Classifying Stress Patterns by Cognitive Complexity

James Rogers
Dept. of Computer Science
Earlham College
and
Dept. of Linguistics and Cognitive Science
University of Delaware

jrogers@cs.earlham.edu

Slide 1

http://cs.earlham.edu/~jrogers/slides/cornell.ho.pdf

Joint work with Jeff Heinz (UDel), Sean Wibel, Maggie Fero and Dakotah Lambert (EC)

Slide 2

Some simple patterns

(1) Primary stress falls on the final syllable

(2) Primary stress falls on the antepenultimate syllable

(3) In words of five or more syllables primary stress falls on the antepenultimate syllable

Some simple patterns

Slide 3

(4) Primary stress falls on the initial syllable if it is heavy, else the peninitial syllable.

(5) Primary stress falls on the leftmost heavy syllable

(6) Secondary stress falls on every third syllable counting from the antepenultimate syllable.

Some simple patterns

Slide 4

(7) Final syllable is heavy

(8) All heavy syllables get some stress

(9) There are always an odd number of heavy syllables

Some simple patterns

Slide 5

(10) Primary stress falls on some syllable. (At least one)

(11) Primary stress falls on at most one syllable.

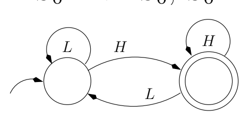
(12) Primary stress falls on exactly one syllable.

Complexity of Simple Patterns

Slide 6

(7) Sequences of 'L's and 'H's which end in 'H':

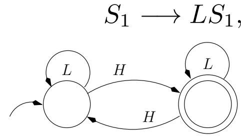
$$S_0 \longrightarrow LS_0,\ S_0 \longrightarrow HS_0,\ S_0 \longrightarrow H$$



$$(L + H)^* H$$

(9) Sequences of 'L's and 'H's which contain an odd number of 'H's:

$$S_0 \longrightarrow LS_0,\ S_0 \longrightarrow HS_1,$$
$$S_1 \longrightarrow LS_1,\ S_1 \longrightarrow HS_0,\ S_1 \longrightarrow \varepsilon$$



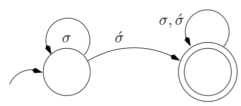
$$(L^* H L^* H L^*)^* L^* H L^*$$

Some More Simple Patterns

Slide 7

(10) Sequences of 'σ's and '$\acute{\sigma}$'s which contain at least one '$\acute{\sigma}$':

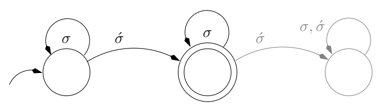
$$S_0 \longrightarrow \sigma\, S_0,\; S_0 \longrightarrow \acute{\sigma}\, S_1,$$
$$S_1 \longrightarrow \sigma\, S_1,\; S_1 \longrightarrow \acute{\sigma}\, S_1,\; S_1 \longrightarrow \varepsilon$$



$$\sigma^*\,\acute{\sigma}(\sigma+\acute{\sigma})^*$$

(12) Sequences of 'σ's and '$\acute{\sigma}$'s which contain exactly one '$\acute{\sigma}$':

$$S_0 \longrightarrow \sigma\, S_0,\; S_0 \longrightarrow \acute{\sigma}\, S_1,$$
$$S_1 \longrightarrow \sigma\, S_1,\; S_1 \longrightarrow \varepsilon$$



$$\sigma^*\,\acute{\sigma}\,\sigma^*$$

Cognitive Complexity from First Principles

Slide 8

What kinds of distinctions does a cognitive mechanism need to be sensitive to in order to classify an event with respect to a pattern?

Reasoning about patterns

- What objects/entities/things are we reasoning about?
- What relationships between them are we reasoning with?

Some Assumptions about Linguistic Behaviors

Slide 9

- Perceive/process/generate linear sequence of (sub)events
- Can model as strings—linear sequence of abstract symbols
  - Discrete linear order (initial segment of $\mathbb{N}$).
  - Labeled with alphabet of events

    Partitioned into subsets, each the set of positions at which some event occurs.

Word models

Slide 10

$$\begin{tikzcd}[column sep=small] & \langle \mathcal{D}, \triangleleft, \triangleleft^+, P_\sigma\rangle_{\sigma\in\Sigma} \arrow[dl] \arrow[dr] & \\ (+1)\quad \langle \mathcal{D}, \triangleleft, P_\sigma\rangle_{\sigma\in\Sigma} & & (<)\quad \langle \mathcal{D}, \triangleleft^+, P_\sigma\rangle_{\sigma\in\Sigma}\end{tikzcd}$$

$$\begin{array}{rcl} \mathcal{D} & — & \text{Finite} \\ \triangleleft^+ & — & \text{Linear order on } \mathcal{D} \\ \triangleleft & — & \text{Successor wrt } \triangleleft^+ \\ P_\sigma & — & \text{Subset of } \mathcal{D} \text{ at which } \sigma \text{ occurs} \\ & & (P_\sigma \text{ partition } \mathcal{D}) \end{array}$$

$$\begin{array}{rcccccl} CCVC = \langle & \{0,1,2,3\}, & \{\langle i, i+1\rangle \mid 0 \leq i < 3\}, & \{0,1,3\}_C, & \{2\}_V & \rangle \\ \langle & \mathcal{D} & \triangleleft & P_C & P_V & \rangle \end{array}$$

