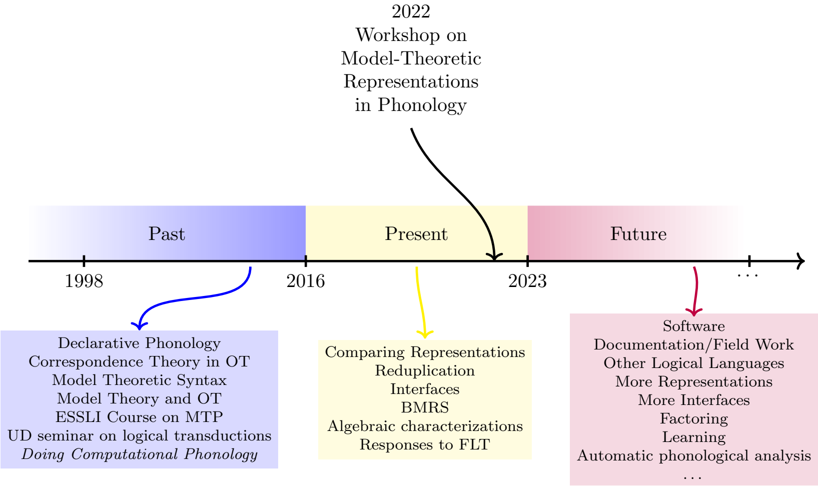
The Past, Present, and Future of Model-Theoretic Phonology

Jeffrey Heinz and Scott Nelson

Stony Brook
September 22, 2022

This Talk



2022
Workshop on
Model-Theoretic
Representations
in Phonology

Declarative Phonology
Correspondence Theory in OT
Model Theoretic Syntax
Model Theory and OT
ESSLI Course on MTP
UD seminar on logical transductions
Doing Computational Phonology

Comparing Representations
Reduplication
Interfaces
BMRS
Algebraic characterizations
Responses to FLT

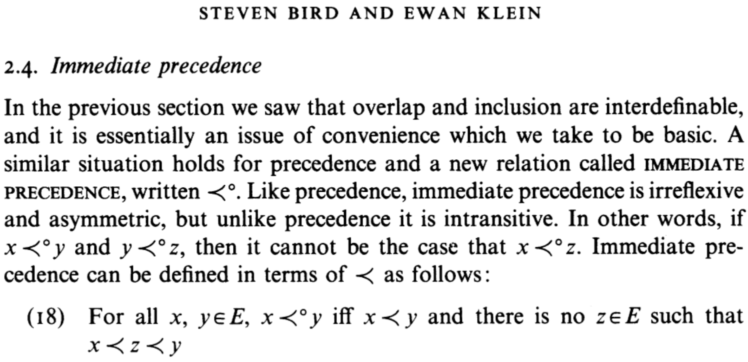
Software
Documentation/Field Work
Other Logical Languages
More Representations
More Interfaces
Factoring
Learning
Automatic phonological analysis
...

Past

SOME KEY IDEAS

1. The use of logic in phonology has ample precedent
2. Model-theoretic analysis of linguistic theories
3. Turning point with logical transductions

1990s: Declarative Phonology



STEVEN BIRD AND EWAN KLEIN

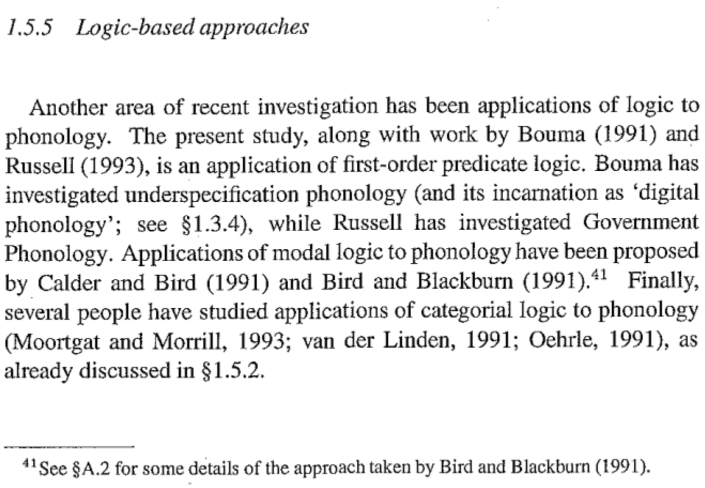
2.4. *Immediate precedence*

In the previous section we saw that overlap and inclusion are interdefinable, and it is essentially an issue of convenience which we take to be basic. A similar situation holds for precedence and a new relation called IMMEDIATE PRECEDENCE, written $\prec^\circ$. Like precedence, immediate precedence is irreflexive and asymmetric, but unlike precedence it is intransitive. In other words, if $x \prec^\circ y$ and $y \prec^\circ z$, then it cannot be the case that $x \prec^\circ z$. Immediate precedence can be defined in terms of $\prec$ as follows:

(18) For all x, $y \in E$, $x \prec^\circ y$ iff $x \prec y$ and there is no $z \in E$ such that $x \prec z \prec y$

Bird and Klein 1990, Phonological Events

1990S: DECLARATIVE PHONOLOGY



1.5.5 Logic-based approaches

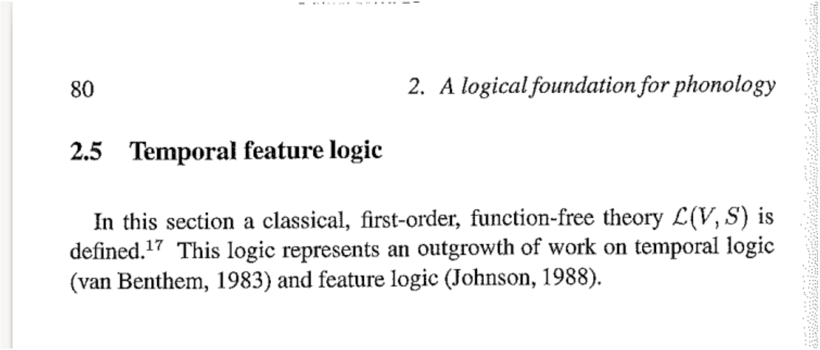
Another area of recent investigation has been applications of logic to phonology. The present study, along with work by Bouma (1991) and Russell (1993), is an application of first-order predicate logic. Bouma has investigated underspecification phonology (and its incarnation as 'digital phonology'; see §1.3.4), while Russell has investigated Government Phonology. Applications of modal logic to phonology have been proposed by Calder and Bird (1991) and Bird and Blackburn (1991).[41] Finally, several people have studied applications of categorial logic to phonology (Moortgat and Morrill, 1993; van der Linden, 1991; Oehrle, 1991), as already discussed in §1.5.2.

