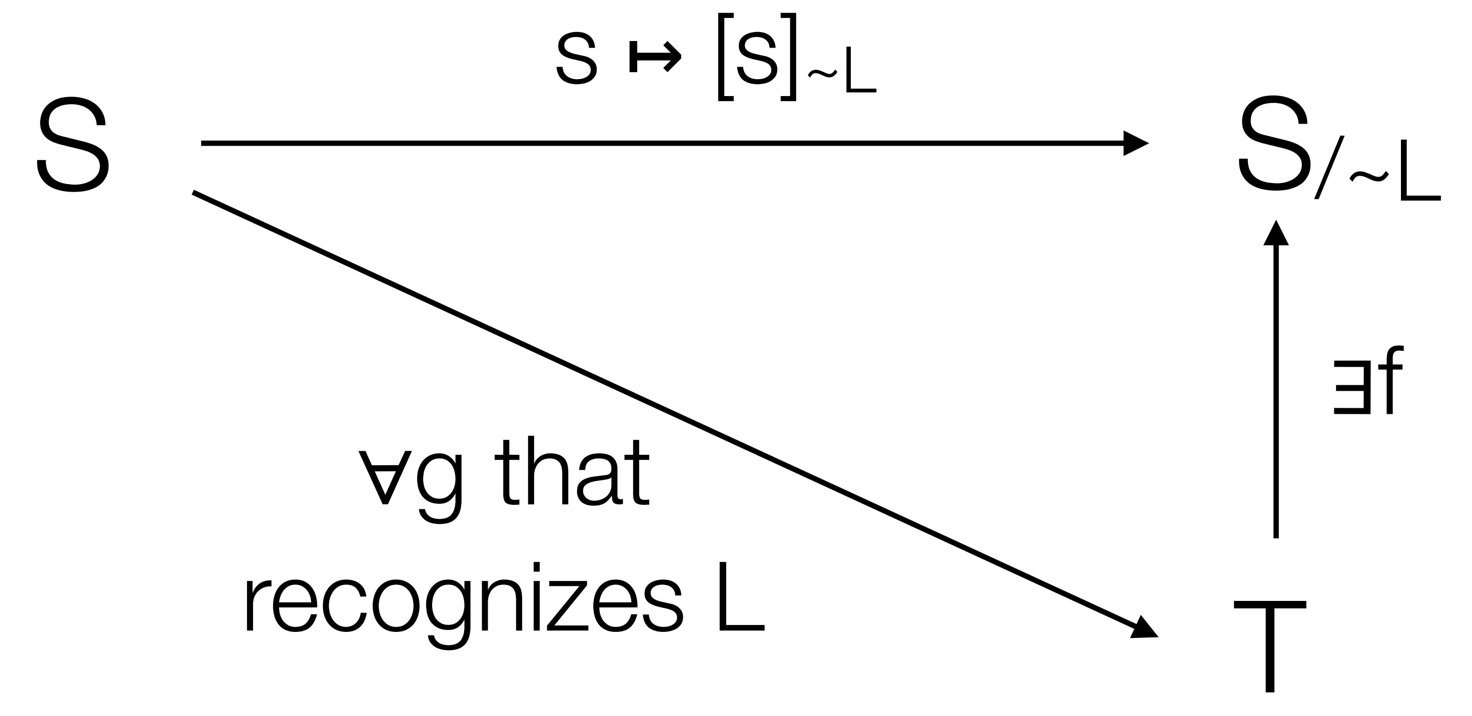
$(S, 1_S, \cdot)$ monoid $\sim\ \subseteq S \times S$ $a \sim b \Rightarrow sat \sim sbt$, for all $s, t \in S$ congruence

$L \subseteq S$ $a \sim_L b \overset{\text{def}}{=} sat \in L \Leftrightarrow sbt \in L$, for all $s, t \in S$ syntactic congruence

$(S_{/\sim L}, [1_S]_{\sim L}, \cdot)$ $[a]_{\sim L} \cdot [b]_{\sim L} \overset{\text{def}}{=} [ab]_{\sim L}$ syntactic monoid

Theorem [Myhill-Nerode]