Local Constraints

Slide 11

- Blocks of adjacent symbols
  - k-factors
- End markers: '$\rtimes$', '$\ltimes$'

$$\begin{aligned} F_2(\rtimes\,\sigma\,\sigma\,\acute{\sigma}\,\ltimes) &= \{\rtimes\,\sigma,\ \sigma\,\sigma,\ \sigma\,\acute{\sigma},\ \acute{\sigma}\,\ltimes\} \\ F_3(\rtimes\,\sigma\,\sigma\,\acute{\sigma}\,\ltimes) &= \{\rtimes\,\sigma\,\sigma,\ \sigma\,\sigma\,\acute{\sigma},\ \sigma\,\acute{\sigma}\,\ltimes\} \\ F_6(\rtimes\,\sigma\,\sigma\,\acute{\sigma}\,\ltimes) &= \{\rtimes\,\sigma\,\sigma\,\acute{\sigma}\,\ltimes\} \end{aligned}$$

Strictly k-Local Constraints

Slide 12

- Co-occurrence of negative atomic local constraints
  - Conjunctions of negated k-factors

(1) Primary stress falls on the final syllable

$$\neg\,\sigma\,\ltimes \qquad (\mathrm{SL}_2)$$

(2) Primary stress falls on the antepenultimate syllable

$$\neg\,\sigma\,\sigma\,\sigma\,\ltimes\ \wedge\ \neg\,\acute{\sigma}\,\sigma\,\ltimes\ \wedge\ \neg\,\acute{\sigma}\,\ltimes \qquad (\mathrm{SL}_4)$$

Cambodian

Slide 13

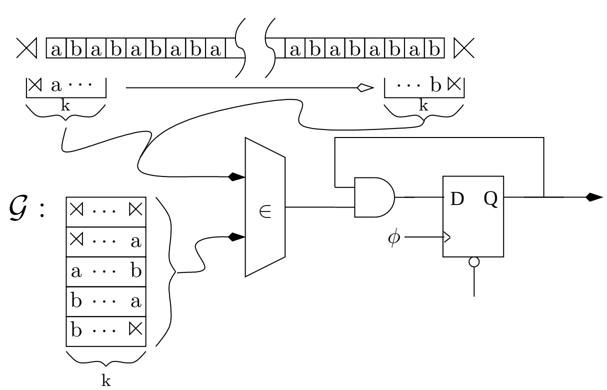
1) In words of all sizes, primary stress falls on the final syllable. $\neg\,\sigma \ltimes \;\wedge\; \neg\,\grave{\sigma} \ltimes$ (SL_2)
2) In words of all sizes, secondary stress falls on all heavy syllables. $\neg\, H$ (SL_1)
3) Light syllables occur only immediately following heavy syllables. $\neg \rtimes \overset{*}{L} \;\wedge\; \neg\, \overset{*}{L}\,\overset{*}{L}$ (SL_2)

[4) Light monosyllables do not occur. $\neg \rtimes \acute{L} \ltimes$ (SL_3)]

Cambodian stress is SL_2.

Scanners

Slide 14



Recognizing an SL_k stringset requires only remembering the k most recently encountered symbols.

Character of Strictly k-Local Sets

Slide 15

Theorem (Suffix Substitution Closure):
A stringset L is strictly k-local iff whenever there is a string x of length $k-1$ and strings w, y, v, and z, such that

$$\begin{array}{ccccccc} w & \cdot & \overbrace{x}^{k-1} & \cdot & y & \in L \\ v & \cdot & x & \cdot & z & \in L \end{array}$$

then it will also be the case that

$$w \cdot x \cdot z \in L$$

$\star\, CCC$ is SL_3

$$\begin{array}{ccccc|c} V & \cdot & CC & \cdot & VC & \in \star\, CCC \\ CV & \cdot & CC & \cdot & V & \in \star\, CCC \\ \hline V & \cdot & CC & \cdot & V & \in \star\, CCC \end{array}$$

But $\star\, CCC$ is not SL_2:

$$\begin{array}{ccccc|c} C & \cdot & C & \cdot & VC & \in \star\, CCC \\ V & \cdot & C & \cdot & CV & \in \star\, CCC \\ \hline C & \cdot & C & \cdot & CV & \notin \star\, CCC \end{array}$$

Alawa

Slide 16

- In words of all sizes, primary stress falls on the penultimate syllable.
- [—Except in monosyllables]

$$\mathcal{G}_{\text{Alawa}} = \{ \quad \rtimes\sigma\sigma,\ \rtimes\sigma\acute{\sigma},\ \rtimes\acute{\sigma}\sigma,\ \sigma\sigma\sigma,\ \sigma\sigma\acute{\sigma},\ \sigma\acute{\sigma}\sigma,\ \acute{\sigma}\sigma\ltimes,\ \rtimes\acute{\sigma}\ltimes \quad \}$$

$$\begin{array}{r|c|l} \rtimes\sigma & \acute{\sigma} & \sigma\ltimes \\ \rtimes & \acute{\sigma} & \ltimes \\ \hline \star\ \rtimes\sigma & \acute{\sigma} & \ltimes \end{array} \qquad \begin{array}{r|c|l} \rtimes\sigma & \sigma & \acute{\sigma}\sigma\ltimes \\ \rtimes\acute{\sigma} & \sigma & \ltimes \\ \hline \star\ \rtimes\sigma & \sigma & \ltimes \end{array}$$

Alawa stress is in $\mathrm{SL}_3 - \mathrm{SL}_2$.

SL Hierarchy

Theorem 1 (SL-Hierarchy)

$$SL_1 \subsetneq SL_2 \subsetneq SL_3 \subsetneq \cdots \subsetneq SL_i \subsetneq SL_{i+1} \subsetneq \cdots \subsetneq SL$$

Slide 17

Every Finite stringset is SL_k for some k: $\mathrm{Fin} \subseteq \mathrm{SL}$.

There is no k for which SL_k includes all Finite languages.

SL_k is learnable in the limit from positive data.

SL is not.

