Computational Phonology - Class3

Jeffrey Heinz (Instructor)
Jon Rawski (TA)

LSA Summer Institute
UC Davis
July 01, 2019

Last Week

1. We learned the successor model for words.
2. We learned how to express constraints in First Order Logic with this model.
3. We encountered some limitations in the expressivity of this class.
4. There are many paths forward from FO($\triangleleft$).
5. We briefly encountered the precedence model for words.

Outstanding Questions

1. What constraints can we write (and not write) with FO(<)?
2. This defines constraints as functions $f : \Sigma^* \rightarrow \{\texttt{true}, \texttt{false}\}$? How do we count violations? Assign probabilities?
3. This defines constraints. How do we define transformations?
4. What happens when we change the signature? What other signatures are there?
5. What happens when we change the logic? What other logics are there?

Today

1. Analysis of FO(<)
2. Weighting Constraints with Semirings

Part I

Analysis of FO(<)

What constraints can we write (and not write) with FO(<)?

Theorem

A constraint is FO-definable with precedence if and only if there is a natural number n such that for all strings x, y, z and $m > n$, either both xy^nz and xy^mz obey the constraint or neither does.

(Schutzenberger 1965, McNaughton and Paper 1971)

Examples

- Even-Nasal is NOT definable with FO(<).
  - Let $x = y = z = na$.
  - Consider $xy^2z = na(nana)na$ obeys Even-Nasal.
  - But $xy^3z = na(nanana)na$ does not.
  - In fact, for any n, $xy^{2n}z$ obeys Even-Nasal but, $xy^{2n+1}z$ does not.

Examples

- Even-Nasal is NOT definable with FO(<).
  - Let $x = y = z = na$.
  - Consider $xy^2z = na(nana)na$ obeys Even-Nasal.
  - But $xy^3z = na(nanana)na$ does not.
  - In fact, for any n, $xy^{2n}z$ obeys Even-Nasal but, $xy^{2n+1}z$ does not.
- Every constraint which can be written in FO($\triangleleft$) can be written in FO(<). This is because "successor" is FO-definable with precedence.

$$x \triangleleft y \stackrel{\text{def}}{=} x < y \wedge \neg \exists z[x < z \wedge z < y]$$

Outstanding Questions

1. **What constraints can we write (and not write) with FO(<)**
2. This defines constraints as functions $f : \Sigma^* \to \{\texttt{true}, \texttt{false}\}$? How do we count violations? Assign probabilities?
3. This defines constraints. How do we define transformations?
4. What happens when we change the signature? What other signatures are there?
5. What happens when we change the logic? What other logics are there?

Outstanding Questions

1. What constraints can we write (and not write) with FO($<$)
2. This defines constraints as functions $f : \Sigma^* \to \{\texttt{true}, \texttt{false}\}$? How do we count violations? Assign probabilities?
3. This defines constraints. How do we define transformations?
4. What happens when we change the signature? What other signatures are there?
5. **What happens when we change the logic? What other logics are there?**

Monadic Second Order Logic

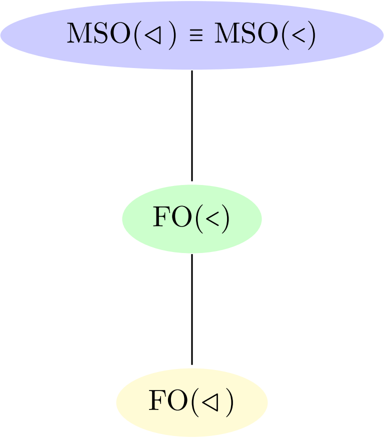
MSO allows quantification over *sets of individuals.*

Additional Symbols in MSO logic	
X, Y, Z	variables which range over sets of elements of the domain
$x \in X$	checks whether an element x belongs to a set of elements X