$L \subseteq S$ is recognisable $\Leftrightarrow$ $S_{/\sim L}$ is finite



all homomorphisms
are surjective

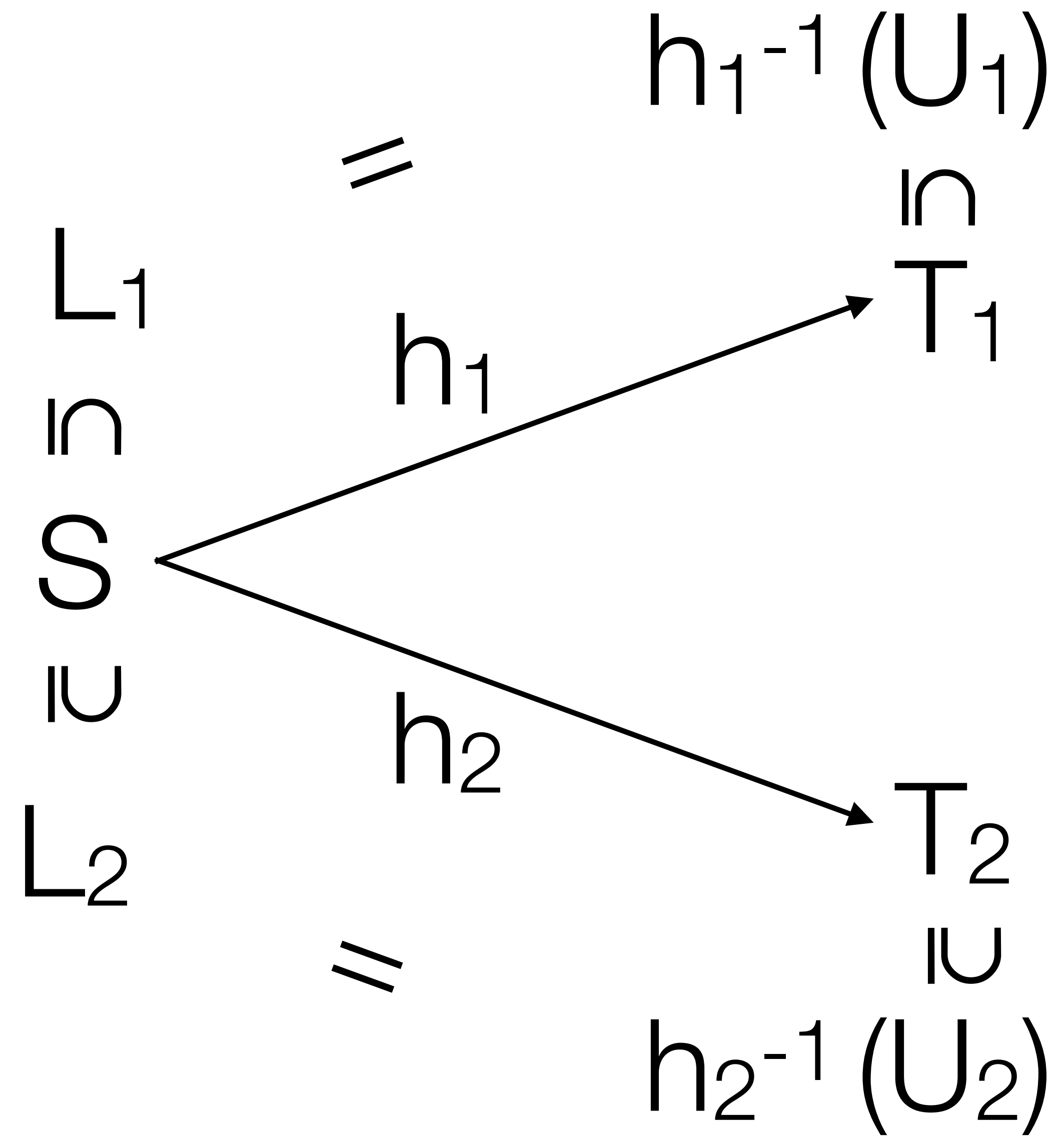
Properties of Recognisable Sets

Recognisable sets are closed under boolean combinations

Properties of Recognisable Sets

Recognisable sets are closed under boolean combinations

intersection



Properties of Recognisable Sets

Recognisable sets are closed under boolean combinations

intersection

$$L_1 = h_1^{-1}(U_1), \quad L_1 \subseteq S, \quad U_1 \subseteq T_1, \quad S \xrightarrow{h_1} T_1$$

