Modeling Long-distance Alternations with Tier-based Strictly Local Functions

Jane Chandlee, Jeffrey Heinz, Adam Jardine, and Kevin McMullin

Haverford College University of Delaware Rutgers University University of Ottawa

LSA Annual Meeting
Austin, TX
January 7, 2016

Introduction

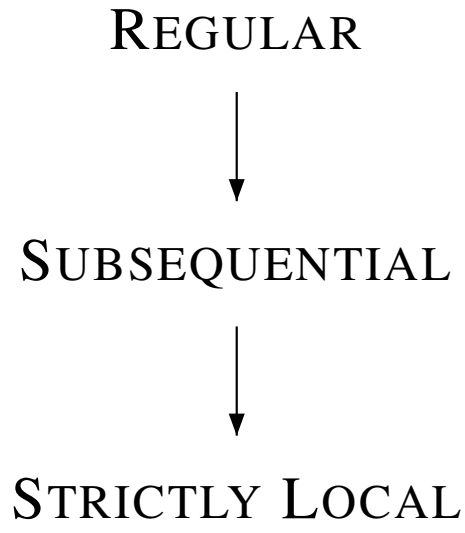
- Previous work has established that 'local' alternations are describable with *Strictly Local functions* (Chandlee, 2014).
  - Computationally restrictive and learnable (Chandlee et al., 2014, 2015a; Jardine et al., 2014).
- Long-distance *phonotactics* are describable with *Tier-based Strictly Local* languages (Heinz et al., 2011; McMullin, 2016).
- We define *Tier-based Strictly Local* ***functions*** and use them to model long-distance alternations.

Maps versus formal languages

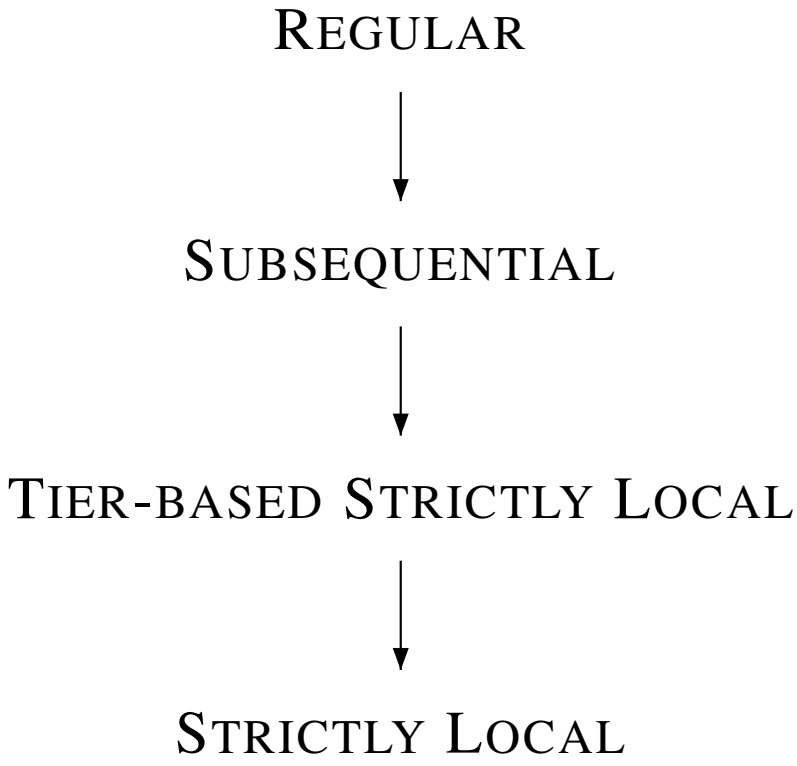
Formal Language	Map
{at, ap, am, ada, ...}	{(ad, at), (ab, ap), (am, am), (ada, ada)...}

- A formal language is a set of well-formed strings, according to some constraint(s) or grammar.
- Maps are functions that map input strings to output strings.
- Analogous to the distinction between phonotactics and the UR to SR transformation.

Hierarchy of maps



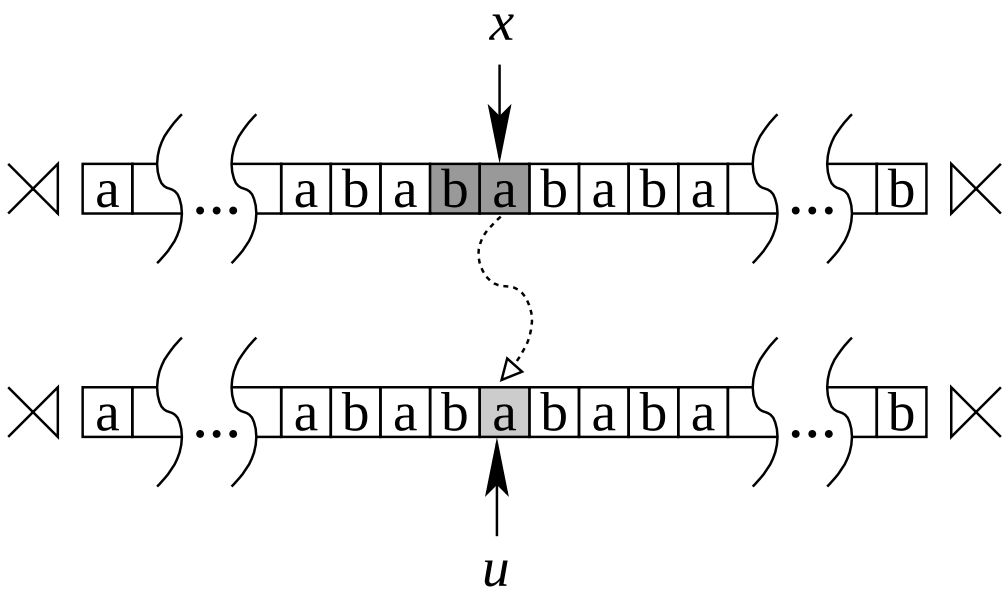
Hierarchy of maps



Strictly Local functions

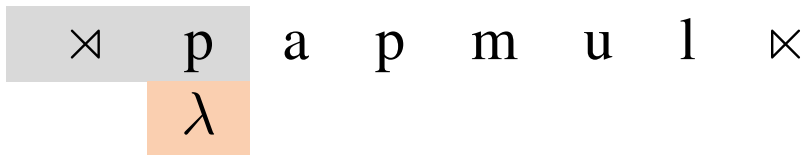
- Output string is computed *locally*, depending only on a bounded number of previous segments.
- Two varieties:
  - Input Strictly Local: output depends on previous input symbols
  - Output Strictly Local: output depends on previous output symbols

Input Strictly Local



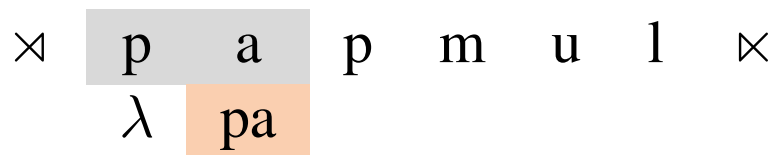
Example: Input Strictly Local (ISL)

(1) Korean (Lee and Pater, 2008)
/pa**pm**ul/ $\mapsto$ [pa**mm**ul] ‘rice water’