Cognitive interpretation of SL

Slide 18

- Any cognitive mechanism that can distinguish member strings from non-members of a (properly) SL_k language must be sensitive, at least, to the length k blocks of consecutive events that occur in the presentation of the string.
- If the strings are presented as sequences of events in time, then this corresponds to being sensitive, at each point in the string, to the immediately prior sequence of $k-1$ events.
- Any cognitive mechanism that is sensitive *only* to the length k blocks of consecutive events in the presentation of a string will be able to recognise *only* SL_k languages.

Strictly Local Stress Patterns

Slide 19

Heinz's Stress Pattern Database (ca. 2007)—109 patterns

9 are SL_2	Abun West, Afrikans, ... Cambodian,... Maranungku
44 are SL_3	Alawa, Arabic (Bani-Hassan),...
24 are SL_4	Dutch,...
3 are SL_5	Asheninca, Bhojpuri, Hindi (Fairbanks)
1 is SL_6	Icua Tupi
28 are not SL	Amele, Bhojpuri (Shukla Tiwari), Arabic (Classical), Hindi (Kelkar), Yidin,...

72% are SL, all $k \leq 6$. 49% are SL_3.

Some-$\acute{\sigma}$

Slide 20

$$\begin{array}{r|c|l} \rtimes\, \sigma & \overbrace{\sigma \cdots \sigma}^{k-1} & \grave{\sigma} \ltimes \\ \rtimes\, \acute{\sigma} & \overbrace{\sigma \cdots \sigma}^{k-1} & \sigma \ltimes \\ \hline \star\; \sigma & \overbrace{\sigma \cdots \sigma}^{k-1} & \sigma \ltimes \end{array}$$

Some-$\acute{\sigma} \notin$ SL

How can any stress pattern be SL?

Locally definable stringsets

$$\begin{array}{ccccc} f \in F_k(\rtimes \cdot \Sigma^* \cdot \ltimes) & & w \models f & \stackrel{\text{def}}{\iff} & f \in F_k(\rtimes \cdot w \cdot \ltimes) \\ \varphi \wedge \psi & & w \models \varphi \wedge \psi & \stackrel{\text{def}}{\iff} & w \models \varphi \text{ and } w \models \psi \\ \neg\varphi & & w \models \neg\varphi & \stackrel{\text{def}}{\iff} & w \not\models \varphi \end{array}$$

Slide 21

$$\varphi \vee \psi \;\equiv\; \neg(\varphi \wedge \psi)$$

$$L = L(\varphi) \stackrel{\text{def}}{=} \{w \in \Sigma^* \mid w \models \varphi\}$$

$$\mathrm{SL}_k \equiv \bigwedge_{f_i \notin \mathcal{G}} [\neg f_i] \subsetneq \mathrm{LT}_k$$

Some-$\boldsymbol{\acute{\sigma}}$ again

Slide 22

$$\text{Some-}\acute{\sigma} = L(\acute{\sigma})$$

$$\text{Some-}\acute{\sigma} \in \mathrm{LT}_1$$

Slide 23

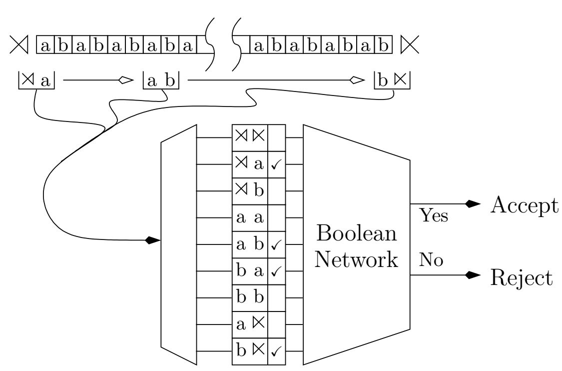
NKL

- Primary stress falls on the final syllable if it is Heavy
- Else on the initial syllable if it is Light
- Else on the penultimate syllable

$$
\begin{array}{ll}
\varphi_{\mathrm{NKL}} = & \\
\quad \acute{H} \ltimes & \text{final syllable if it is Heavy} \\
\quad \vee\, (\neg\, \acute{H} \ltimes \wedge \rtimes \acute{L}) & \text{Else on the initial if it is Light} \\
\quad \vee\, (\neg\, \acute{H} \ltimes \wedge \neg{\rtimes}\, \acute{L} \wedge \acute{\sigma}\, \overset{*}{\sigma} \ltimes) & \text{Else on the penultimate syllable}
\end{array}
$$

Slide 24

LT Automata



Membership in an LT_k stringset depends only on the set of k-Factors which occur in the string.

Recognizing an LT_k stringset requires only remembering which k-factors occur in the string.

Character of Locally Testable sets

Theorem 2 (k-Test Invariance) *A stringset L is Locally Testable iff*

there is some k such that, for all strings x and y,

Slide 25

if $\rtimes \cdot x \cdot \ltimes$ and $\rtimes \cdot y \cdot \ltimes$ have exactly the same set of k-factors

then either both x and y are members of L or neither is.

Definition 1 (k-Local Equivalence)

$$w \equiv_k^L v \stackrel{def}{\Longleftrightarrow} F_k(\rtimes w \ltimes) = F_k(\rtimes v \ltimes).$$

LT Hierarchy

Theorem 3 (LT-Hierarchy)

$$LT_1 \subsetneq LT_2 \subsetneq LT_3 \subsetneq \cdots \subsetneq LT_i \subsetneq LT_{i+1} \subsetneq \cdots \subsetneq LT$$

Slide 26

$$\mathrm{SL}_k \subseteq \mathrm{LT}_k$$

$$\mathrm{LT}_k \subseteq \mathrm{LT}_{k+1}$$

$$\mathrm{LT}_k \not\subseteq \mathrm{SL}_{k+1}$$

$$\mathrm{SL}_{k+1} \not\subseteq \mathrm{LT}k$$

Cognitive interpretation of LT

Slide 27

- Any cognitive mechanism that can distinguish member strings from non-members of a (properly) LT_k language must be sensitive, at least, to the *set* of length k contiguous blocks of events that occur in the presentation of the string—both those that do occur and those that do not.
- If the strings are presented as sequences of events in time, then this corresponds to being sensitive, at each point in the string, to the set of length k blocks of events that occurred at any prior point.
- Any cognitive mechanism that is sensitive *only* to the occurrence or non-occurrence of length k contiguous blocks of events in the presentation of a string will be able to recognise *only* LT_k languages.