$$L_2 = h_2^{-1}(U_2), \quad L_2 \subseteq S, \quad U_2 \subseteq T_2, \quad S \xrightarrow{h_2} T_2$$

$$(s_1,s_2) \otimes (t_1,t_2) \stackrel{\text{def}}{=} (s_1 t_1,\ s_2 t_2)$$

$$S \xrightarrow{h \stackrel{\text{def}}{=} (h_1,\ h_2)} (T_1 x T_2,\ (1_{T1}, 1_{T2}),\ \otimes)$$

$$L_1 \cap L_2 \subseteq S, \qquad h^{-1}(U_1 \text{ x } U_2) \subseteq T_1 x T_2$$

$$L_1 \cap L_2 \quad = \quad h^{-1}(U_1 \text{ x } U_2)$$

Properties of Recognisable Sets

Recognisable sets are closed under boolean combinations

complement

$$\begin{array}{ccc} S & \xrightarrow{\quad h \quad} & T \\ \rotatebox{90}{$\subseteq$} & & \rotatebox{90}{$\subseteq$} \\ L & = & h^{-1}(U) \end{array}$$

Properties of Recognisable Sets

Recognisable sets are closed under boolean combinations

complement

$$\begin{array}{ccc} S & \xrightarrow{\ h\ } & T \\ \cup\!| & & \cup\!| \\ L & = & h^{-1}(U) \end{array} \qquad \begin{array}{ccc} S & \xrightarrow{\ h\ } & T \\ \cup\!| & & \cup\!| \\ S\setminus L & = & h^{-1}(T\setminus U) \end{array}$$

Properties of Recognisable Sets

Recognisable sets are closed under inverse homomorphisms

$$\begin{array}{ccc} S & \xrightarrow{\quad h \quad} & T \\ \rotatebox{90}{$\subseteq$} & & \rotatebox{90}{$\subseteq$} \\ L & = & h^{-1}(U) \end{array}$$

Properties of Recognisable Sets

Recognisable sets are closed under inverse homomorphisms

$$S' \xrightarrow{\;g\;} S \xrightarrow{\;h\;} T$$

$$\begin{array}{ccccc} S' & \xrightarrow{g} & S & \xrightarrow{h} & T \\ \cup\!\!\shortmid & & \cup\!\!\shortmid & & \cup\!\!\shortmid \\ g^{-1}(L) = (h \circ g)^{-1}(U) & & L & = & h^{-1}(U) \end{array}$$

Logic in General

$\mathcal{R} = r_1, r_2, \ldots$ signature of relation symbols

Logic in General

$\mathcal{R} = r_1, r_2, ...$ signature of relation symbols

$\forall x$, $\exists x$	$\phi \wedge \psi$, $\phi \vee \psi$, $\neg \psi$	$r(x_1,\ldots,x_n)$	$x=y$
quantification over individuals	boolean operations	relations from the signature	equality

First Order Logic (FO)

Logic in General

$\mathcal{R} = r_1, r_2, \ldots$ signature of relation symbols

$\forall x, \exists x$ quantification over individuals

$\phi \wedge \psi, \phi \vee \psi, \neg \psi$ boolean operations

$r(x_1, \ldots, x_n)$ relations from the signature

$x = y$ equality

First Order Logic (FO)

$\forall X, \exists X$ quantification over sets

$x \in X$ membership

Monadic Second Order Logic (MSO)

Logic in General

$\mathcal{R} = r_1, r_2, \ldots$ signature of relation symbols

$S = (\mathcal{U}, r^S_1, r^S_2, \ldots)$ relational structure

universe

interpretation of relation symbols

Logic in General

$\mathcal{R} = r_1, r_2, \ldots$ signature of relation symbols

$\mathcal{S} = (\mathcal{U}, r^{\mathcal{S}}_1, r^{\mathcal{S}}_2, \ldots)$ relational structure

universe

interpretation of relation symbols

$\mathcal{S}, u_1, \ldots, u_n, \mathcal{U}_1, \ldots, \mathcal{U}_m \models \phi(x_1, \ldots, x_n, \mathcal{X}_1, \ldots \mathcal{X}_m)$ the formula ϕ is true in $\mathcal{S}$ when

- x_i is mapped to the element u_i
- $\mathcal{X}_j$ is mapped to the set $\mathcal{U}_j$

Logic for Words (the precedence model)