Example: Input Strictly Local (ISL)

(1) Korean (Lee and Pater, 2008)
/pa**pm**ul/ $\mapsto$ [pa**mm**ul] ‘rice water’



Example: Input Strictly Local (ISL)

(1) Korean (Lee and Pater, 2008)
/pa**pm**ul/ $\mapsto$ [pa**mm**ul] ‘rice water’

$\rtimes$	p	a	p	m	u	l	$\ltimes$
	λ	pa	λ				

Example: Input Strictly Local (ISL)

(1) Korean (Lee and Pater, 2008)
/pa**pm**ul/ $\mapsto$ [pa**mm**ul] 'rice water'

$\rtimes$	p	a	p	m	u	l	$\ltimes$
	λ	pa	λ	mm			

Example: Input Strictly Local (ISL)

(1) Korean (Lee and Pater, 2008)
/pa**pm**ul/ $\mapsto$ [pa**mm**ul] ‘rice water’

$\rtimes$	p	a	p	m	u	l	$\ltimes$
	λ	pa	λ	mm	u		

Example: Input Strictly Local (ISL)

(1) Korean (Lee and Pater, 2008)
/pa**pm**ul/ $\mapsto$ [pa**mm**ul] ‘rice water’

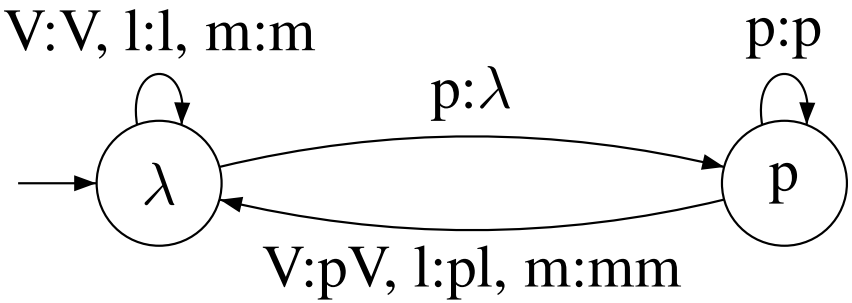
$\rtimes$	p	a	p	m	u	l	$\ltimes$
	λ	pa	λ	mm	u	l	

Example: Input Strictly Local (ISL)

(1) Korean (Lee and Pater, 2008)
/pa**pm**ul/ $\mapsto$ [pa**mm**ul] ‘rice water’

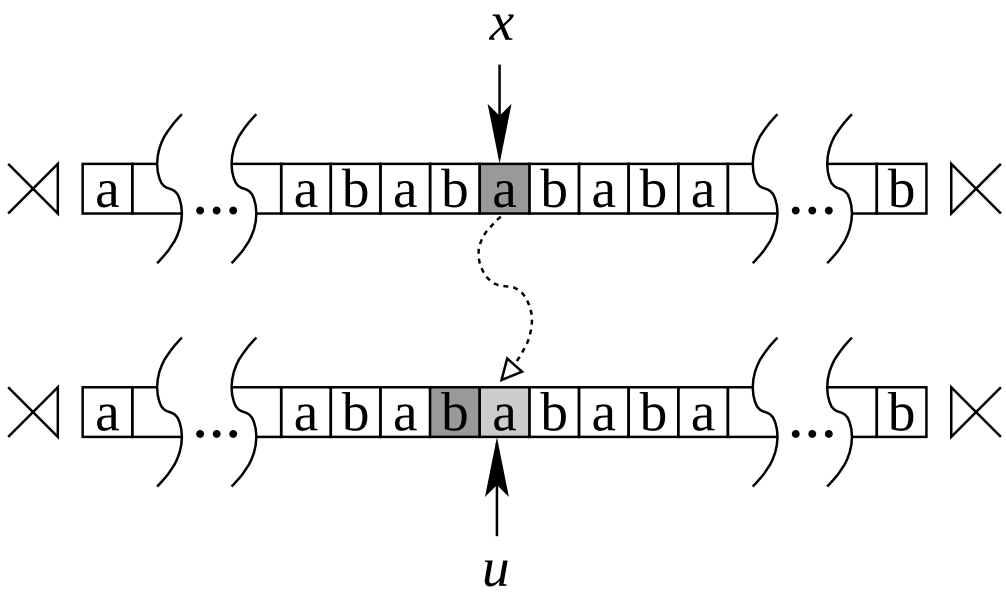
Window size is 2: this map is 2-ISL.

Input Strictly Local (ISL): FST characterization



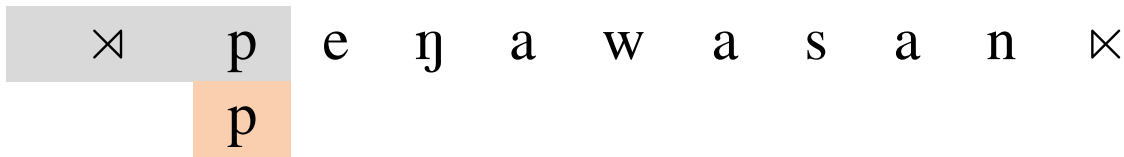
For a 2-ISL map, the states are of length ≤ 1.

Output Strictly Local



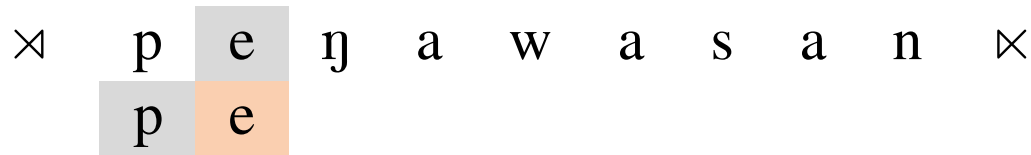
Example: Output Strictly Local (OSL)

(2) Johore Malay (Onn, 1980)
/pəŋawasan/ ↦ [pəŋãwãsan] ‘supervision’



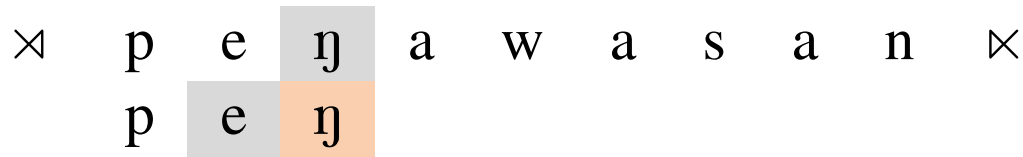
Example: Output Strictly Local (OSL)

(2) Johore Malay (Onn, 1980)
/pəŋawasan/ ↦ [pəŋãwãsan] ‘supervision’



Example: Output Strictly Local (OSL)

(2) Johore Malay (Onn, 1980)
/pəŋawasan/ ↦ [pəŋãwãsan] ‘supervision’



Example: Output Strictly Local (OSL)

(2) Johore Malay (Onn, 1980)
/pəŋawasan/ ↦ [pəŋãwãsan] ‘supervision’

⋊	p	e	ŋ	a	w	a	s	a	n	⋉
	p	e	ŋ	ã						

Example: Output Strictly Local (OSL)

(2) Johore Malay (Onn, 1980)
/pəŋawasan/ ↦ [pəŋãwãsan] ‘supervision’