---

[41] See §A.2 for some details of the approach taken by Bird and Blackburn (1991).

Bird 1995, Computational Phonology: A Constraint-Based Approach

1990s: Declarative Phonology



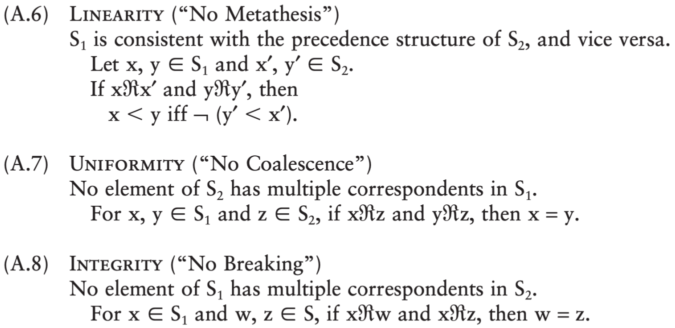
80 *2. A logical foundation for phonology*

2.5 Temporal feature logic

In this section a classical, first-order, function-free theory $\mathcal{L}(V, S)$ is defined.[17] This logic represents an outgrowth of work on temporal logic (van Benthem, 1983) and feature logic (Johnson, 1988).

Bird 1995, Computational Phonology: A Constraint-Based Approach

1990s: Logic in Optimality Theory



(A.6) LINEARITY ("No Metathesis")
S_1 is consistent with the precedence structure of S_2, and vice versa.
Let $x, y \in S_1$ and $x', y' \in S_2$.
If $x \mathfrak{R} x'$ and $y \mathfrak{R} y'$, then
$x < y$ iff $\neg\,(y' < x')$.

(A.7) UNIFORMITY ("No Coalescence")
No element of S_2 has multiple correspondents in S_1.
For $x, y \in S_1$ and $z \in S_2$, if $x \mathfrak{R} z$ and $y \mathfrak{R} z$, then $x = y$.

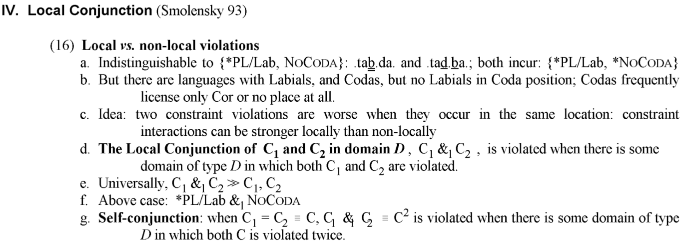
(A.8) INTEGRITY ("No Breaking")
No element of S_1 has multiple correspondents in S_2.
For $x \in S_1$ and $w, z \in S$, if $x \mathfrak{R} w$ and $x \mathfrak{R} z$, then $w = z$.

- Faithfulness constraints in correspondence theory are defined essentially using first order logic

---

McCarthy & Prince, 1995/1999

1990S: LOGIC IN OPTIMALITY THEORY



IV. Local Conjunction (Smolensky 93)

(16) **Local *vs.* non-local violations**

- a. Indistinguishable to {*PL/Lab, NOCODA}: .tab.da. and .tad.ba.; both incur: {*PL/Lab, *NOCODA}
- b. But there are languages with Labials, and Codas, but no Labials in Coda position; Codas frequently license only Cor or no place at all.
- c. Idea: two constraint violations are worse when they occur in the same location: constraint interactions can be stronger locally than non-locally
- d. **The Local Conjunction of $\mathbf{C_1}$ and $\mathbf{C_2}$ in domain *D***, $C_1 \&_1 C_2$, is violated when there is some domain of type *D* in which both C_1 and C_2 are violated.
- e. Universally, $C_1 \&_1 C_2 \gg C_1, C_2$
- f. Above case: *PL/Lab $\&_1$ NOCODA
- g. **Self-conjunction**: when $C_1 = C_2 \equiv C$, $C_1 \&_1 C_2 \equiv C^2$ is violated when there is some domain of type *D* in which both C is violated twice.

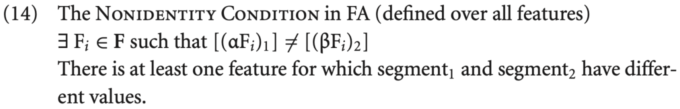
Smolensky 1995; On the Internal Structure of the Constraint Component *Con* of UG

1990S: LOGIC IN OPTIMALITY THEORY

In the following sections, we show that phonological behaviour associated with macro-constraints reveals striking parallels between constraint co-ordination and familiar operations of classical propositional logic. We exploit the analogy with logic extensively in developing our model, and argue that, just as between arguments in a Boolean expression, the relationship between co-ordinated constraints is symmetrical in some cases, asymmetrical in others.

Crowhurst & Hewitt 1997; Boolean Operations and Constraint Interactions in Optimality Theory

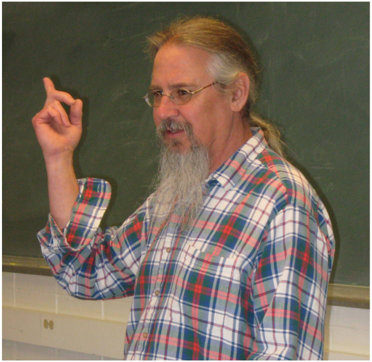
2000s: Quantification in Structural Descriptions

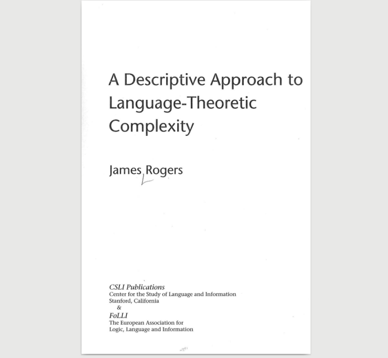


(14) The Nonidentity Condition in FA (defined over all features)
$\exists\ F_i \in \mathbf{F}$ such that $[(\alpha F_i)_1] \neq [(\beta F_i)_2]$
There is at least one feature for which segment_1 and segment_2 have different values.