$\mathcal{L} = \{a, b, c, \ldots\}$

finite alphabet

$x \leq y$

position x is before position y

$a(x)$

position x has letter a

Logic for Words (the precedence model)

$\mathcal{E} = \{a,b,c,...\}$ finite alphabet

$x \leq y$

position x is before position y

$a(x)$

position x has letter a

$\mathcal{L}(\phi) = \{ w \in \mathcal{E} \mid w \models \phi \}$ language of a sentence

Logic for Words (the precedence model)

$\mathcal{E} = \{a, b, c, \ldots\}$ finite alphabet

$x \leq y$

position x is before position y

$a(x)$

position x has letter a

$\mathcal{L}(\phi) = \{ w \in \mathcal{E} \mid w \models \phi \}$ language of a sentence

$\mathcal{L} \subseteq \mathcal{E}^*$ is FO (MSO) definable iff $\mathcal{L} = \mathcal{L}(\phi)$ for ϕ a FO (MSO) sentence

Some FO-definable Languages

$$x{+}1{=}y \stackrel{\text{def}}{=} \underbrace{x \leqslant y \wedge \neg x{=}y}_{x<y} \wedge \forall z \,.\, x \leqslant z \wedge z \leqslant y \Rightarrow x{=}z \vee z{=}y$$

Some FO-definable Languages

$$x{+}1{=}y \stackrel{\text{def}}{=} \underbrace{x{\leqslant}y \land \neg x{=}y}_{x<y} \land \forall z\ .\ x{\leqslant}z \land z{\leqslant}y \Rightarrow x{=}z \lor z{=}y$$

a^*bc^* $\quad \exists x\ .\ b(x) \land \forall y\ .\ y<x \Rightarrow a(x) \land \forall y\ .\ x<y \Rightarrow c(x)$

Some FO-definable Languages

$$x{+}1{=}y \stackrel{\text{def}}{=} \underbrace{x \leq y \wedge \neg x{=}y}_{x<y} \wedge \forall z \,.\, x \leq z \wedge z \leq y \Rightarrow x{=}z \vee z{=}y$$

a^*bc^* $\qquad \exists x \,.\, b(x) \wedge \forall y \,.\, y<x \Rightarrow a(x) \wedge \forall y \,.\, x<y \Rightarrow c(x)$

$first(x) \stackrel{\text{def}}{=} \forall y \,.\, x \leq y \qquad last(x) \stackrel{\text{def}}{=} \forall y \,.\, y \leq x$

Some FO-definable Languages

$$x+1=y \stackrel{\text{def}}{=} \underbrace{x\leq y \land \neg x=y}_{x<y} \land \forall z \,.\, x\leq z \land z\leq y \Rightarrow x=z \lor z=y$$

a^*bc^* $\quad \exists x \,.\, b(x) \land \forall y \,.\, y<x \Rightarrow a(x) \land \forall y \,.\, x<y \Rightarrow c(x)$

$$first(x) \stackrel{\text{def}}{=} \forall y \,.\, x\leq y \qquad last(x) \stackrel{\text{def}}{=} \forall y \,.\, y\leq x$$

$(ab)^+$ $\quad \exists x \,.\, first(x) \land a(x) \land \forall y \,.\, a(y) \Rightarrow \neg last(y) \land b(y+1) \land \forall y \,.\, b(y) \land \neg last(y) \Rightarrow a(y+1)$

Some MSO-definable Languages

$$\mathcal{X}\subseteq\mathcal{Y} \stackrel{\text{def}}{=} \forall x \,.\, x\in\mathcal{X} \Rightarrow x\in\mathcal{Y}$$

$$\mathcal{X}\subseteq a \stackrel{\text{def}}{=} \forall x \,.\, x\in\mathcal{X} \Rightarrow a(x)$$

$$sing(\mathcal{X}) \stackrel{\text{def}}{=} \mathcal{X}\neq\emptyset \wedge \forall\mathcal{Y} \,.\, \mathcal{Y}\subseteq\mathcal{X} \Rightarrow \mathcal{Y}=\emptyset \vee \mathcal{Y}=\mathcal{X}$$

$$\mathcal{X}\leqslant\mathcal{Y} \stackrel{\text{def}}{=} \forall x\, \forall y \,.\, x\in\mathcal{X} \wedge y\in\mathcal{Y} \Rightarrow x\leqslant y$$