⋊	p	e	ŋ	a	w	a	s	a	n	⋉
	p	e	ŋ	ã	w̃					

Example: Output Strictly Local (OSL)

(2) Johore Malay (Onn, 1980)
/pəŋawasan/ ↦ [pəŋãw̃ãsan] ‘supervision’

⋊	p	e	ŋ	a	w	a	s	a	n	⋉
	p	e	ŋ	ã	w̃	ã				

Example: Output Strictly Local (OSL)

(2) Johore Malay (Onn, 1980)
/pəŋawasan/ ↦ [pəŋãwãsan] ‘supervision’

⋊	p	e	ŋ	a	w	a	s	a	n	⋉
	p	e	ŋ	ã	w̃	ã	s			

Example: Output Strictly Local (OSL)

(2) Johore Malay (Onn, 1980)
/pəŋawasan/ ↦ [pəŋãwãsan] ‘supervision’

⋊	p	e	ŋ	a	w	a	s	a	n	⋉
	p	e	ŋ	ã	w̃	ã	s	a		

Example: Output Strictly Local (OSL)

(2) Johore Malay (Onn, 1980)
/pəŋawasan/ ↦ [pəŋãwãsan] ‘supervision’

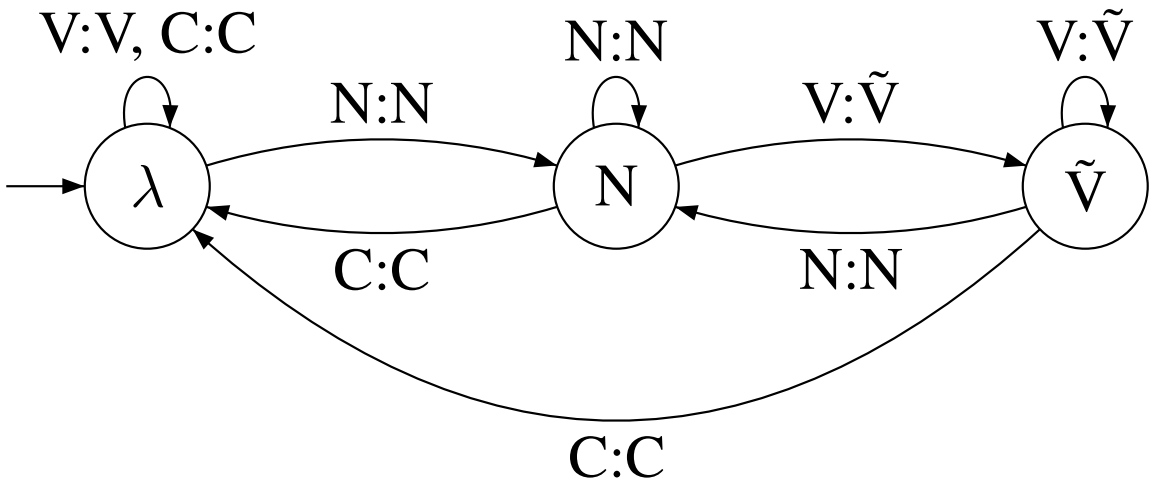
⋊	p	e	ŋ	a	w	a	s	a	n	⋉
	p	e	ŋ	ã	w̃	ã	s	a	n	

Example: Output Strictly Local (OSL)

(2) Johore Malay (Onn, 1980)
/pəŋawasan/ $\mapsto$ [pəŋãwãsan] ‘supervision’

Window size is 2: this map is 2-OSL.

Output Strictly Local (OSL): FST characterization



Summary: ISL and OSL

- ISL: processes with bounded conditioning environments (Chandlee, 2014; Chandlee et al., 2014)
  - Assimilation, dissimilation, deletion, insertion, metathesis
  - Over 90% of rules in P-base (Mielke, 2008) are ISL
  - A variety of process interactions, including opaque maps, are ISL (Chandlee et al., 2015b).
- OSL: processes with bounded conditioning environments on the output side (Chandlee et al., 2015a)
  - iterative spreading
- Neither ISL nor OSL can model long-distance alternations.

Long-distance (unbounded) assimilation

Kikongo (Meinof, 1932; Odden, 1994; Rose and Walker, 2004)

a. /tu**n**iki**d**i/ $\mapsto$ [tu**n**iki**n**i] ‘we ground’

b. /kudu**m**ukisi**l**a/ $\mapsto$ [kudu**m**ukisi**n**a] ‘to cause to jump for’

Long-distance (unbounded) assimilation

Kikongo (Meinof, 1932; Odden, 1994; Rose and Walker, 2004)

a. /tu**n**iki**d**i/ ↦ [tu**n**iki**n**i] ‘we ground’

b. /kudu**m**ukisi**l**a/ ↦ [kudu**m**ukisi**n**a] ‘to cause to jump for’

⋊	t	u	n	i	k	i	d	i	⋉
	t	u	n	i	k	i	n		

⋊	k	u	d	u	m	u	k	i	s	i	l	a	⋉
	k	u	d	u	m	u	k	i	s	i	l		

Long-distance (unbounded) assimilation

Kikongo (Meinof, 1932; Odden, 1994; Rose and Walker, 2004)

a. /tu**n**iki**d**i/ $\mapsto$ [tu**n**iki**n**i] ‘we ground’

b. /kudu**m**ukisi**l**a/ $\mapsto$ [kudu**m**ukisi**n**a] ‘to cause to jump for’

LDA is not SL for any k.

Two approaches to long-distance phenomena

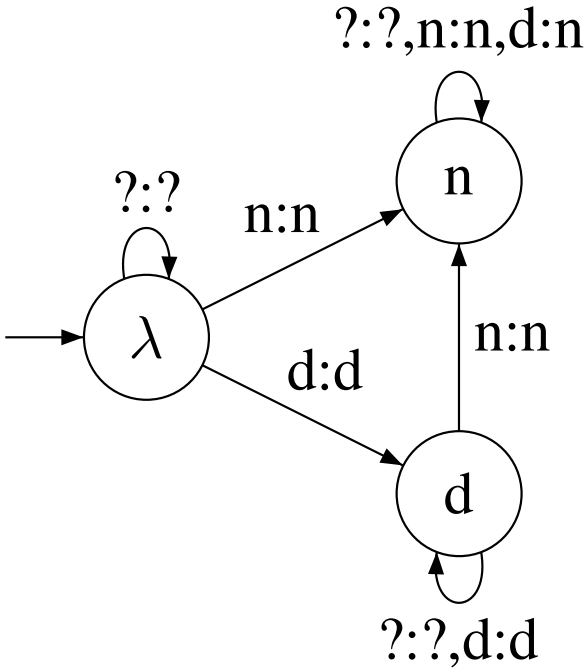
- Strictly Piecewise (Heinz, 2010): precedence instead of succesor
- Tier-based Strictly Local (Heinz et al., 2011; McMullin, 2016): SL over a subset of the alphabet (the 'tier')
- Both have been applied to the modeling of LD phonotactics.
- To model LD alternations (or maps), here we take the latter approach and define TSL *functions*.