Reiss 2003a/b, (see also *Towards a Theory of Fundamental Phonological Relations*)

1990S: MODEL-THEORETIC SYNTAX





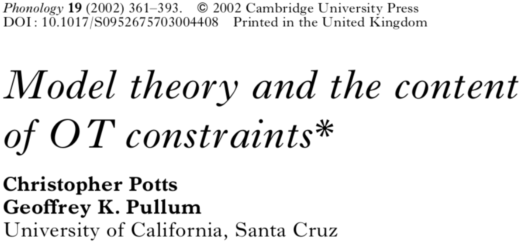
A Descriptive Approach to
Language-Theoretic
Complexity

James Rogers

CSLI Publications
Center for the Study of Language and Information
Stanford, California
&
FoLLI
The European Association for
Logic, Language and Information

The 2017 SIGMOL S.-Y. Kuroda Prize is awarded to James Rogers (Earlham College). James Rogers's 1998 book, "A Descriptive Approach to Language-Theoretic Complexity," was the first comprehensive work to apply monadic second-order logic to the analysis of linguistic theories. . . `http://molweb.org/mol/award-2017.html`

2000s: Modal Logic for Constraints

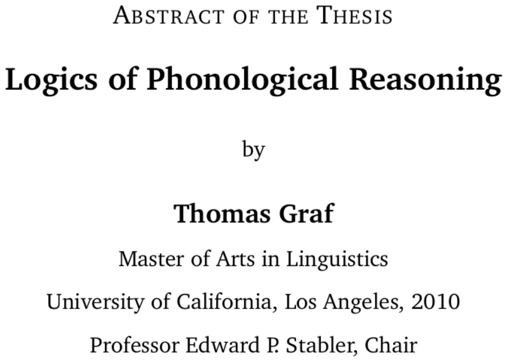


Phonology **19** (2002) 361–393. © 2002 Cambridge University Press
DOI: 10.1017/S0952675703004408 Printed in the United Kingdom

*Model theory and the content of OT constraints**

Christopher Potts
Geoffrey K. Pullum
University of California, Santa Cruz

2010s: Logical Analysis of SPE and Government Phonology



Abstract of the Thesis

Logics of Phonological Reasoning

by

Thomas Graf

Master of Arts in Linguistics

University of California, Los Angeles, 2010

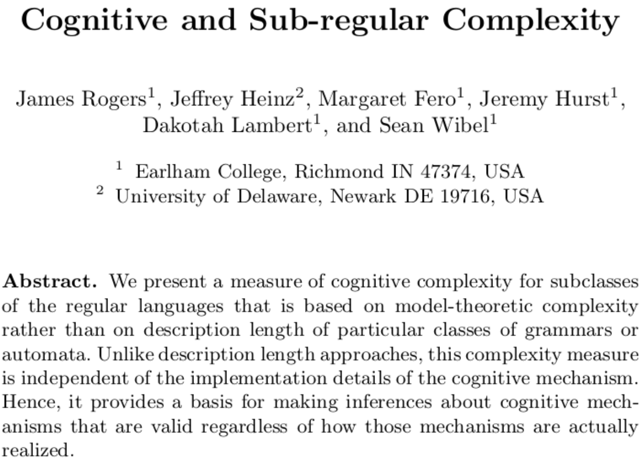
Professor Edward P. Stabler, Chair

2010s: Logical Analysis of SPE and Government Phonology

> *MTP is based on the insight that many attributes of a theory are reflected in the properties of the weakest language one can use to describe it. In model-theoretic approaches, this description language is some logic chosen from the array of logics one encounters in mathematics and computer science.*

Graf 2010:1-2

2010S: LOGICAL ANALYSES OF PHONOTACTIC/STRESS PATTERNS



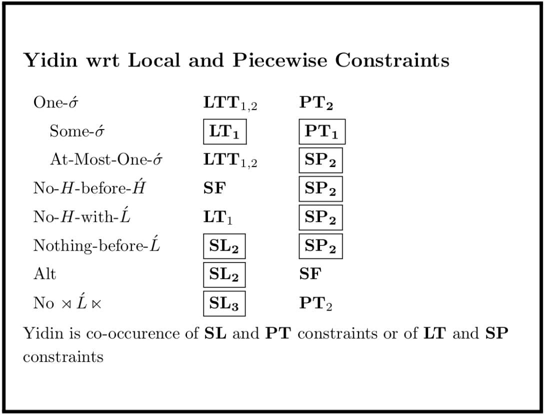
Cognitive and Sub-regular Complexity

James Rogers[1], Jeffrey Heinz[2], Margaret Fero[1], Jeremy Hurst[1],
Dakotah Lambert[1], and Sean Wibel[1]

[1] Earlham College, Richmond IN 47374, USA
[2] University of Delaware, Newark DE 19716, USA

Abstract. We present a measure of cognitive complexity for subclasses of the regular languages that is based on model-theoretic complexity rather than on description length of particular classes of grammars or automata. Unlike description length approaches, this complexity measure is independent of the implementation details of the cognitive mechanism. Hence, it provides a basis for making inferences about cognitive mechanisms that are valid regardless of how those mechanisms are actually realized.

2010S: LOGICAL ANALYSES OF PHONOTACTIC/STRESS PATTERNS



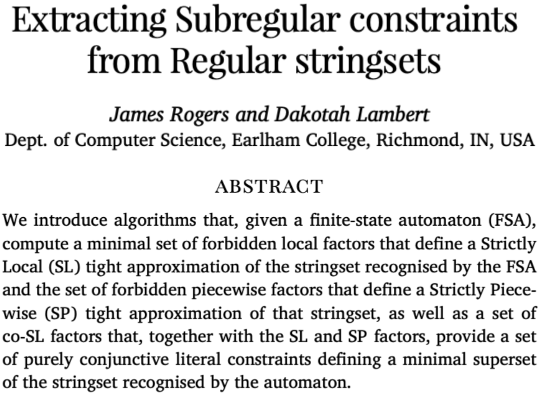
Yidin wrt Local and Piecewise Constraints

One-$\acute{\sigma}$	$\mathbf{LTT}_{1,2}$	$\mathbf{PT_2}$
Some-$\acute{\sigma}$	$\boxed{\mathbf{LT_1}}$	$\boxed{\mathbf{PT_1}}$
At-Most-One-$\acute{\sigma}$	$\mathbf{LTT}_{1,2}$	$\boxed{\mathbf{SP_2}}$
No-H-before-$\acute{H}$	**SF**	$\boxed{\mathbf{SP_2}}$
No-H-with-$\acute{L}$	$\mathbf{LT}_1$	$\boxed{\mathbf{SP_2}}$
Nothing-before-$\acute{L}$	$\boxed{\mathbf{SL_2}}$	$\boxed{\mathbf{SP_2}}$
Alt	$\boxed{\mathbf{SL_2}}$	**SF**
No $\rtimes \acute{L} \ltimes$	$\boxed{\mathbf{SL_3}}$	$\mathbf{PT}_2$