Some MSO-definable Languages

$(aa)^+$ all words of non-zero even length over the alphabet $\mathcal{L} = \{a\}$

$$\exists X \,.\, \forall x \,.\, first(x) \Rightarrow x{\in}X \wedge last(x) \Rightarrow x{\notin}X \wedge \forall y \,.\, x{=}y{+}1 \Rightarrow (x{\in}X \Leftrightarrow y{\notin}X)$$

all words having a on even positions over the alphabet $\mathcal{L} = \{a, b\}$

$$\exists X \,.\, \forall x \,.\, first(x) \Rightarrow x{\in}X \wedge \forall y \forall z \,.\, y{\in}X \wedge z{=}y{+}1 \Rightarrow z{\notin}X \wedge X{\subseteq}a$$

As we shall see later, these languages are not FO-definable ...

Logic for Words (the successor model)

$\mathcal{L} = \{a, b, c, ...\}$ finite alphabet

$x+1=y$ — position y is the successor of x

$a(x)$ — position x has letter a

Makes a difference for FO-definability but not for MSO-definability

$$x \leq y \stackrel{\text{def}}{=} \forall X \,.\, (x \in X \wedge \underbrace{(\forall y \forall z \,.\, y \in X \wedge y+1=z \Rightarrow z \in X)}_{\text{closed under successors}}) \Rightarrow y \in X$$

closed under successors

MSO-definability = recognisability for words

Theorem [Trakhtenbrot-Büchi-Elgot]

Let $\mathcal{E}$ be a finite alphabet and L ⊆ $\mathcal{E}^*$ be a language.

L is recognisable ⇔ L is MSO-definable.

"⇒" L is recognisable ⇒ L is the language of a finite automaton A with n states

$\exists X_1 \ldots \exists X_n$. *the sets* $X_1 \ldots X_n$ *describe an accepting run of* A

MSO-definability = recognisability for words

Theorem [Trakhtenbrot-Büchi-Elgot]

Let $\mathcal{E}$ be a finite alphabet and L ⊆ $\mathcal{E}^*$ be a language.

L is recognisable ⇔ L is MSO-definable.

"⇐" $X \subseteq Y$ $X \subseteq a$ $sing(X)$ $X \leqslant Y$

each MSO formula can be written as $\phi(X_1 \ldots X_n)$

MSO-definability = recognisability for words

Theorem [Trakhtenbrot-Büchi-Elgot]

Let $\mathcal{L}$ be a finite alphabet and L ⊆ $\mathcal{L}^*$ be a language.

L is recognisable ⇔ L is MSO-definable.

"⇐" $X \subseteq Y$ $X \subseteq a$ $sing(X)$ $X \leq Y$

each MSO formula can be written as $\phi(X_1 \dots X_n)$

the language of $\phi(X_1 \dots X_n)$ is the set of words w
over the alphabet $\mathcal{L} \times \{0,1\}^n$: $w, U_1, \dots, U_n \models \phi(X_1 \dots X_n)$
where $i \in U_j \Leftrightarrow$ the j-th bit on position i is 1