TSL functions: FST characterization

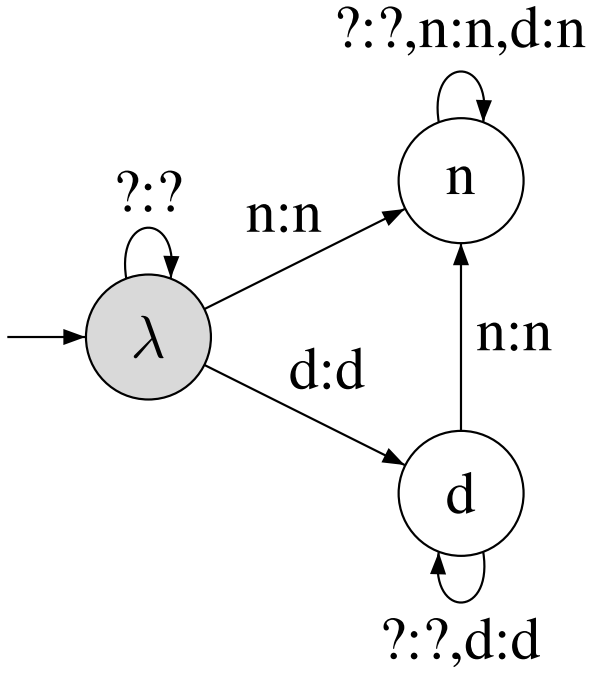
- Given an alphabet Σ, we designate a subset of the alphabet as the *tier*, T.
- The states of the FST are all possible sequences of segments in T up to length $k - 1$.
- All transitions for non-tier segments ($\Sigma - T$) are self-loops.