Murik

Slide 28

- Primary stress falls on the leftmost heavy syllable
- else the initial syllable

Murik

- Primary stress falls on the leftmost heavy syllable
- else the initial syllable
- No more than one heavy syllable occurs in any word

Slide 29

Murik

- Primary stress falls on the leftmost heavy syllable
- else the initial syllable
- No more than one heavy syllable occurs in any word

Slide 30

$$L_{\text{Murik}} = \neg H \ \wedge (\acute{H} \vee \rtimes \acute{\sigma}) \ \wedge \cdots$$

Murik

- Primary stress falls on the leftmost heavy syllable
- else the initial syllable
- No more than one heavy syllable occurs in any word

Slide 31

$$\rtimes \overbrace{L \cdots L}^{k-1} \acute{H} \overbrace{L \cdots L}^{k-1} \ltimes \qquad \in L_{\mathrm{Murik}}$$

$$\rtimes \overbrace{L \cdots L}^{k-1} \acute{H} \overbrace{L \cdots L}^{k-1} \acute{H} \overbrace{L \cdots L}^{k-1} \ltimes \quad \notin L_{\mathrm{Murik}}$$

$$\mathrm{F}_k(\rtimes \overbrace{L \cdots L}^{k-1} \acute{H} \overbrace{L \cdots L}^{k-1} \ltimes) = \mathrm{F}_k(\rtimes \overbrace{L \cdots L}^{k-1} \acute{H} \overbrace{L \cdots L}^{k-1} \acute{H} \overbrace{L \cdots L}^{k-1} \ltimes)$$
$$= \{\rtimes \overbrace{L \cdots L}^{k-1},\ \overbrace{L \cdots L}^{k-1} \acute{H},\ \ldots,\ \acute{H} \overbrace{L \cdots L}^{k-1},\ \overbrace{L \cdots L}^{k-1} \ltimes\}$$

(no-more-than) One-$\acute{\sigma}$

Slide 32

$$\rtimes \overbrace{\sigma \cdots \sigma}^{k-1} \acute{\sigma} \overbrace{\sigma \cdots \sigma}^{k-1} \ltimes \qquad \in L_{\mathrm{One}-\acute{\sigma}}$$

$$\rtimes \overbrace{\sigma \cdots \sigma}^{k-1} \acute{\sigma} \overbrace{\sigma \cdots \sigma}^{k-1} \acute{\sigma} \overbrace{\sigma \cdots \sigma}^{k-1} \ltimes \quad \notin L_{\mathrm{One}-\acute{\sigma}}$$

One-$\acute{\sigma}$ is not LT (hence not SL)

Slide 33

FO(+1)

Models: $\langle \mathcal{D}, \triangleleft, P_\sigma \rangle_{\sigma\in\Sigma}$

First-order Quantification (over positions in the strings)

$$
\begin{array}{lrcl}
x \triangleleft y & w, [x \mapsto i, y \mapsto j] \models x \triangleleft y & \stackrel{\text{def}}{\iff} & j = i + 1 \\
P_\sigma(x) & w, [x \mapsto i] \models P_\sigma(x) & \stackrel{\text{def}}{\iff} & i \in P_\sigma \\
\varphi \wedge \psi & \vdots & & \\
\neg\varphi & \vdots & & \\
(\exists x)[\varphi(x)] & w, s \models (\exists x)[\varphi(x)] & \stackrel{\text{def}}{\iff} & w, s[x \mapsto i] \models \varphi(x)] \\
 & & & \text{for some } i \in \mathcal{D}
\end{array}
$$

FO(+1)-Definable Stringsets: $L(\varphi) \stackrel{\text{def}}{=} \{w \mid w \models \varphi\}$.

$$\varphi_{\text{One-}\acute{\sigma}} = (\exists x)[\acute{\sigma}(x) \wedge (\forall y)[\acute{\sigma}(y) \to x \approx y]$$

Slide 34

Character of the FO(+1) Definable Stringsets

Definition 2 (Locally Threshold Testable) *A set L is* Locally Threshold Testable *(LTT) iff there is some k and t such that, for all $w, v \in \Sigma^*$:*

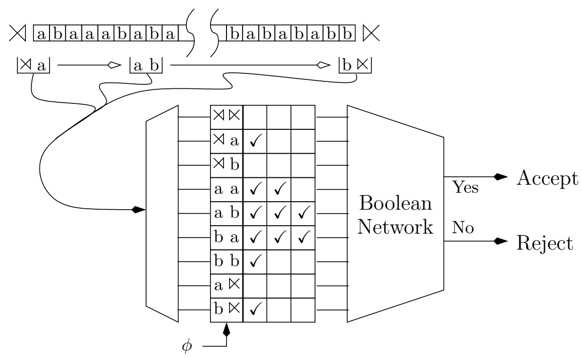
if for all $f \in F_k(\rtimes \cdot w \cdot \ltimes) \cup F_k(\rtimes \cdot v \cdot \ltimes)$
either $|w|_f = |v|_f$ or both $|w|_f \geq t$ and $|v|_f \geq t$,

then $w \in L \iff v \in L$.