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

Course at ESSLI 2014 by Rogers & Heinz

2010s: Logical Analyses of Phonotactic/Stress Patterns



Extracting Subregular constraints from Regular stringsets

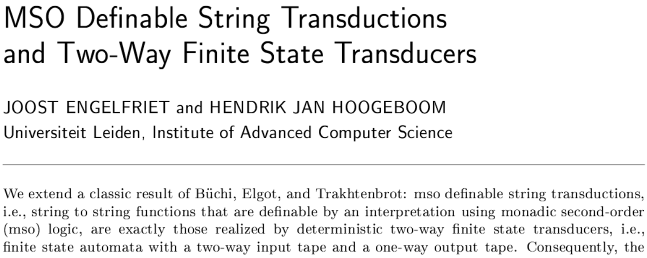
James Rogers and Dakotah Lambert
Dept. of Computer Science, Earlham College, Richmond, IN, USA

ABSTRACT

We introduce algorithms that, given a finite-state automaton (FSA), compute a minimal set of forbidden local factors that define a Strictly Local (SL) tight approximation of the stringset recognised by the FSA and the set of forbidden piecewise factors that define a Strictly Piecewise (SP) tight approximation of that stringset, as well as a set of co-SL factors that, together with the SL and SP factors, provide a set of purely conjunctive literal constraints defining a minimal superset of the stringset recognised by the automaton.

2019; *Journal of Language Modeling*

2015: UD SEMINAR ON LOGICAL TRANSDUCTIONS



MSO Definable String Transductions and Two-Way Finite State Transducers

JOOST ENGELFRIET and HENDRIK JAN HOOGEBOOM
Universiteit Leiden, Institute of Advanced Computer Science

We extend a classic result of Büchi, Elgot, and Trakhtenbrot: mso definable string transductions, i.e., string to string functions that are definable by an interpretation using monadic second-order (mso) logic, are exactly those realized by deterministic two-way finite state transducers, i.e., finite state automata with a two-way input tape and a one-way output tape. Consequently, the

---

ACM Transations on Computational Logic, Vol. 1, No. 4, April 2001, Pages 1-38.

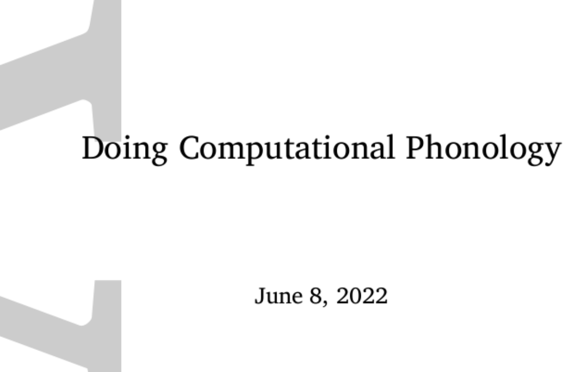
2015: UD SEMINAR ON LOGICAL TRANSDUCTIONS

We realized we could analyze morpho-phonological transformations

- Using logic
- Using linguistic representations

and that it was sufficiently powerful to capture phenomena like total reduplication, which hitherto (to us) had been outside of what we thought was possible with finite-state models.

2015: UD Seminar on Logical Transductions



Doing Computational Phonology

June 8, 2022

Present

SOME KEY IDEAS

1. Non-string representations
2. Rational vs Regular Relations
3. Quantifier Free Logic and Locality
4. Comparing Different Representations
5. Connections to Connectionism
6. Boolean Monadic Recursive Schemes
7. Algebraic Characterizations

REPRESENTATIONS

1. Autosegmental structures (Jardine 2016, 2017, Chandlee and Jardine 2019, Oakden 2020)
2. Syllable structures (Strother-Garcia 2018, 2019)
3. Prosodic structures (Dolatian 2020)
4. Morphological structures (Dolatian 2020)
5. Signed structures (Rawski 2017)
6. Articulatory Phonology structures (Chadwick 2020, Nelson 2022)
7. Features (Strother Garcia et al. 2016, Nelson 2021)

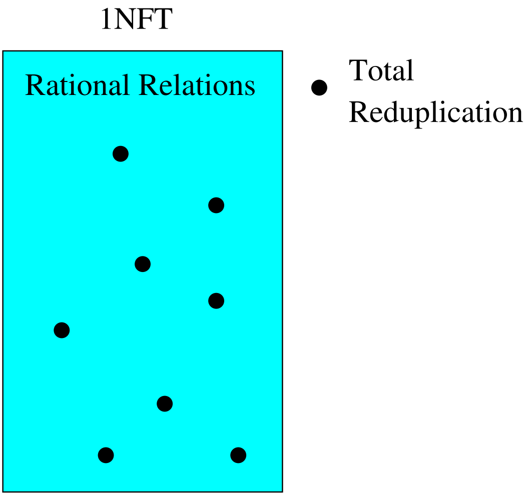
STRING-TO-STRING FUNCTIONS

The established, foundational view (Roark and Sproat 2007)

Rational Relations	non-Rational Relations
Affixation	Total Reduplication
Truncation	
Root and pattern	
Umlaut/Ablaut	
Partial Reduplication	
Phonological Processes	
…	

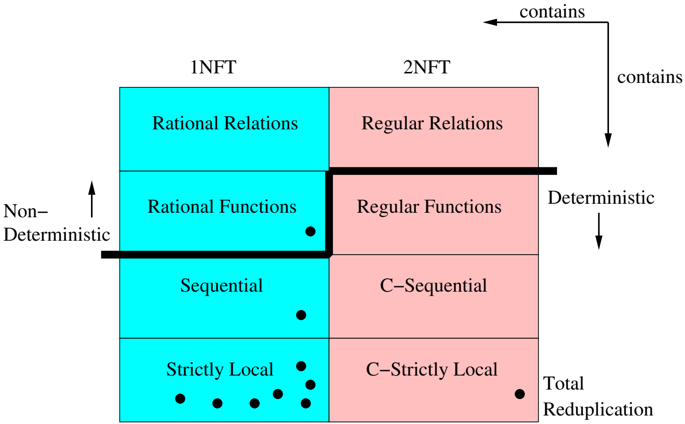
String-to-String Functions

In pictures:



STRING-TO-STRING FUNCTIONS

A more articulated view:



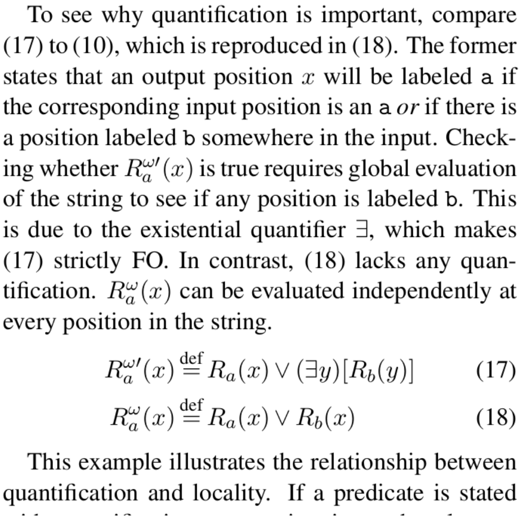
(Chandlee 2017, Dolatian and Heinz 2021)

QUANTIFIER FREE LOGIC

More generally, our main result is that ISL functions (without null cycles) are ***quantifier-free logical interpretations over strings with adjacency.***

Chandlee & Lindell (June 2021 version), Logical perspectives on strictly local transformations

How QF is related to being local



To see why quantification is important, compare (17) to (10), which is reproduced in (18). The former states that an output position x will be labeled `a` if the corresponding input position is an `a` *or* if there is a position labeled `b` somewhere in the input. Checking whether $R_a^{\omega\prime}(x)$ is true requires global evaluation of the string to see if any position is labeled `b`. This is due to the existential quantifier $\exists$, which makes (17) strictly FO. In contrast, (18) lacks any quantification. $R_a^{\omega}(x)$ can be evaluated independently at every position in the string.

$$R_a^{\omega\prime}(x) \overset{\text{def}}{=} R_a(x) \vee (\exists y)[R_b(y)] \qquad (17)$$

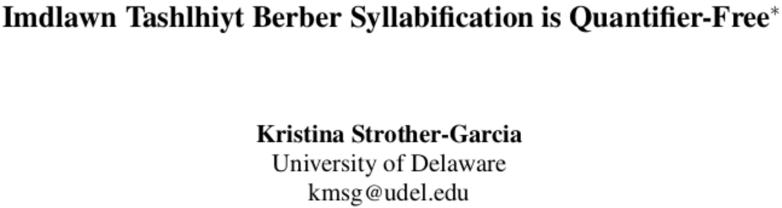
$$R_a^{\omega}(x) \overset{\text{def}}{=} R_a(x) \vee R_b(x) \qquad (18)$$

This example illustrates the relationship between quantification and locality. If a predicate is stated

---

Strother-Garcia 2018, Imdlawn Tashlhiyt Berber Syllabification is Quantifier-Free

QF (Local) Transformations



Imdlawn Tashlhiyt Berber Syllabification is Quantifier-Free[*]

Kristina Strother-Garcia
University of Delaware
kmsg@udel.edu

Strother-Garcia 2018, Imdlawn Tashlhiyt Berber Syllabification is Quantifier-Free

QF (LOCAL) TRANSFORMATIONS

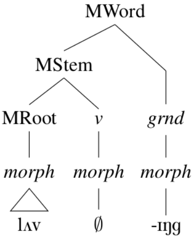
The goal of this paper has been to apply a rigorous, independently motivated notion of locality to investigate the notion that ARs "make non-local patterns local." The fact that in Table 1 common tone patterns fill out the logical possibilities of ISL and A-ISL shows that this statement is not automatically true. In fact, we found some patterns that look intuitively 'local' with ARs but, under a rigorous definition, are not. This opens up a rich line of investigation into what definitions of locality allow ARs to "make non-local patterns local," and under what conditions these definitions hold.

Chandlee and Jardine 2019, Autosegmental Input Strictly Local Functions

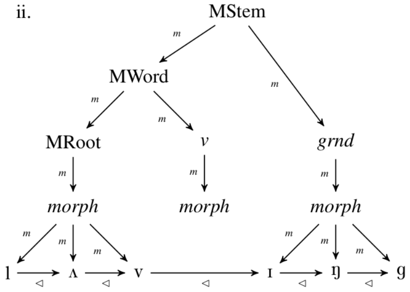
PROSODIC MORPHO-PHONOLOGY

b. *loving*

i.



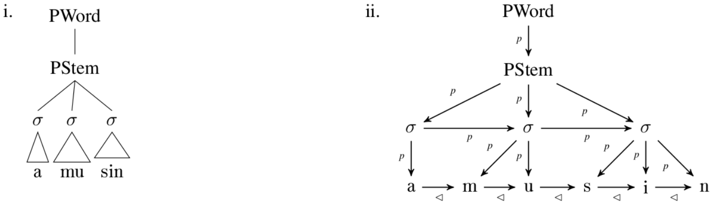
ii.



Dolatian 2020: chap. 4, Computational locality of cyclic phonology in Armenian

PROSODIC MORPHO-PHONOLOGY

(237) a. amusín ‘husband’



Dolatian 2020: chap. 4, Computational locality of cyclic phonology in Armenian

MORPHOLOGY-PHONOLOGY INTERFACE

> *I showed that the brunt of the interface consists of computationally simple or local processes. . . . In fact, by assuming interactionism and the SETTINGS factorization, nearly the entire interface is QF. Non-locality is restricted to the generation of tiers, different types of allomorphy, the need for settings examination for cophonology selection, and post-cyclic prosody.*
>
> *. . . The goal of this dissertation was not to make everything in morphology-phonology become computationally local. Rather the goal is to understand which representations and analytical choices can create locality or non-locality.*

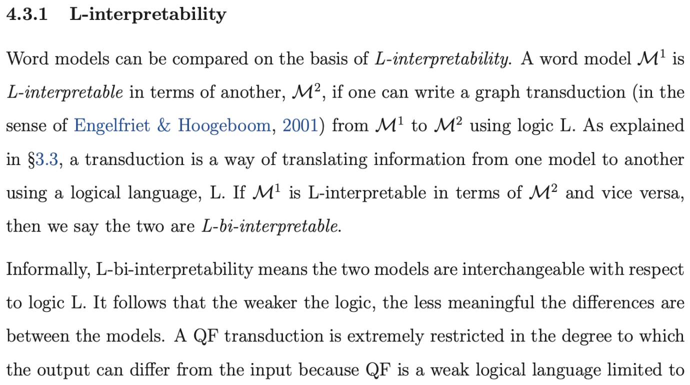
Dolatian 2020:357, Computational locality of cyclic phonology in Armenian

COMPARING PHONOLOGICAL REPRESENTATIONS

> *This dissertation investigates the computational properties of syllable-based phenomena using tools from Model Theory...After introducing the necessary formalisms from Model Theory...I show that three types of syllable structure representations from the literature are **notationally equivalent**, meaning we can 'translate' between them very easily without loss of information.*