LDA as TSL

- $T = \{\text{d}, \text{n}\}$
- ? abbreviates all non-tier segments

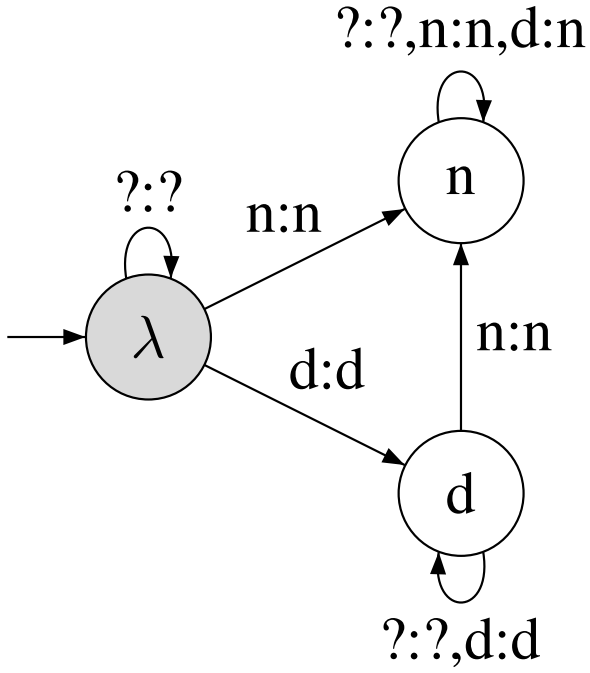


LDA as TSL



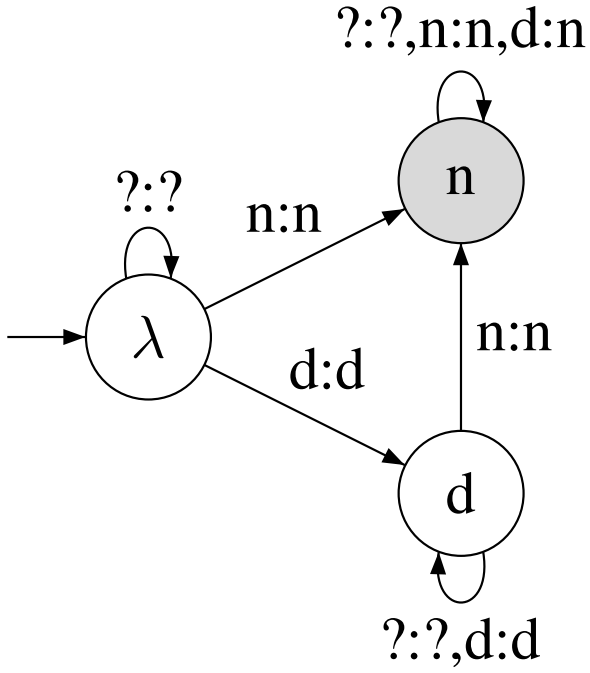
⋊	t	u	n	i	k	i	d	i	⋉
	t								

LDA as TSL



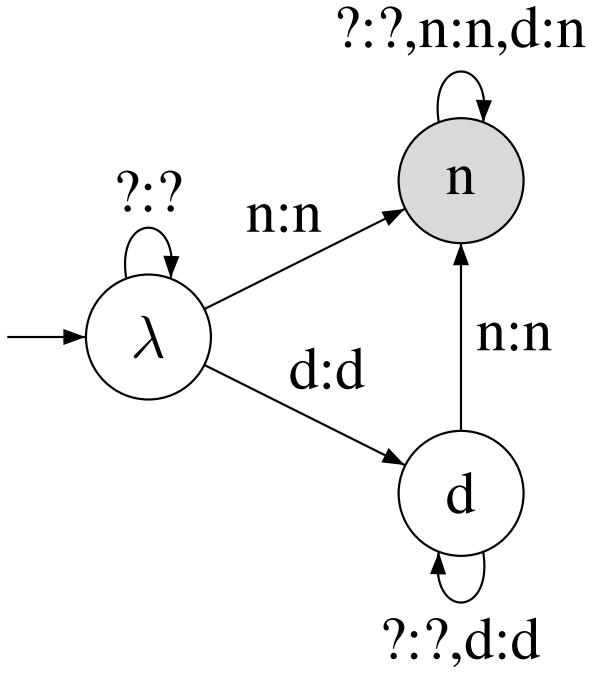
⋊	t	u	n	i	k	i	d	i	⋉
	t	u							

LDA as TSL



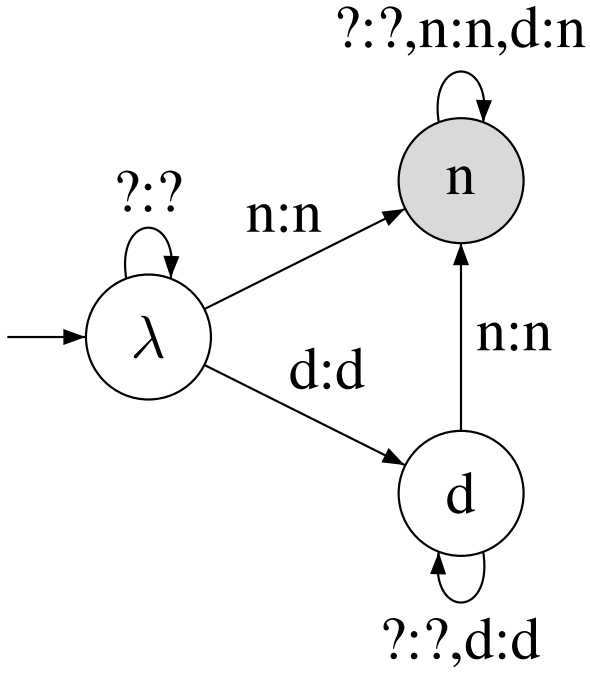
⋊	t	u	n	i	k	i	d	i	⋉
	t	u	n						

LDA as TSL



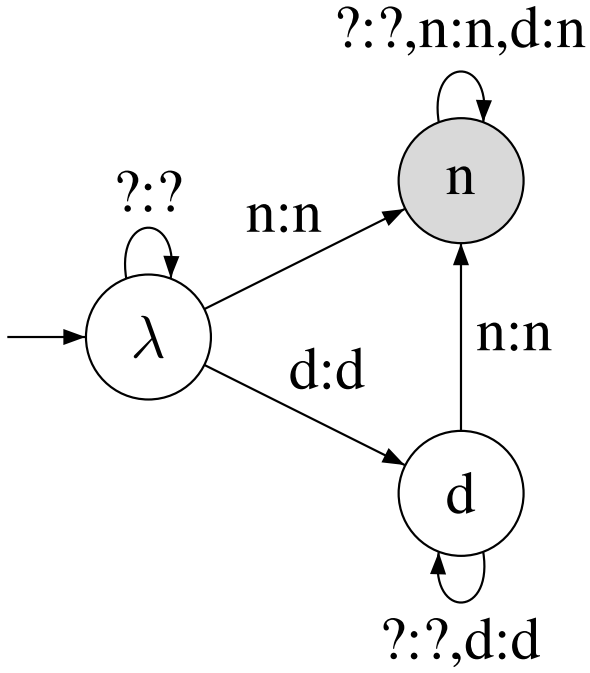
⋊	t	u	n	i	k	i	d	i	⋉
	t	u	n	i					

LDA as TSL



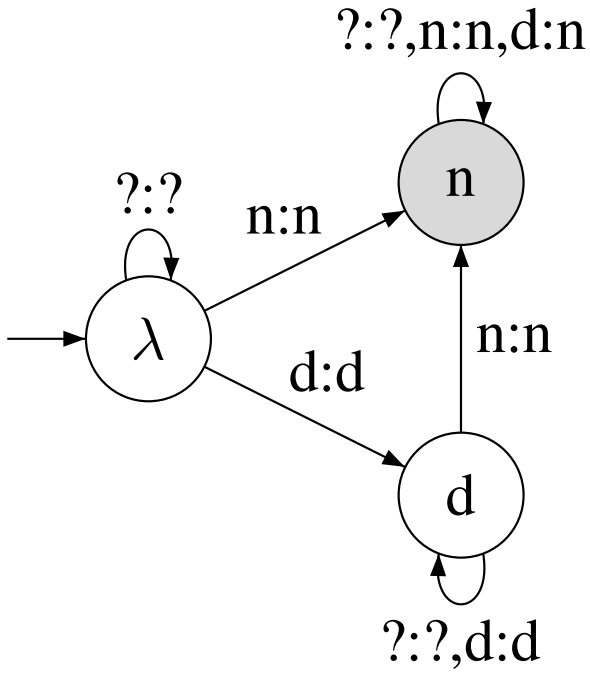
⋊	t	u	n	i	k	i	d	i	⋉
	t	u	n	i	k				

LDA as TSL



⋊	t	u	n	i	k	i	d	i	⋉
	t	u	n	i	k	i			

LDA as TSL



⋊	t	u	n	i	k	i	d	i	⋉
	t	u	n	i	k	i	n		

LDA as TSL

- FST has states of length 1; so the map is 2-TSL.
- The transitions follow the output, so it is *Output TSL*.

LDA with blocking as TSL

Slovenian (Jurgec, 2011; McMullin, 2016)

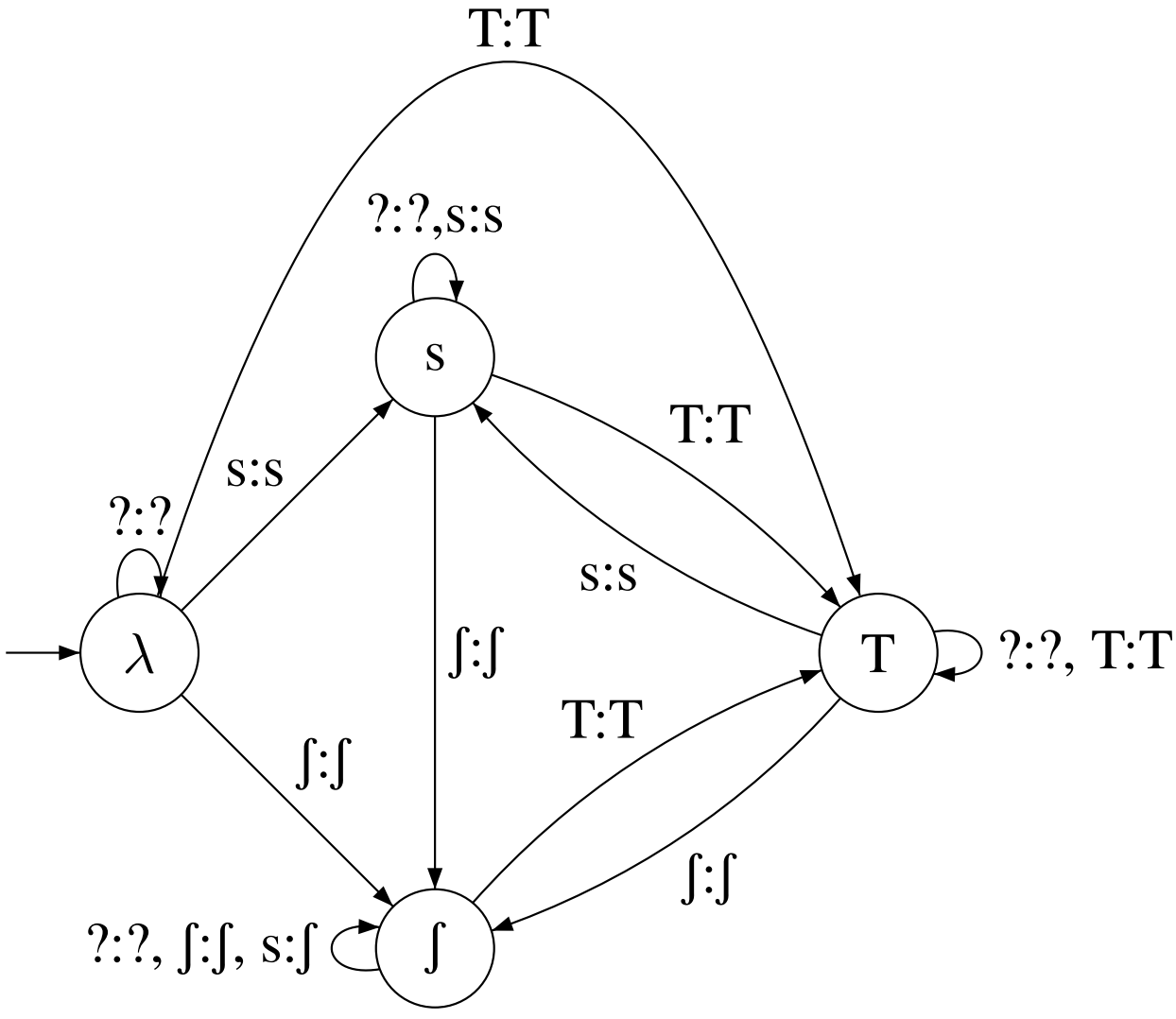
a. **s**piʃ ↦ ʃpiʃ ‘(you) sleep’

b. **z**aklɔniʃtʃe ↦ **ʒ**aklɔniʃtʃe ‘bomb shelter’