Theorem 4 (Thomas) *A set of strings is First-order definable over $\langle \mathcal{D}, \triangleleft, P_\sigma \rangle_{\sigma\in\Sigma}$ iff it is* Locally Threshold Testable.

Slide 35

LTT Automata



Membership in an FO(+1) definable stringset depends only on the multiplicity of the k-factors, up to some fixed finite threshold, which occur in the string.

Slide 36

Cognitive interpretation of FO(+1)

- Any cognitive mechanism that can distinguish member strings from non-members of a (properly) FO(+1) stringset must be sensitive, at least, to the multiplicity of the length k blocks of events, for some fixed k, that occur in the presentation of the string, distinguishing multiplicities only up to some fixed threshold t.
- If the strings are presented as sequences of events in time, then this corresponds to being able count up to some fixed threshold.
- Any cognitive mechanism that is sensitive *only* to the multiplicity, up to some fixed threshold, (and, in particular, not to the order) of the length k blocks of events in the presentation of a string will be able to recognize *only* FO(+1) stringsets.

No-H-before-$\acute{H}$

- Primary stress falls on the leftmost heavy syllable
- (Murik), Maori, Yidin, Kashmiri, ...

Slide 37

$$\star\ H \ldots \acute{H}$$

$$\rtimes \overbrace{\grave{L}\,L \cdots \grave{L}\,L}^{2kt}\ \acute{H}\,H\ \overbrace{\grave{L}\,L \cdots \grave{L}\,L}^{2kt}\ \grave{H}\,H\ \overbrace{\grave{L}\,L \cdots \grave{L}\,L}^{2kt} \ltimes$$

$$\equiv^{L}_{k,t}$$

$$\star \rtimes \underbrace{\grave{L}\,L \cdots \grave{L}\,L}_{2kt}\ \grave{H}\,H\ \underbrace{\grave{L}\,L \cdots \grave{L}\,L}_{2kt}\ \acute{H}\,H\ \underbrace{\grave{L}\,L \cdots \grave{L}\,L}_{2kt} \ltimes$$

First-Order($<$) definable stringsets

$\langle \mathcal{D}, \triangleleft^{+}, P_\sigma \rangle_{\sigma \in \Sigma}$

First-order Quantification over positions in the strings

Slide 38

$$\begin{array}{ccrcl}
x \triangleleft^{+} y & & w, [x \mapsto i, y \mapsto j] \models x \triangleleft^{+} y & \stackrel{\text{def}}{\iff} & i < j \\
P_\sigma(x) & & w, [x \mapsto i] \models P_\sigma(x) & \stackrel{\text{def}}{\iff} & i \in P_\sigma \\
\varphi \wedge \psi & & \vdots & & \\
\neg\varphi & & \vdots & & \\
(\exists x)[\varphi(x)] & & w, s \models (\exists x)[\varphi(x)] & \stackrel{\text{def}}{\iff} & w, s[x \mapsto i] \models \varphi(x) \\
 & & & & \text{for some } i \in \mathcal{D}
\end{array}$$

$$\varphi_{\text{no-}H\text{-before-}\acute{H}} = \neg(\exists x, y)[x \triangleleft^{+} y \wedge H(x) \wedge \acute{H}(y)]$$

Star-Free stringsets

Definition 3 (Star-Free Set) *The class of* Star-Free Sets *(SF) is the smallest class of languages satisfying:*

Slide 39

- *Fin* $\subseteq$ *SF.*
- *If* $L_1, L_2 \in SF$ *then:* $L_1 \cdot L_2 \in SF$,
  $L_1 \cup L_2 \in SF$,
  $\overline{L_1} \in SF$.

Theorem 5 (McNauthton and Papert) *A set of strings is First-order definable over* $\langle \mathcal{D}, \triangleleft^{+}, P_\sigma \rangle_{\sigma \in \Sigma}$ *iff it is* Star-Free.

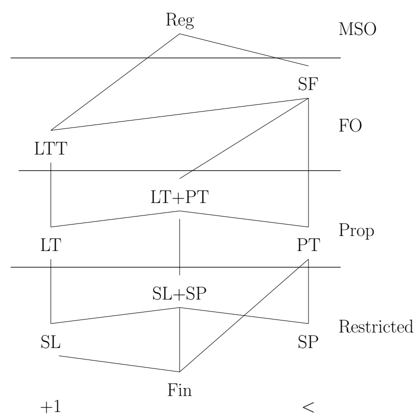
Cognitive interpretation of SF (FO(<))

Slide 40

- Any cognitive mechanism that can distinguish member strings from non-members of a (properly) SF language must be sensitive, at least, to both the order and the multiplicity of the length k blocks of events, for some fixed k, that occur in the presentation of the string, distinguishing multiplicities only up to some fixed threshold t.
- If the strings are presented as sequences of events in time, then this corresponds to being able not only to count events up to some threshold but also to track the sequence in which those events occur.
- Any cognitive mechanism that is sensitive *only* to the order and the multiplicity of the length k blocks of events, for some fixed k, that occur in the presentation of the string, distinguishing multiplicities only up to some fixed threshold t will be able to recognise *only* SF languages.