Strother-Garcia 2019, Using Model Theory in Phonology: A Novel Characterization of Syllable Structure and Syllabification

Comparing Phonological Representations



4.3.1 L-interpretability

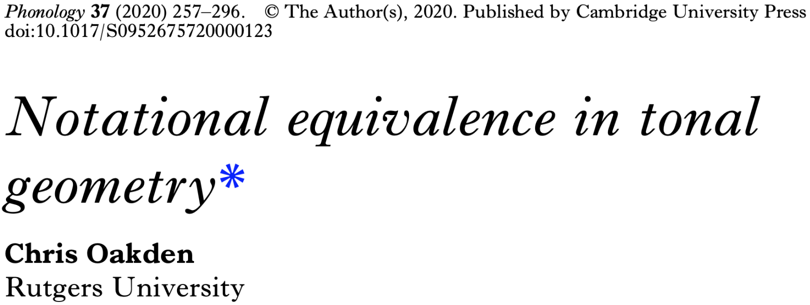
Word models can be compared on the basis of *L-interpretability*. A word model $\mathcal{M}^1$ is *L-interpretable* in terms of another, $\mathcal{M}^2$, if one can write a graph transduction (in the sense of Engelfriet & Hoogeboom, 2001) from $\mathcal{M}^1$ to $\mathcal{M}^2$ using logic L. As explained in §3.3, a transduction is a way of translating information from one model to another using a logical language, L. If $\mathcal{M}^1$ is L-interpretable in terms of $\mathcal{M}^2$ and vice versa, then we say the two are *L-bi-interpretable*.

Informally, L-bi-interpretability means the two models are interchangeable with respect to logic L. It follows that the weaker the logic, the less meaningful the differences are between the models. A QF transduction is extremely restricted in the degree to which the output can differ from the input because QF is a weak logical language limited to

---

Strother-Garcia 2019, Using Model Theory in Phonology: A Novel Characterization of Syllable Structure and Syllabification

Comparing Phonological Representations

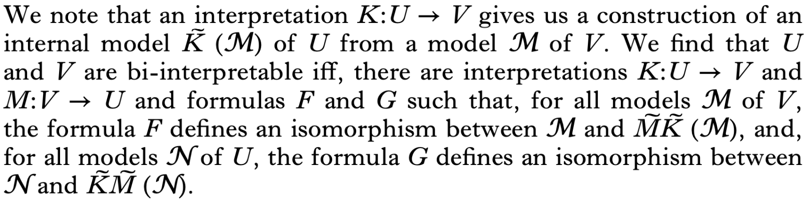


Phonology **37** (2020) 257–296.
doi:10.1017/S0952675720000123

*Notational equivalence in tonal geometry**

Chris Oakden
Rutgers University

COMPARING PHONOLOGICAL REPRESENTATIONS



We note that an interpretation $K{:}U \to V$ gives us a construction of an internal model $\tilde{K}(\mathcal{M})$ of U from a model $\mathcal{M}$ of V. We find that U and V are bi-interpretable iff, there are interpretations $K{:}U \to V$ and $M{:}V \to U$ and formulas F and G such that, for all models $\mathcal{M}$ of V, the formula F defines an isomorphism between $\mathcal{M}$ and $\widetilde{M}\widetilde{K}(\mathcal{M})$, and, for all models $\mathcal{N}$ of U, the formula G defines an isomorphism between $\mathcal{N}$ and $\widetilde{K}\widetilde{M}(\mathcal{N})$.

- New definition of bi-interpretability to ensure contrast preservation across translations

---

Oakden 2020, Notational Equivalence in Tonal Geometry
Harvey & Visser 2014, When bi-interpretability implies synonymy

Comparing Phonological Representations

> *Our results show that ARs and QRs are equivalent with respect to their expressivity, but there are ways in which the two representational theories are conceptually distinct...Our article thus follows in the spirit of Kornai and Pullum (1990), who formally analyze X-bar theory in order to distinguish its true novel theoretical contributions from specious differences with context-free grammars...This thus highlights the value to phonological theory of the rigorous model-theoretic analysis of phonological structure.*

Jardine, Danis, & Iacoponi 2021, A Formal Investigation of Q-Theory in Comparison to Autosegmental Representations

COMPARING PHONOLOGICAL REPRESENTATIONS

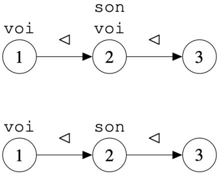
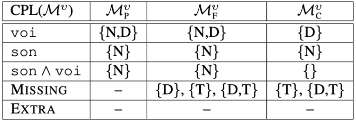
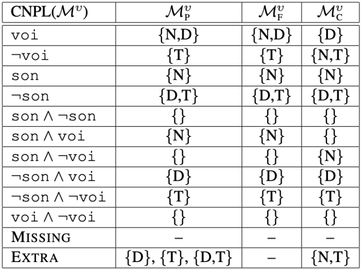


Figure 3: Models for the string DNT using models $\mathcal{M}^{v}_{\text{P}} = \mathcal{M}^{v}_{\text{F}}$ (top) and $\mathcal{M}^{v}_{\text{C}}$ (bottom)



CPL($\mathcal{M}^{v}$)	$\mathcal{M}^{v}_{\text{P}}$	$\mathcal{M}^{v}_{\text{F}}$	$\mathcal{M}^{v}_{\text{C}}$
voi	{N,D}	{N,D}	{D}
son	{N}	{N}	{N}
son ∧ voi	{N}	{N}	{}
MISSING	–	{D}, {T}, {D,T}	{T}, {D,T}
EXTRA	–	–	–



CNPL($\mathcal{M}^{v}$)	$\mathcal{M}^{v}_{\text{P}}$	$\mathcal{M}^{v}_{\text{F}}$	$\mathcal{M}^{v}_{\text{C}}$
voi	{N,D}	{N,D}	{D}
¬voi	{T}	{T}	{N,T}
son	{N}	{N}	{N}
¬son	{D,T}	{D,T}	{D,T}
son ∧ ¬son	{}	{}	{}
son ∧ voi	{N}	{N}	{}
son ∧ ¬voi	{}	{}	{N}
¬son ∧ voi	{D}	{D}	{D}
¬son ∧ ¬voi	{T}	{T}	{T}
voi ∧ ¬voi	{}	{}	{}
MISSING	–	–	–
EXTRA	{D}, {T}, {D,T}	–	{N,T}