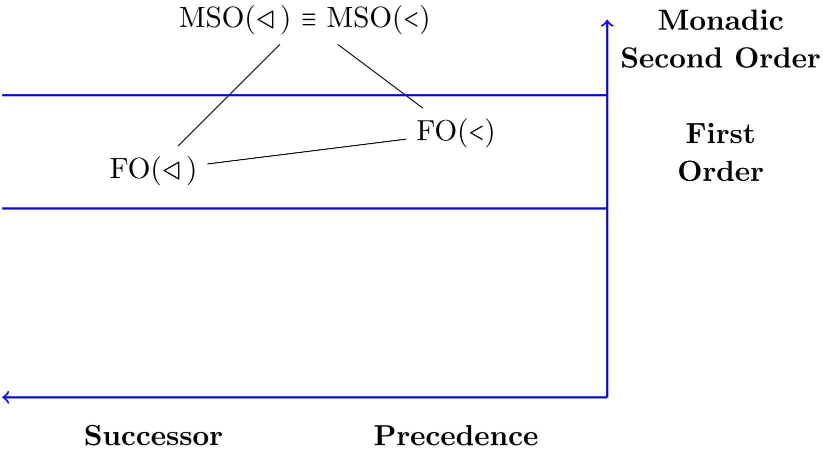
Some Important Facts

- Precedence is MSO definable with successor.
- Even-Nasal is MSO definable with successor (and precedence).

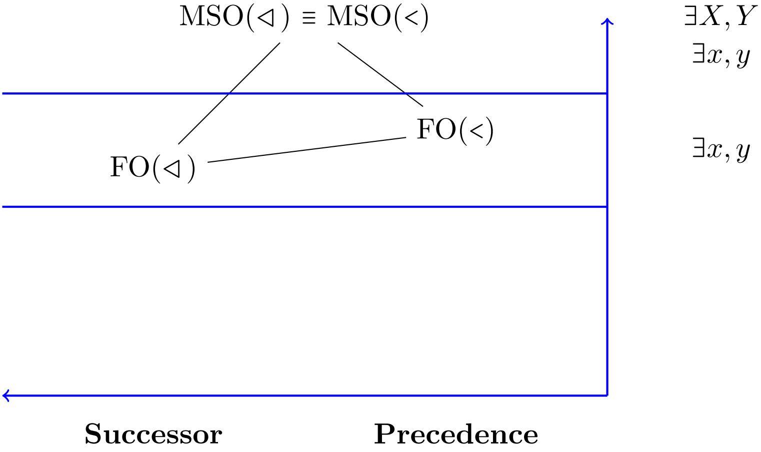
Summary of the expressivity of the logical languages (so far)



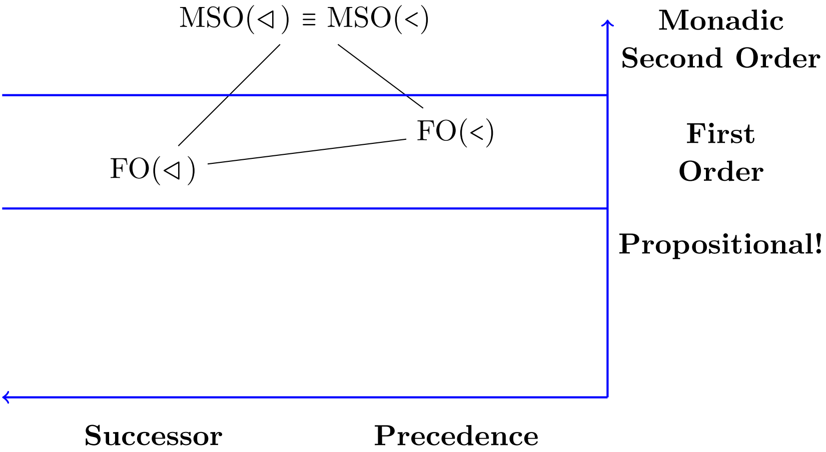
Another view



Another view



Another view



Outstanding Questions

1. What constraints can we write (and not write) with FO(<)
2. **This defines constraints as functions** $f : \Sigma^* \to \{\texttt{true}, \texttt{false}\}$**? How do we count violations? Assign probabilities?**
3. This defines constraints. How do we define transformations?
4. What happens when we change the signature? What other signatures are there?
5. What happens when we change the logic? What other logics are there?