Slide 41

Sub-Regular Hierarchies



Slide 42

Yidin

- Primary stress on the leftmost heavy syllable, else the initial syllable
- Secondary stress iteratively on every second syllable in both directions from primary stress
- No light monosyllables

Explicitly:

- Exactly one $\acute{\sigma}$ (One-$\acute{\sigma}$)
- $\acute{L}$ implies no H (No-H -with-$\acute{L}$)
- σ and $\overset{+}{\sigma}$ alternate (Alt)
- First H gets primary stress (No-H -before-$\acute{H}$)
- $\acute{L}$ only if initial (Nothing-before-$\acute{L}$)
- No $\acute{L}$ monosyllables (No $\rtimes \acute{L} \ltimes$)

Classifying Conjunctive Constraints

Slide 43

- One-$\acute{\sigma}$ $\quad (\exists! x)[\acute{\sigma}(x)] \quad$ ($\mathrm{LTT}_{1,2}$)
- No-H -before-$\acute{H}$ $\quad \neg(\exists x, y)[x \triangleleft^{+} y \wedge H(x) \wedge \acute{H}(y)] \quad$ (SF)
- No-H -with-$\acute{L}$ $\quad \neg(H \wedge \acute{L}) \quad$ (LT_1)
- Nothing-before-$\acute{L}$ $\quad \neg\, \sigma\, \acute{L} \quad$ (SL_2)
- Alt $\quad \neg\, \sigma\, \sigma \wedge \neg\, \acute{\sigma}\, \acute{\sigma} \wedge \neg\, \acute{\sigma}\, \grave{\sigma} \wedge \neg\, \grave{\sigma}\, \acute{\sigma} \wedge \neg\, \grave{\sigma}\, \grave{\sigma} \quad$ (SL_2)
- No $\rtimes \acute{L} \ltimes$ $\quad \neg \rtimes \acute{L} \ltimes \quad$ (SL_3)

Yidin is SF

Precedence—Subsequences

Slide 44

Definition 4 (Subsequences)

$$v \sqsubseteq w \overset{def}{\Longleftrightarrow} v = \sigma_1 \cdots \sigma_n \ \ and \ w \in \Sigma^* \cdot \sigma_1 \cdot \Sigma^* \cdots \Sigma^* \cdot \sigma_n \cdot \Sigma^*$$

$$P_k(w) \overset{def}{=} \{v \in \Sigma^k \mid v \sqsubseteq w\}$$

$$P_{\leq k}(w) \overset{def}{=} \{v \in \Sigma^{\leq k} \mid v \sqsubseteq w\}$$

$$\sigma \ \ \sigma \ \ \acute{\sigma} \ \ \sigma \ \ \grave{\sigma} \ \ \sigma$$

$\sigma\ \sigma$, $\sigma\ \acute{\sigma}$, $\acute{\sigma}\sigma$, $\sigma\ \grave{\sigma}$, $\grave{\sigma}\sigma$
$\sigma\ \acute{\sigma}$, $\sigma\ \sigma$, $\acute{\sigma}\grave{\sigma}$
$\sigma\ \sigma$, $\sigma\ \grave{\sigma}$, $\acute{\sigma}\sigma$
$\sigma\ \grave{\sigma}$, $\sigma\ \sigma$
$\sigma\ \sigma$

$$P_2(\sigma\, \sigma\, \acute{\sigma}\, \sigma\, \grave{\sigma}\, \sigma) = \{\sigma\, \sigma, \sigma\, \acute{\sigma}, \sigma\, \grave{\sigma}, \acute{\sigma}\, \sigma, \acute{\sigma}\, \grave{\sigma}, \grave{\sigma}\, \sigma\}$$

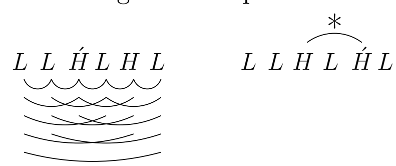
$$P_{\leq 2}(\sigma\, \sigma\, \acute{\sigma}\, \sigma\, \grave{\sigma}\, \sigma) = \{\varepsilon, \sigma, \acute{\sigma}, \grave{\sigma}, \sigma\, \sigma, \sigma\, \acute{\sigma}, \sigma\, \grave{\sigma}, \acute{\sigma}\, \sigma, \acute{\sigma}\, \grave{\sigma}, \grave{\sigma}\, \sigma\}$$

Strictly Piecewise Stringsets—SP

Strictly k-Piecewise Definitions

- Co-occurrence of negative atomic piecewise constraints
  - Conjunctions of negated k-sequences

Slide 45



Membership in an SP_k stringset depends only on the individual ($\leq k$)-subsequences which do and do not occur in the string.

Character of the Strictly k-Piecewise Sets

Theorem 6 *A stringset L is Strictly k-Piecewise Testable iff it is closed under subsequence:*

$$w\sigma v \in L \Rightarrow wv \in L$$

Slide 46

Every naturally occurring stress pattern requires Primary Stress

$\Rightarrow$

No naturally occurring stress pattern is SP.

But SP can forbid multiple primary stress: $\neg\,\acute{\sigma}\,\acute{\sigma}$

Yidin constraints wrt SP

Slide 47

• One-$\acute{\sigma}$ is not SP	$\star\ \sigma\,\sigma\ \sqsubseteq\ \sigma\,\acute{\sigma}\,\sigma$
• No-H -before-$\acute{H}$ is SP_2	$\neg\, H\,\acute{H}$
• No-H -with-$\acute{L}$ is SP_2	$\neg\, H\,\acute{L} \wedge \neg\, \acute{L}\,H$
• Nothing-before-$\acute{L}$ is SP_2	$\neg \sigma\,\acute{L}$
• Alt is not SP	$\star\ \sigma\,\sigma\,\acute{\sigma}\ \sqsubseteq\ \sigma\,\grave{\sigma}\,\sigma\,\acute{\sigma}$
• No $\rtimes\,\acute{L}\,\ltimes$ is not SP	$\star\ \acute{L}\ \sqsubseteq \acute{L}\,L$