Nelson 2022, A Model Theoretic Perspective on Phonological Feature Systems

COMPARING PHONOLOGICAL REPRESENTATIONS

> *The goal of this paper was not to find the correct feature system. Rather, the goal was to…better understand what the differences between each system are… For example, privative feature systems can be represented most simply as they minimally require univalent primitives and CPL logic. In order to describe a full feature system there needs to be either an increase in logical power (CNPL) or an increase in representational primitives (bivalent primitives). A contrastive feature system is the least flexible in how it can be represented as it requires CPL and bivalent primitives.*

Nelson 2022, A Model Theoretic Perspective on Phonological Feature Systems

CONNECTIONS TO DISTRIBUTED REPRESENTATIONS

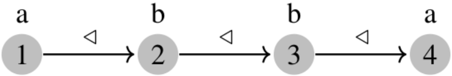
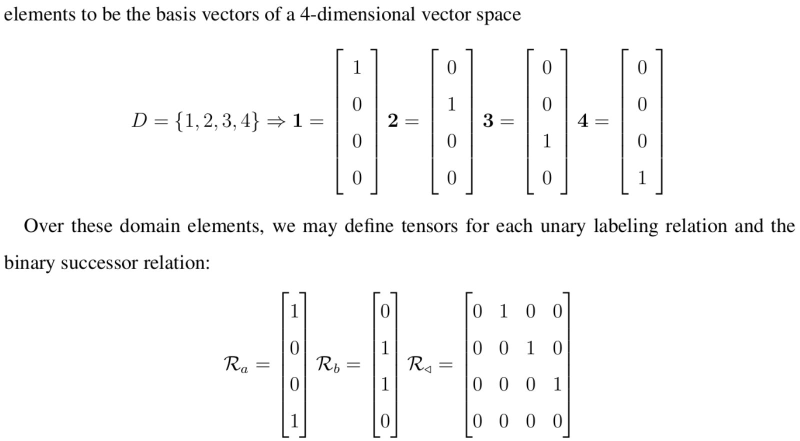


Figure 5.5: Successor Word model for *abba*

Rawski 2021:chap. 5: Structure and Learning in Natural Language
Sato 2017, Embedding Tarskian Semantics in Vector Spaces

Connections to Distributed Representations



elements to be the basis vectors of a 4-dimensional vector space

$$D = \{1, 2, 3, 4\} \Rightarrow \mathbf{1} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \mathbf{2} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \mathbf{3} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \mathbf{4} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

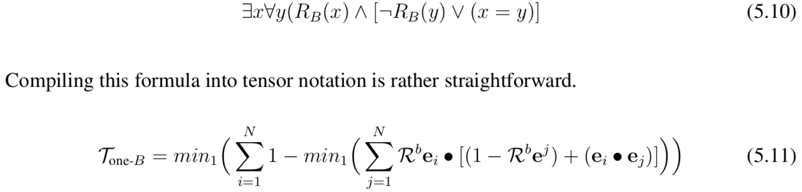
Over these domain elements, we may define tensors for each unary labeling relation and the binary successor relation:

$$\mathcal{R}_a = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \mathcal{R}_b = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \mathcal{R}_{\triangleleft} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Rawski 2021:chap. 5: Structure and Learning in Natural Language
Sato 2017, Embedding Tarskian Semantics in Vector Spaces

CONNECTIONS TO DISTRIBUTED REPRESENTATIONS

Exactly-one B



$$\exists x \forall y (R_B(x) \wedge [\neg R_B(y) \vee (x = y)]) \tag{5.10}$$

Compiling this formula into tensor notation is rather straightforward.

$$\mathcal{T}_{\text{one-}B} = min_1\Big(\sum_{i=1}^{N} 1 - min_1\Big(\sum_{j=1}^{N} \mathcal{R}^b \mathbf{e}_i \bullet [(1 - \mathcal{R}^b \mathbf{e}^j) + (\mathbf{e}_i \bullet \mathbf{e}_j)]\Big)\Big) \tag{5.11}$$

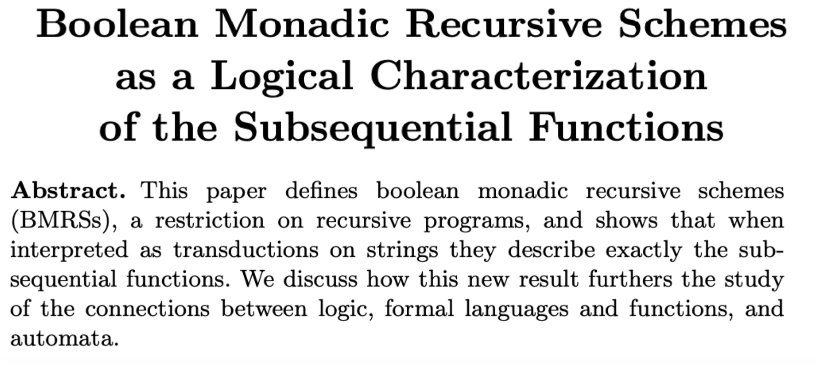
Rawski 2021:chap. 5: Structure and Learning in Natural Language
Sato 2017, Embedding Tarskian Semantics in Vector Spaces

CONNECTIONS TO DISTRIBUTED REPRESENTATIONS

Model-theoretic descriptions of relational structures were embedded in Euclidean vector spaces, and statements in first-order logic over these structures were compiled into tensor formulas. Semantic evaluation was given via tensor contraction over tensors implementing a specific model. This method can easily be extended to consider other relational structures, and to other logics.

Rawski 2021:chap. 5: Structure and Learning in Natural Language
Sato 2017, Embedding Tarskian Semantics in Vector Spaces

BOOLEAN MONADIC RECURSIVE SCHEMES

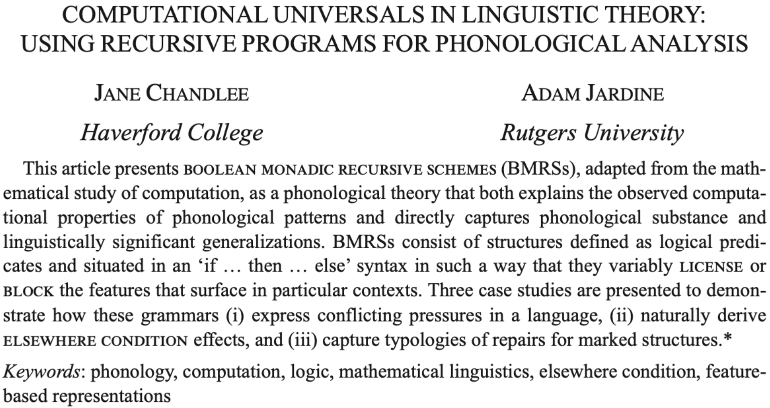


Boolean Monadic Recursive Schemes as a Logical Characterization of the Subsequential Functions

Abstract. This paper defines boolean monadic recursive schemes (BMRSs), a restriction on recursive programs, and shows that when interpreted as transductions on strings they describe exactly the subsequential functions. We discuss how this new result furthers the study of the connections between logic, formal languages and functions, and automata.