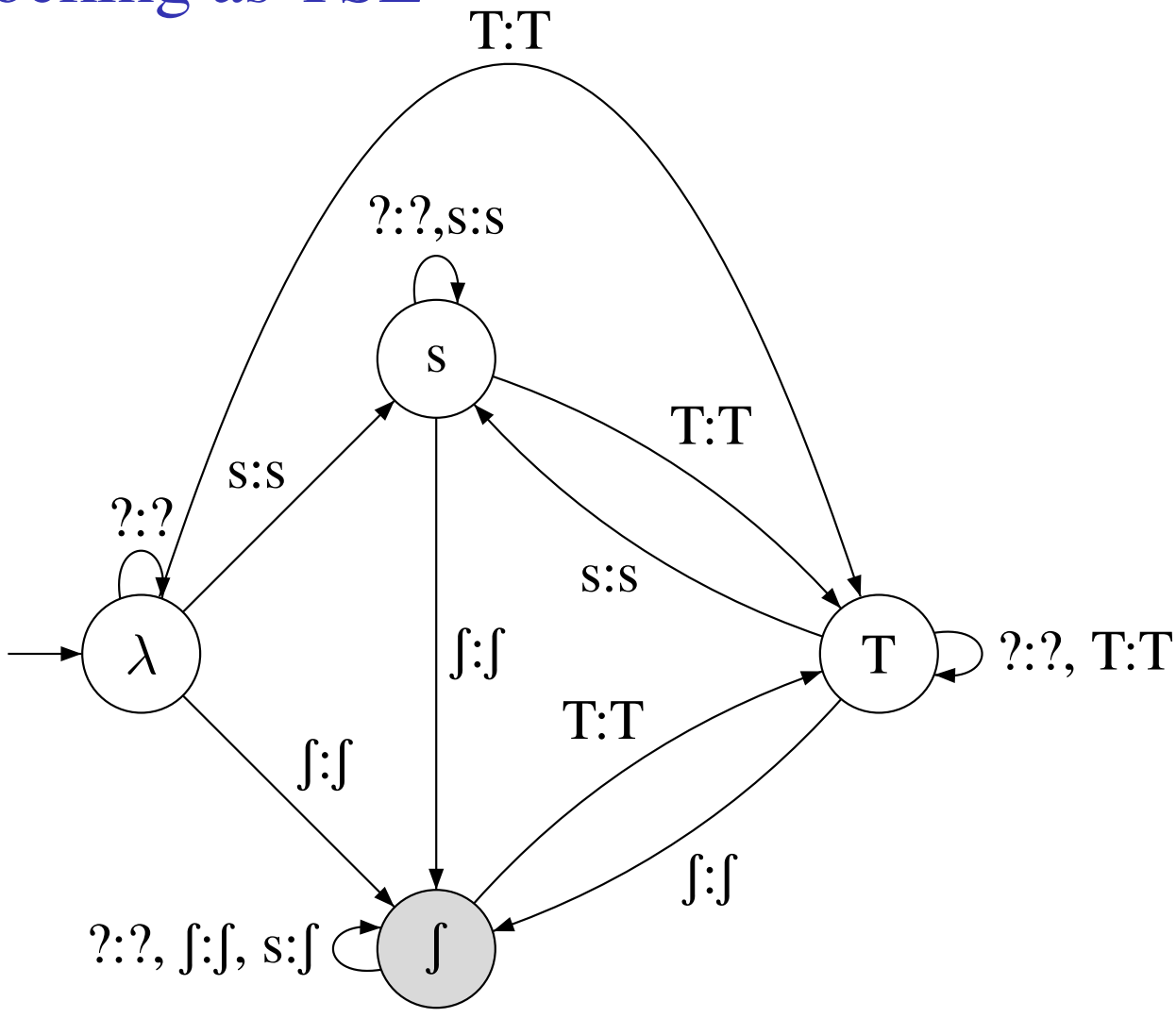
c. na**s**iti**ʃ** ↦ na**s**iti**ʃ** ‘(you) feed’

- The blocking segments go on the tier.