Cognitive interpretation of SP

Slide 48

- Any cognitive mechanism that can distinguish member strings from non-members of a (properly) SP_k stringset must be sensitive, at least, to the length k (not necessarily consecutive) sequences of events that occur in the presentation of the string.
- If the strings are presented as sequences of events in time, then this corresponds to being sensitive, at each point in the string, to up to $k-1$ events distributed arbitrarily among the prior events.
- Any cognitive mechanism that is sensitive *only* to the length k sequences of events in the presentation of a string will be able to recognize *only* SP_k stringsets.

k-Piecewise Testable Stringsets

PT$_k$-expressions

$$\begin{array}{ccc} p \in \Sigma^{\leq k} & w \models p \stackrel{\text{def}}{\iff} & p \sqsubseteq w \\ \varphi \wedge \psi & w \models \varphi \wedge \psi \stackrel{\text{def}}{\iff} & w \models \varphi \text{ and } w \models \psi \\ \neg\varphi & w \models \neg\varphi \stackrel{\text{def}}{\iff} & w \not\models \varphi \end{array}$$

Slide 49

k-Piecewise Testable Languages (PT$_k$):

$$L(\varphi) \stackrel{\text{def}}{=} \{w \in \Sigma^* \mid w \models \varphi\}$$

$$\text{One-}\acute{\sigma} = L(\acute{\sigma} \wedge \neg\, \acute{\sigma}\, \acute{\sigma})$$

Membership in an PT$_k$ stringset depends only on the set of ($\leq k$)-subsequences which occur in the string.

SP$_k$ is equivalent to $\bigwedge_{p_i \notin \mathcal{G}}[\neg p_i]$

Character of Piecewise Testable sets

Slide 50

Theorem 7 (k-Subsequence Invariance) *A stringset L is Piecewise Testable iff*

there is some k such that, for all strings x and y,

if x and y have exactly the same set of ($\leq k$)-subsequences

then either both x and y are members of L or neither is.

$$w \equiv_k^P v \stackrel{\text{def}}{\iff} P_{\leq k}(w) = P_{\leq k}(v).$$

Slide 51

Yidin constraints wrt PT

- One-$\acute{\sigma}$ is PT_2 — $\acute{\sigma} \wedge \neg\, \acute{\sigma}\, \acute{\sigma}$
- No-H -before-$\acute{H}$ is SP_2 — $\neg\, H\, \acute{H}$
- No-H -with-$\acute{L}$ is SP_2 — $\neg\, H\, \acute{L} \wedge \neg\, \acute{L}\, H$
- Nothing-before-$\acute{L}$ is SP_2 — $\neg \sigma\, \acute{L}$
- Alt is not PT — $\star\ \overbrace{\sigma\, \grave{\sigma} \cdots \sigma\, \grave{\sigma}}^{2k} \equiv^{P}_{k} \overbrace{\sigma\, \grave{\sigma} \cdots \sigma\, \grave{\sigma}}^{2k} \grave{\sigma}$
- No $\rtimes \acute{L} \ltimes$ is PT_2 — $\acute{L} \rightarrow (\sigma\, \acute{L} \vee \acute{L}\, \sigma)$

Slide 52

Cognitive interpretation of PT

- Any cognitive mechanism that can distinguish member strings from non-members of a (properly) PT_k stringset must be sensitive, at least, to the set of length k subsequences of events that occur in the presentation of the string—both those that do occur and those that do not.
- If the strings are presented as sequences of events in time, then this corresponds to being sensitive, at each point in the string, to the set of all length k subsequences of the sequence of prior events.
- Any cognitive mechanism that is sensitive *only* to the set of length k subsequences of events in the presentation of a string will be able to recognize *only* PT_k stringsets.

Slide 53

Yidin wrt Local and Piecewise Constraints

One-$\acute{\sigma}$	$\mathrm{LTT}_{1,2}$	PT_2
Some-$\acute{\sigma}$	LT_1	PT_1
At-Most-One-$\acute{\sigma}$	$\mathrm{LTT}_{1,2}$	SP_2
No-H -before-$\acute{H}$	$\boxed{\mathrm{SF}}$	SP_2
No-H -with-$\acute{L}$	LT_1	SP_2
Nothing-before-$\acute{L}$	SL_2	SP_2
Alt	SL_2	$\boxed{\mathrm{SF}}$
No $\rtimes \acute{L} \ltimes$	SL_3	PT_2

Yidin is SF with either local or piecewise constraints.

Slide 54

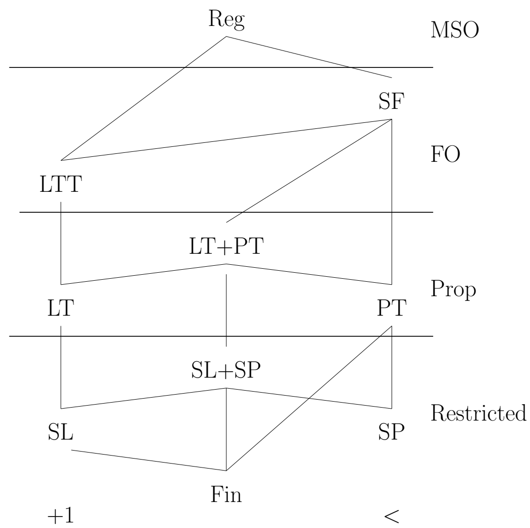
Yidin wrt Local and Piecewise Constraints

One-$\acute{\sigma}$	$\mathrm{LTT}_{1,2}$	$\mathrm{PT}_{\mathbf{2}}$
Some-$\acute{\sigma}$	$\boxed{\mathrm{LT}_{\mathbf{1}}}$	$\boxed{\mathrm{PT}_{\mathbf{1}}}$
At-Most-One-$\acute{\sigma}$	$\mathrm{LTT}_{1,2}$	$\boxed{\mathrm{SP}_{\mathbf{2}}}$
No-H -before-$\acute{H}$	SF	$\boxed{\mathrm{SP}_{\mathbf{2}}}$
No-H -with-$\acute{L}$	LT_1	$\boxed{\mathrm{SP}_{\mathbf{2}}}$
Nothing-before-$\acute{L}$	$\boxed{\mathrm{SL}_{\mathbf{2}}}$	$\boxed{\mathrm{SP}_{\mathbf{2}}}$
Alt	$\boxed{\mathrm{SL}_{\mathbf{2}}}$	SF
No $\rtimes \acute{L} \ltimes$	$\boxed{\mathrm{SL}_{\mathbf{3}}}$	PT_2