Bhaskar, Chandlee, Jardine & Oakden 2020

Boolean Monadic Recursive Schemes



COMPUTATIONAL UNIVERSALS IN LINGUISTIC THEORY: USING RECURSIVE PROGRAMS FOR PHONOLOGICAL ANALYSIS

Jane Chandlee
Haverford College

Adam Jardine
Rutgers University

This article presents boolean monadic recursive schemes (BMRSs), adapted from the mathematical study of computation, as a phonological theory that both explains the observed computational properties of phonological patterns and directly captures phonological substance and linguistically significant generalizations. BMRSs consist of structures defined as logical predicates and situated in an ‘if … then … else’ syntax in such a way that they variably license or block the features that surface in particular contexts. Three case studies are presented to demonstrate how these grammars (i) express conflicting pressures in a language, (ii) naturally derive elsewhere condition effects, and (iii) capture typologies of repairs for marked structures.*

Keywords: phonology, computation, logic, mathematical linguistics, elsewhere condition, feature-based representations

---

2021, *Language*

ALGEBRAIC CHARACTERIZATIONS

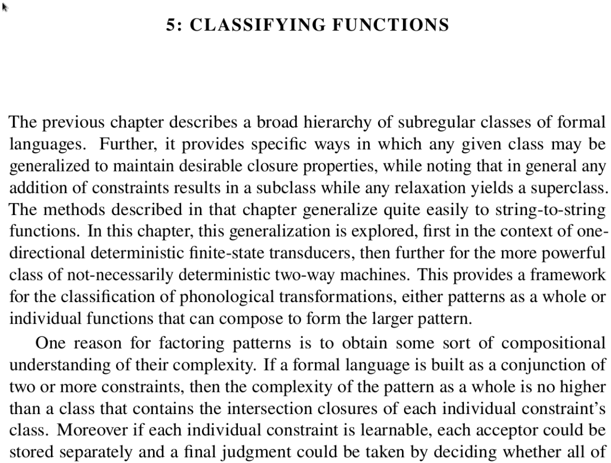
These:

- are based on the syntactic monoid of a finite-state machine
- characterize the behaviors of sequential input
- provide a unified way to classify string sets, functions and relations (including 1-way/2-way and deterministic/non-deterministic machines),

(Subsequential functions and regular string sets have canonical forms, but non-subsequential regular relations do not.)

Lambert 2022, Unifying Classification Schemes for Languages and Processes With Attention to Locality and Relativizations Thereof

Algebraic Advantages



5: CLASSIFYING FUNCTIONS

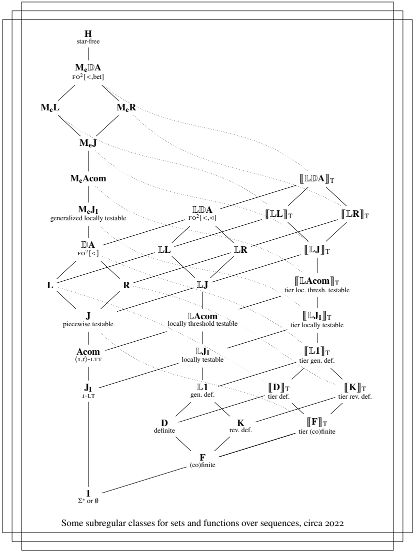
The previous chapter describes a broad hierarchy of subregular classes of formal languages. Further, it provides specific ways in which any given class may be generalized to maintain desirable closure properties, while noting that in general any addition of constraints results in a subclass while any relaxation yields a superclass. The methods described in that chapter generalize quite easily to string-to-string functions. In this chapter, this generalization is explored, first in the context of one-directional deterministic finite-state transducers, then further for the more powerful class of not-necessarily deterministic two-way machines. This provides a framework for the classification of phonological transformations, either patterns as a whole or individual functions that can compose to form the larger pattern.

One reason for factoring patterns is to obtain some sort of compositional understanding of their complexity. If a formal language is built as a conjunction of two or more constraints, then the complexity of the pattern as a whole is no higher than a class that contains the intersection closures of each individual constraint's class. Moreover if each individual constraint is learnable, each acceptor could be stored separately and a final judgment could be taken by deciding whether all of

---

Lambert 2022:chap. 5, Unifying Classification Schemes for Languages and Processes With Attention to Locality and Relativizations Thereof

AN ALGEBRAIC HIERARCHY OF APERIODIC PATTERNS (LAMBERT 2022)



Some subregular classes for sets and functions over sequences, circa 2022

HIGHLIGHTS

- Total Reduplication is 1, along with affixation.
- ISL functions are Definite functions.
- OSL functions are not an algebraic class (every algebraic class shown includes some OSL function).
- Spreading is Tier Definite.
- Long-distance Harmony is Tier Reverse Definite.
- Tutrugbu vowel harmony (McCollum et al. 2021) is Tier-based Locally Testable.
- All the algebraic classes are closed under Boolean operations and direct product (but not composition)

Lambert 2022:chap. 5, Unifying Classification Schemes for Languages and Processes With Attention to Locality and Relativizations Thereof

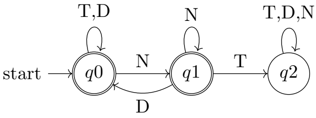
RELATIONSHIP BETWEEN LOGIC, FINITE STATE MACHINES, AND ALGEBRA

"No voiceless stop after a nasal"

Logic:

$$\neg\exists x, y\Big[x \lhd y \wedge N(x) \wedge T(y)\Big]$$

Finite State:



Algebra:

·	λ	T	N	D	TN	NT
λ	λ	T	N	D	TN	NT
T	T	T	TN	T	TN	NT
N	N	NT	N	D	NT	NT
D	D	D	N	D	N	NT
TN	TN	NT	TN	T	NT	NT
NT	NT	NT	NT	NT	NT	NT

Future

THE FUTURE IS NOW

- Software tools
- Classes of representations
- Other logical languages
- More interfaces
- Factoring patterns algebraically
- Learning grammars in a variety of scenarios
- Automatic phonological analysis
- Documentation/field work
- …

THE END OF TIME

Thank You