Part II

Weighting Constraints with Semirings

Semirings

Semirings abstract away from the details of certain mathematical sets. A semiring is a set S

1. with two operations $\oplus$ and $\otimes$ over each of its elements, and
2. S contains two *identity* elements: 0 for $\oplus$ and 1 for $\otimes$.

Name	S	$\oplus$	$\otimes$	0	1
Boolean	$\{\texttt{true}, \texttt{false}\}$	$\vee$	$\wedge$	`false`	`true`
Natural	$\mathbb{N}$	$+$	$\times$	0	1
Real Interval	$[0, 1]$	$+$	$\times$	0	1

(See handout for full set of properties. And see Droste and Gastin (2009).)

Syntax Weighted MSO Logic

Fix a relational model signature $\mathbb{M}$.

Base Cases

(B1)	$\boldsymbol{s}$, for each $s \in S$	(**atomic semiring element**)
(B2)	$\boldsymbol{x = y}$	(**equality**)
(B3)	$\boldsymbol{\neg(x = y)}$	(**non-equality**)
(B4)	$\boldsymbol{x \in X}$	(**membership**)
(B5)	$\boldsymbol{\neg(x \in X)}$	(**non-membership**)
(B6)	$\boldsymbol{R(\vec{x})}$, for each $R \in \mathbb{M}$	(**positive relational atom**)
(B7)	$\boldsymbol{\neg R(\vec{x})}$, for each $R \in \mathbb{M}$	(**negative relational atom**)

Semantics of WMSO

The base cases

$$
\begin{array}{lll}
\text{(B1)} & [\![\boldsymbol{s}]\!](\mathbb{S}, w) & \overset{\text{def}}{=} \quad s \\
\text{(B2)} & [\![(\boldsymbol{x = y})]\!](\mathbb{S}[x \mapsto e_1, y \mapsto e_2], w) & \overset{\text{def}}{=} \quad \begin{array}{l} 1 \text{ iff } e_1 = e_2 \\ 0 \text{ otherwise} \end{array} \\
\text{(B3)} & [\![(\neg\boldsymbol{(x = y)})]\!](\mathbb{S}[x \mapsto e_1, y \mapsto e_2], w) & \overset{\text{def}}{=} \quad \begin{array}{l} 0 \text{ iff } e_1 = e_2 \\ 1 \text{ otherwise} \end{array} \\
\text{(B4)} & [\![\boldsymbol{x \in X}]\!](\mathbb{S}[x \mapsto e, X \mapsto E], w) & \overset{\text{def}}{=} \quad \begin{array}{l} 1 \text{ iff } e \in E \\ 0 \text{ otherwise} \end{array} \\
\text{(B5)} & [\![\neg\boldsymbol{(x \in X)}]\!](\mathbb{S}[x \mapsto e, X \mapsto E], w) & \overset{\text{def}}{=} \quad \begin{array}{l} 0 \text{ iff } e \in E \\ 1 \text{ otherwise} \end{array} \\
\text{(B6)} & [\![\boldsymbol{R(\vec{x})}]\!](\mathbb{S}[\vec{x} \mapsto \vec{e}\,], w) & \overset{\text{def}}{=} \quad \begin{array}{l} 1 \text{ iff } \mathcal{M}_w \models R(\vec{e}\,) \\ 0 \text{ otherwise} \end{array} \\
\text{(B7)} & [\![\neg\boldsymbol{R(\vec{x})}]\!](\mathbb{S}[\vec{x} \mapsto \vec{e}\,], w) & \overset{\text{def}}{=} \quad \begin{array}{l} 0 \text{ iff } \mathcal{M}_w \models R(\vec{e}\,) \\ 1 \text{ otherwise} \end{array}
\end{array}
$$