LDA with blocking as TSL



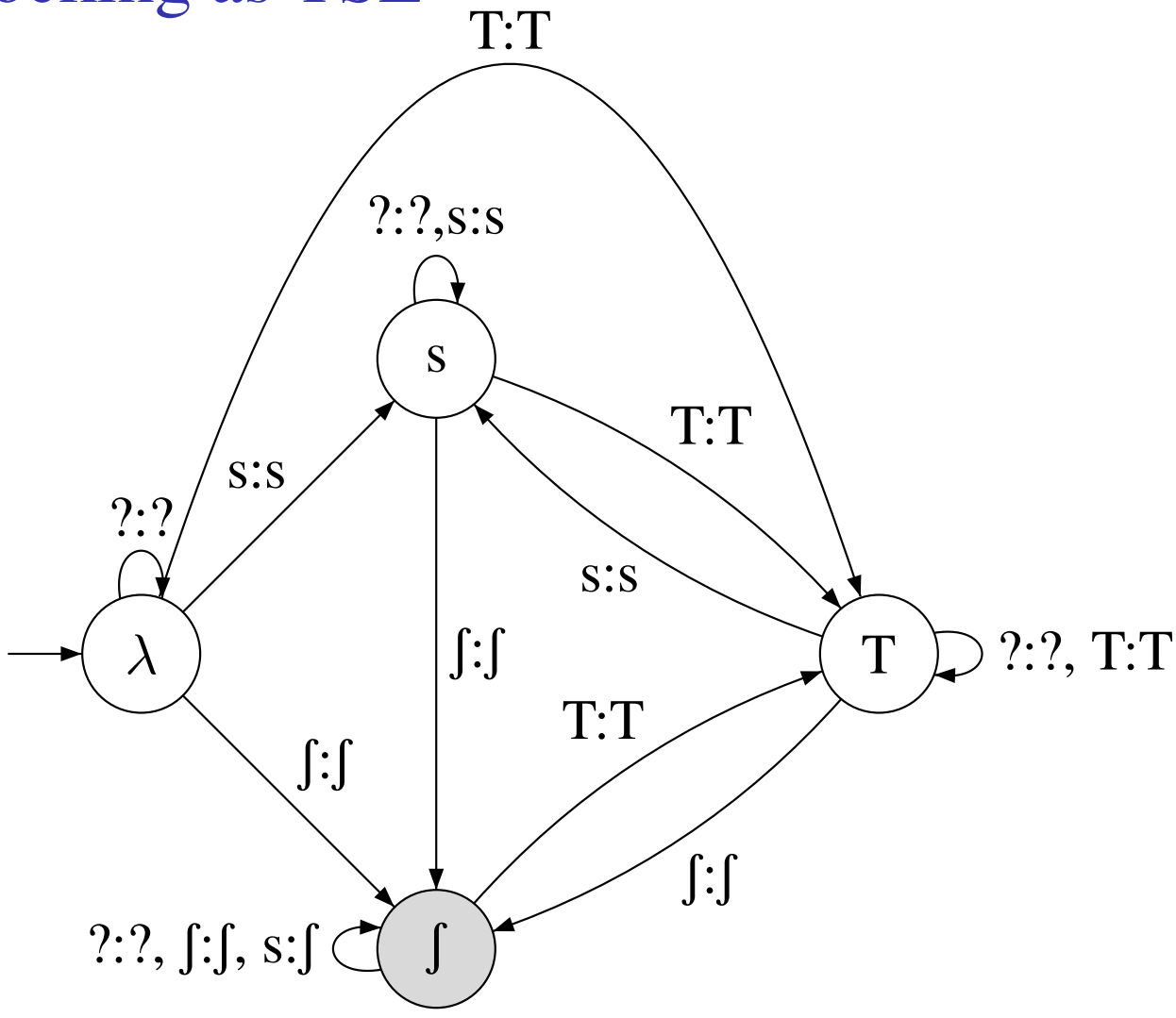
LDA with blocking as TSL



s p i ∫

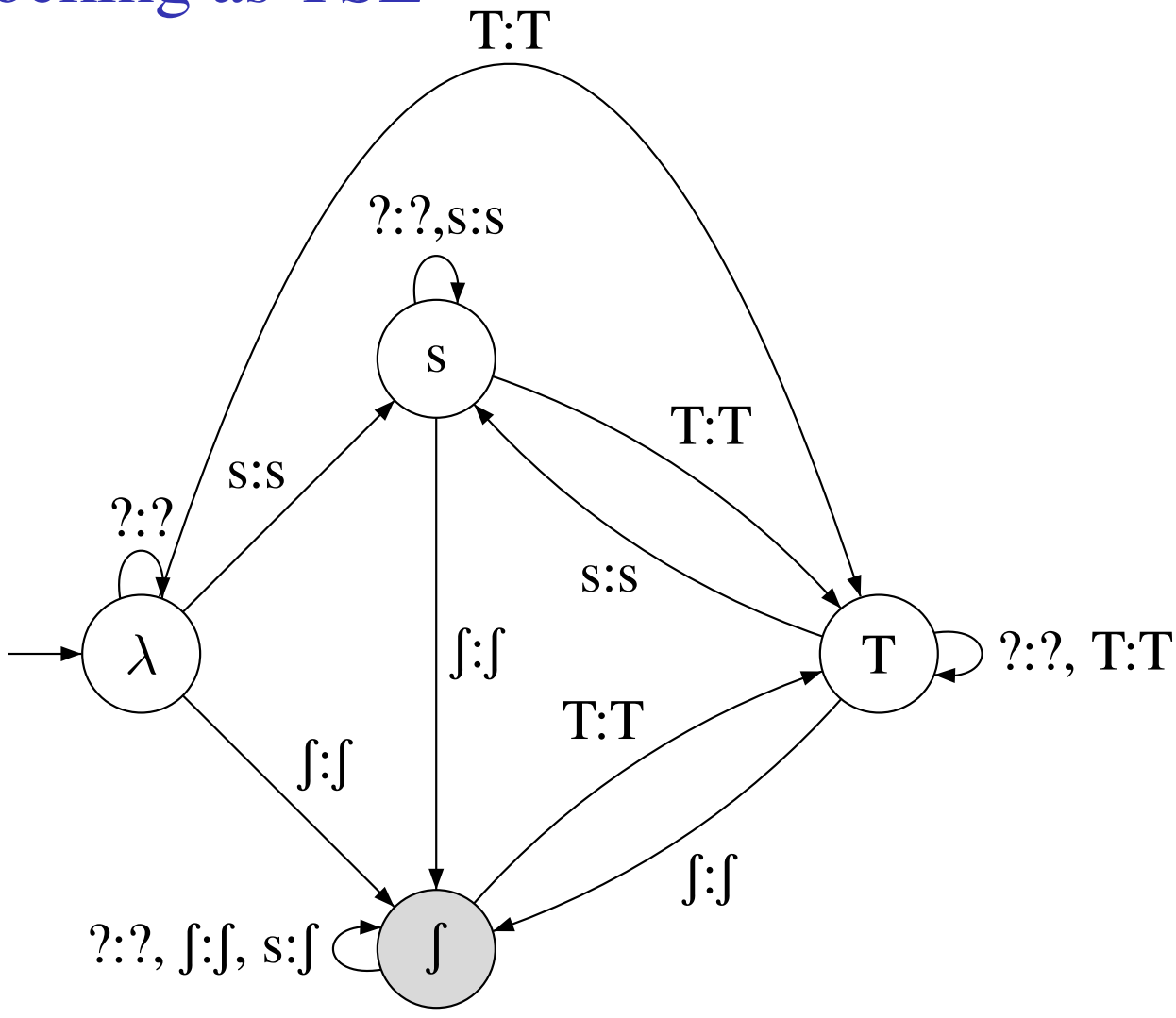
∫

LDA with blocking as TSL



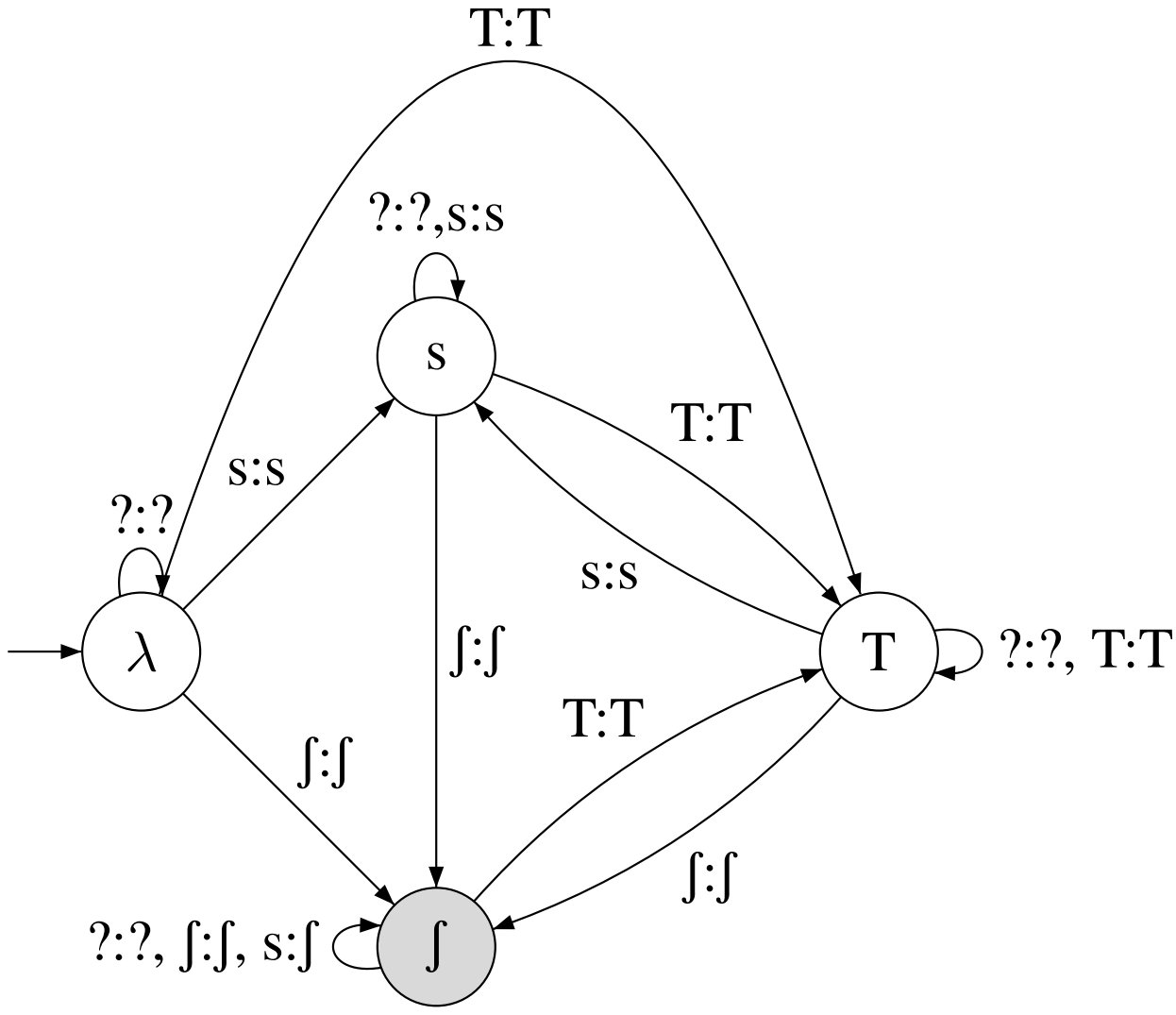
s	p	i	∫
		i	∫

LDA with blocking as TSL



s	p	i	∫
	p	i	∫

LDA with blocking as TSL



s	p	i	∫
∫	p	i	∫

Long-distance dissimilation with blocking

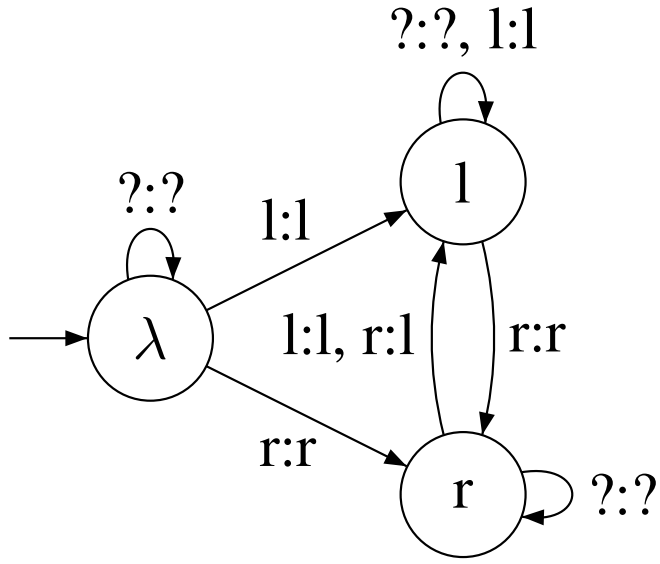
Georgian (Fallon, 1993; Odden, 1994; Bennett, 2013)

a. dan-u**r**i $\mapsto$ dan-u**r**i ‘Danish’

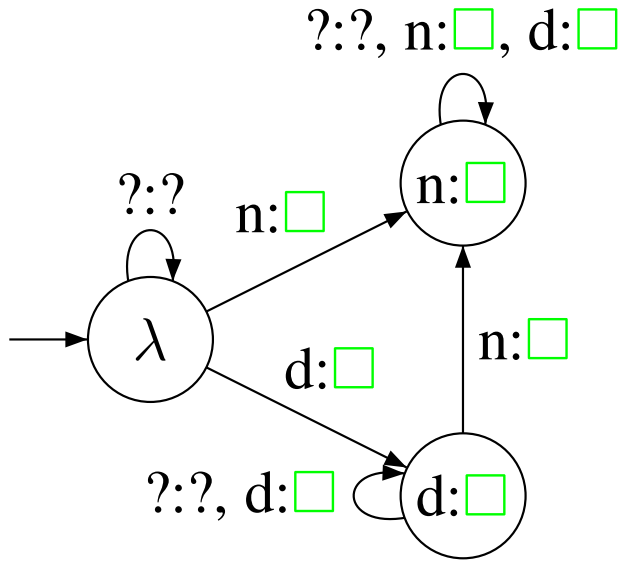
b. ung**r**-u**r**i $\mapsto$ ung**r**-u**l**i ‘Hungarian’

c. ka**rtl**-u**r**i $\mapsto$ ka**rtl**-u**r**i ‘Kartvelian’