X_1	0	1	⋯	0
X_2	1	0	⋯	1
⋮	⋮	⋮		⋮
X_n	0	0	⋯	1
$\mathcal{L}$	a_1	a_2	...	a_m

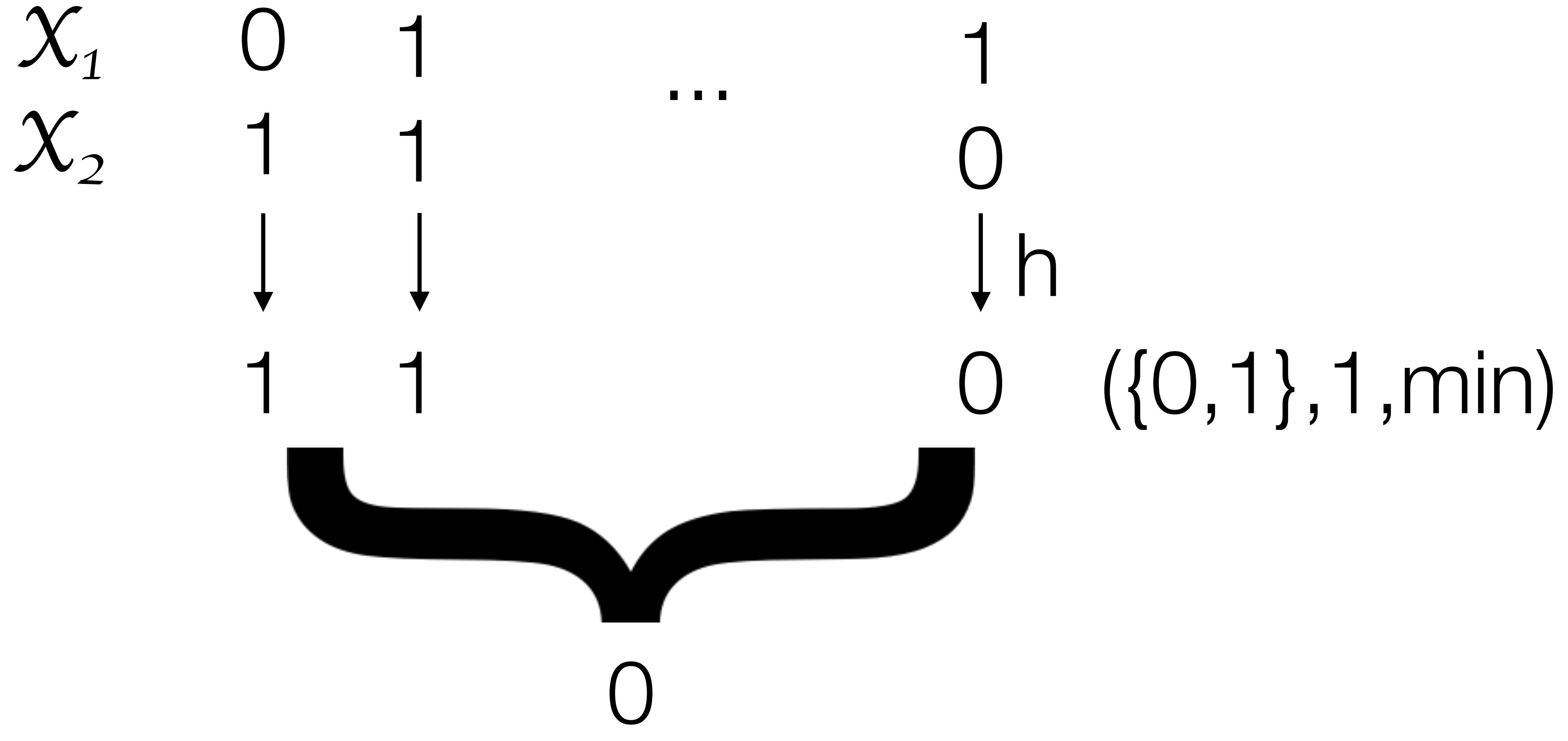
MSO-definability = recognisability for words

Theorem [Trakhtenbrot-Büchi-Elgot]

Let $\mathscr{E}$ be a finite alphabet and L ⊆ $\mathscr{E}^*$ be a language.

L is recognisable ⇔ L is MSO-definable.