Syntax Weighted MSO Logic

Inductive Cases If φ, ψ are formulas of MSO logic, then so are

(I1)	$\boldsymbol{(\varphi \vee \psi)}$	(**disjunction**)
(I2)	$\boldsymbol{(\varphi \wedge \psi)}$	(**conjunction**)
(I3)	$\boldsymbol{(\exists x)[\varphi]}$	(**existential quant. for individuals**)
(I4)	$\boldsymbol{(\exists X)[\varphi]}$	(**existential quant. for sets of individuals**)
(I5)	$\boldsymbol{(\forall x)[\varphi]}$	(**universal quant. for individuals**)
(I6)	$\boldsymbol{(\forall X)[\varphi]}$	(**universal quant. for sets of individuals**)

Nothing else is a formula of MSO logic. Note negation only applies to the base cases.

Semantic of WMSO

The inductive cases.

$$
\begin{array}{llcl}
\text{(I1)} & [\![(\varphi \vee \psi)]\!](\mathbb{S}, w) & \stackrel{\text{def}}{=} & [\![\varphi]\!](\mathbb{S}, w) \oplus [\![\varphi]\!](\mathbb{S}, w) \\
\text{(I2)} & [\![(\varphi \wedge \psi)]\!](\mathbb{S}, w) & \stackrel{\text{def}}{=} & [\![\varphi]\!](\mathbb{S}, w) \otimes [\![\varphi]\!](\mathbb{S}, w) \\
\text{(I3)} & [\![(\exists x)[\varphi]]\!](\mathbb{S}, w) & \stackrel{\text{def}}{=} & \bigoplus_{e \in D} [\![\varphi]\!](\mathbb{S}[x \mapsto e], w) \\
\text{(I4)} & [\![(\exists X)[\varphi]]\!](\mathbb{S}, w) & \stackrel{\text{def}}{=} & \bigoplus_{E \in D} [\![\varphi]\!](\mathbb{S}[X \mapsto E], w) \\
\text{(I5)} & [\![(\forall x)[\varphi]]\!](\mathbb{S}, w) & \stackrel{\text{def}}{=} & \bigotimes_{e \in D} [\![\varphi]\!](\mathbb{S}[x \mapsto e], w) \\
\text{(I6)} & [\![(\forall X)[\varphi]]\!](\mathbb{S}, w) & \stackrel{\text{def}}{=} & \bigotimes_{E \in D} [\![\varphi]\!](\mathbb{S}[X \mapsto E], w)
\end{array}
$$

Main points:

1. existential quantification and disjunction are interpreted with $\oplus$
2. universal quantification and conjunction are interpreted with $\otimes$

Example: *c (categorical)

Model signature: $\mathbb{M} = \langle D \mid \mathtt{a}, \mathtt{b}, \mathtt{c}, \triangleleft \rangle$
Boolean Semiring

Name	S	$\oplus$	$\otimes$	0	1
Boolean	$\{\mathtt{true}, \mathtt{false}\}$	$\vee$	$\wedge$	`false`	`true`

A WMSO Sentence for *c

$$^{*}c \stackrel{\text{def}}{=} \forall x[\neg \mathtt{c}(x)] \tag{1}$$

Examples

1. $[\![^{*}c]\!](acbc) = \neg \mathtt{c}(1) \wedge \neg \mathtt{c}(2) \wedge \neg \mathtt{c}(3) \wedge \neg \mathtt{c}(4) = \mathtt{false}$
2. $[\![^{*}c]\!](abba) = \neg \mathtt{c}(1) \wedge \neg \mathtt{c}(2) \wedge \neg \mathtt{c}(3) \wedge \neg \mathtt{c}(4) = \mathtt{true}$