Yidin is co-occurence of SL and PT constraints or of LT and SP constraints

Slide 55

Local and Piecewise Hierarchies



Slide 56

MSO definable stringsets

$\langle \mathcal{D}, \triangleleft, \triangleleft^{+}, P_{\sigma} \rangle_{\sigma \in \Sigma}$

First-order Quantification (positions)

Monadic Second-order Quantification (sets of positions)

$\triangleleft^{+}$ is MSO-definable from $\triangleleft$.

Character of the MSO-definable sets

Slide 57

Theorem 8 (Medvedev, Büchi, Elgot) *A set of strings is MSO-definable over* $\langle \mathcal{D}, \triangleleft, \triangleleft^{+}, P_\sigma \rangle_{\sigma \in \Sigma}$ *iff it is regular.*

Theorem 9 (Chomsky Schützenberger) *A set of strings is Regular iff it is a homomorphic image of a Strictly 2-Local set.*

Theorem 10 $MSO = \exists MSO$ *over strings.*

Cognitive interpretation of Finite-state

Slide 58

- Any cognitive mechanism that can distinguish member strings from non-members of a finite-state stringset must be capable of classifying the events in the input into a finite set of abstract categories and are sensitive to the sequence of those categories.
- Subsumes *any* recognition mechanism in which the amount of information inferred or retained is limited by a fixed finite bound.
- Any cognitive mechanism that has a fixed finite bound on the amount of information inferred or retained in processing sequences of events will be able to recognize *only* finite-state stringsets.

Slide 59

Stress Patterns wrt Local Constraints

- SL — 89 of 109 patterns
- LT

  None
- LTT

  Alawa, Bulgarian, Murik
- SF

  Amele, Arabic (Classical), Buriat, Cheremis (East), Cheremis (Meadow), Chuvash, Golin, Komi, Kuuku Yau, Lithuanian, Mam, Maori, K. Mongolian (Street), K. Mongolian (Stuart), K. Mongolian (Bosson), Nubian, Yidin
- Reg

  Arabic (Cairene)?, Arabic (Negev Bedouin), Arabic (Cyrenaican Bedouin), Klamath

Slide 60

Stress Patterns wrt Piecewise Constraints

- SP

  None
- PT

  Amele, Bulgarian, Chuvash, Golin, Lithuanian, Maori K. Mongolian (Street), Murik,
- SF

  Alawa, Arabic (Classical), Buriat, Cheremis (East), Cheremis (Meadow), Komi, Kuuku Lau, Mam, K. Mongolian (Bosson), K. Mongolian (Stuart), Nubian, Yidin
- Reg

  Arabic (Cairene)?, Arabic (Negev Bedouin), Arabic (Cyrenaican Bedouin), Klamath

Stress Patterns wrt Co-occurrence of Local and Piecewise Constraints

Slide 61

- SL + SP — 89 of 109 patterns
- SL + PT — Komi, Kuuku Lau, Yidin
- LT + SP

  Alawa Amele, Arabic (Classical), Bulgarian, Buriat, Cheremis (East), Cheremis (Meadow), Chuvash, Golin, Komi, Kuuku Lau, Lithuanian, Mam, Maori K. Mongolian (Bosson), K. Mongolian (Street), K. Mongolian (Stuart), Murik, Nubian, Yidin
- SF — None
- Reg

  Arabic (Cairene)?, Arabic (Negev Bedouin), Arabic (Cyrenaican Bedouin), Klamath

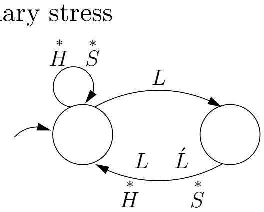
Arabic (Negev Bedouin)

- In sequences of light syllables, secondary stress falls on the even numbered syllables, counting from the left edge of the sequence.
- This pattern is used only for the sake of defining main stress. Secondary stress is absent on the surface.

Slide 62

Without reference to secondary stress

- Odd number of unstressed light syllables precedes a light syllable with primary stress



Arabic (Negev Bedouin) with explicit secondary stress

Slide 63

$$\varphi_{\text{Lalt}} = \neg\, L\, L \wedge \neg\, \grave{L}\, \grave{L} \wedge \neg\, \grave{L}\, \acute{L} \wedge \neg\, \acute{L}\, \grave{L} \wedge \neg\, \overset{*}{H}\, L \wedge \neg\, \overset{*}{S}\, L$$

If secondary stress is explicit, then Arabic (Negev Bedouin) is LT

Some Constraints

Slide 64

- Forbidden syllables (SL_1, SP_1)
  - No heavy syllables
- Required syllables (LT_1, PT_1)
  - Some primary stress
- Forbidden initial/final syllables (SL_2, SF)
  - Cannot start with unstressed light
  - Cannot start with unstressed heavy
  - Cannot end with stressed light
- Forbidden adjacent pairs (SL_2, SF)
  - No adjacent unstressed
  - No adhacent secondary stress
  - No heavy immediately following a stressed light

  . . .

Slide 65

Properly Regular Constraints

- Alternation (Reg)
  - Arabic (Negev Bedouin), ...
  - This class of constraints accounts for all properly regular stress patterns (that are known to us).