"⇐" $X_1 \subseteq X_2$



$X_1 \subseteq a$

similar construction

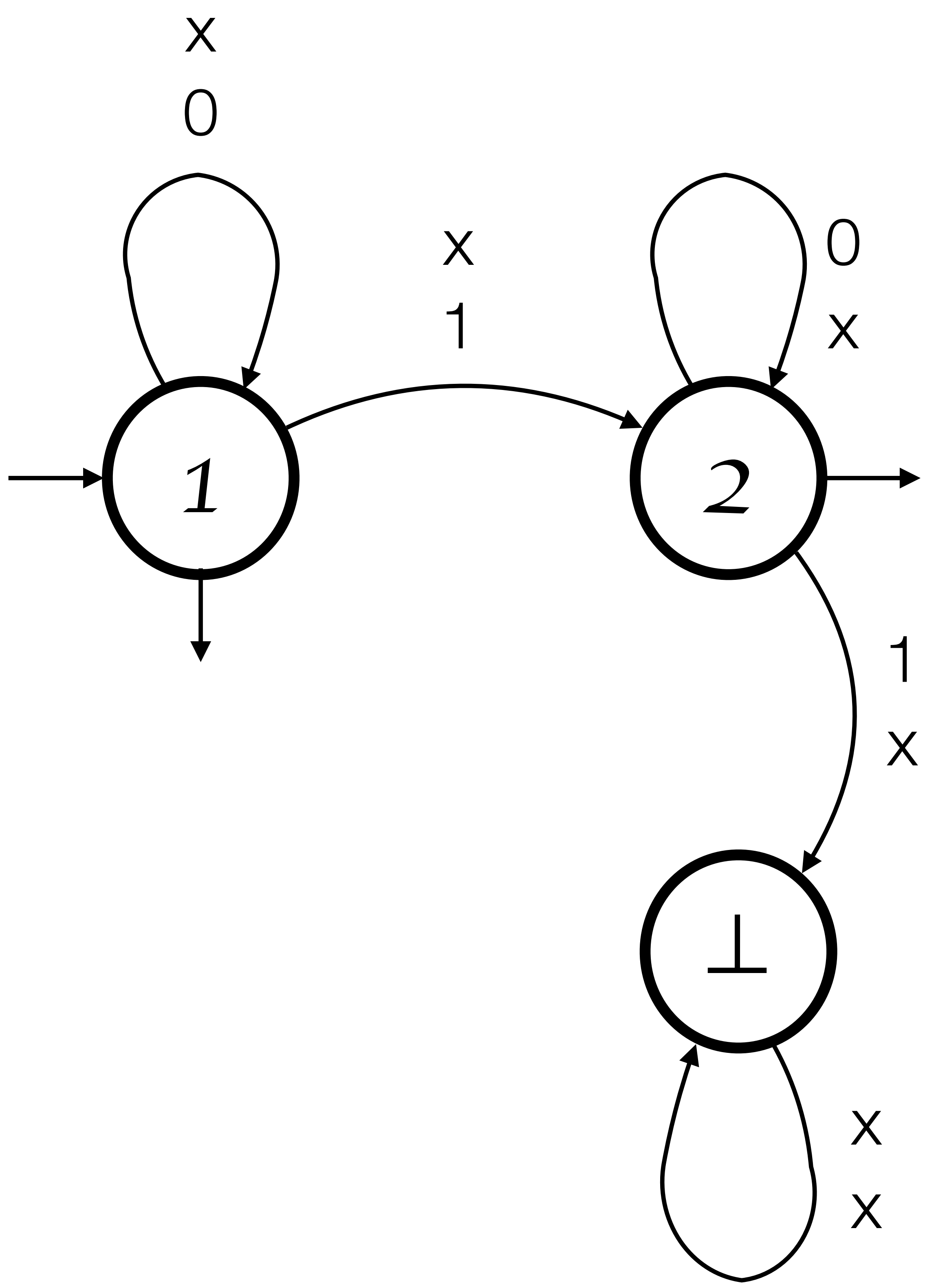
MSO-definability = recognisability for words

Theorem [Trakhtenbrot-Büchi-Elgot]

Let $\mathscr{L}$ be a finite alphabet and L ⊆ $\mathscr{L}^*$ be a language.

L is recognisable ⇔ L is MSO-definable.

"⇐" $\mathcal{X}_1 \leqslant \mathcal{X}_2$



MSO-definability = recognisability for words

Theorem [Trakhtenbrot-Büchi-Elgot]

Let $\mathcal{L}$ be a finite alphabet and L ⊆ $\mathcal{L}^*$ be a language.

L is recognisable ⇔ L is MSO-definable.

"⇐" $\phi \wedge \psi$, $\phi \vee \psi$, $\neg \psi$ closure of recognizable sets under boolean operations

$$(\mathcal{L} \times \{0,1\}^{n-1})^* \xleftarrow{\pi} (\mathcal{L} \times \{0,1\}^{n})^* \xrightarrow{h} M$$

$$\phi(X_1 \dots X_n) \subseteq (\mathcal{L} \times \{0,1\}^{n})^*, \quad N \subseteq M$$

$$\phi(X_1 \dots X_n) = h^{-1}(N)$$

$$(\mathcal{L} \times \{0,1\}^{n-1})^* \xrightarrow{H(w)=\{h(v) \mid \pi(v)=w\}} 2^M$$

$$\exists X_n . \phi(X_1 \dots X_n) \subseteq (\mathcal{L} \times \{0,1\}^{n-1})^*, \quad \{P \subseteq M \mid P \cap N \neq \varnothing\} \subseteq 2^M$$

$$\exists X_n . \phi(X_1 \dots X_n) = H^{-1}(\{P \subseteq M \mid P \cap N \neq \varnothing\})$$