Example: *c (counting violations)

Model signature: $\mathbb{M} = \langle D \mid \mathsf{a}, \mathsf{b}, \mathsf{c}, \lhd \rangle$
Integer Semiring

Name	S	$\oplus$	$\otimes$	0	1
Natural	$\mathbb{N}$	$+$	$\times$	0	1

A WMSO Sentence for *c

$$^{*}c \overset{\text{def}}{=} \exists x[\mathsf{c}(x)] \tag{2}$$

Examples

1. $[\![{}^{*}c]\!](acbc) = \mathsf{c}(1) + \mathsf{c}(2) + \mathsf{c}(3) + \mathsf{c}(4) = 2$
2. $[\![{}^{*}c]\!](abba) = \mathsf{c}(1) + \mathsf{c}(2) + \mathsf{c}(3) + \mathsf{c}(4) = 0$

Example: *c (probabilities)

Model signature: $\mathbb{M} = \langle D \mid \mathtt{a}, \mathtt{b}, \mathtt{c}, \triangleleft \rangle$

Integer Semiring

Name	S	$\oplus$	$\otimes$	0	1
Natural	$\mathbb{N}$	$+$	$\times$	0	1

A WMSO Sentence for *c

$$^{*}c \overset{\text{def}}{=} \exists x\Big[\big(\mathtt{a}(x) \wedge 0.4\big) \vee \big(\mathtt{b}(x) \wedge 0.4\big) \vee \big(\mathtt{c}(x) \wedge 0.2\big)\Big] \quad (3)$$

Example

$$\begin{aligned}
[\![{}^{*}c]\!](acbc) &= \Big(\big(\mathtt{a}(1) \times 0.4\big) + \big(\mathtt{b}(1) \times 0.4\big) + \big(\mathtt{c}(1) \times 0.2\big)\Big) \\
&\quad \times\Big(\big(\mathtt{a}(2) \times 0.4\big) + \big(\mathtt{2}(1) \times 0.4\big) + \big(\mathtt{2}(1) \times 0.2\big)\Big) \\
&\quad \times\Big(\big(\mathtt{a}(3) \times 0.4\big) + \big(\mathtt{3}(1) \times 0.4\big) + \big(\mathtt{3}(1) \times 0.2\big)\Big) \\
&\quad \times\Big(\big(\mathtt{a}(4) \times 0.4\big) + \big(\mathtt{4}(1) \times 0.4\big) + \big(\mathtt{4}(1) \times 0.2\big)\Big) \\
&= 0.4 \times 0.2 \times 0.4 \times 0.2 = 0.0064
\end{aligned}$$

WMSO Summary

1. In weighted logic, negation only applies to the base cases.
2. In the base cases, model satisfaction yields semiring 1, 0 otherwise.
3. Conjunction and disjunction are interpreted as semiring multiplication $\otimes$ and addition $\oplus$.
4. There are *many* semirings, including ones for strings.
5. They form an important chapter in processing in NLP.
6. Use of logic does not preclude studying phenomena deemed gradient or probabilistic.
7. See handout for more examples with *NT.

Where we are

Today

1. Analysis of FO(<)
2. WMSO with semirings.

Next Class

- So far, we have looked at generalizations over structures (constraints).
- How do we define transformations?

Homework

1. Let $\mathbb{M} = \langle D \mid \mathtt{a}, \mathtt{b}, \mathtt{c}, < \rangle$. Using the natural semiring, consider the WMSO sentence

$$^{*}\mathsf{N..L} \overset{\text{def}}{=} \exists x, y[x < y \land \mathtt{nasal}(x) \land \mathtt{cons}(y) \land \neg\mathtt{voiced}(y)].$$

What is $[\![^{*}\mathsf{N..L}]\!](nilanola)$? Show your work.