Long-distance dissimilation with blocking



Learnability

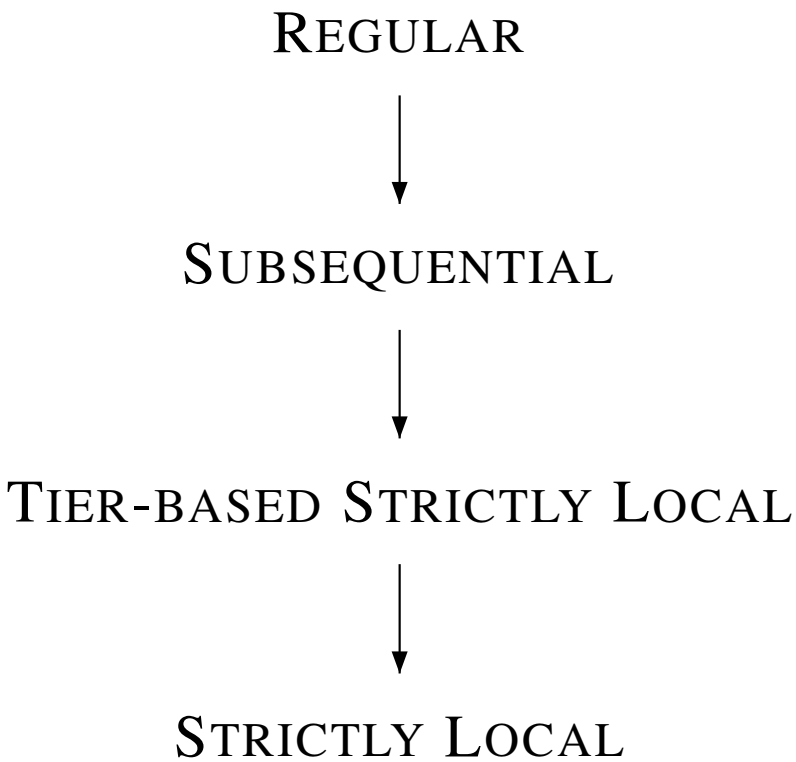


Output-empty FST for LD nasal assimilation

- If the tier is known *a priori*, maps can be learned from the structure of the FST (Chandlee et al., 2015a).
- To learn tiers, we may be able to draw from earlier results on learning TSL languages without a priori knowledge of the tier (Jardine and Heinz, 2016; Jardine and McMullin, to appear).

Conclusion

- The goal is to identify the computational classes that best characterize phonological maps.
- TSL functions provide a good fit to the range of LD maps, and their inherent structure yields learnability results.

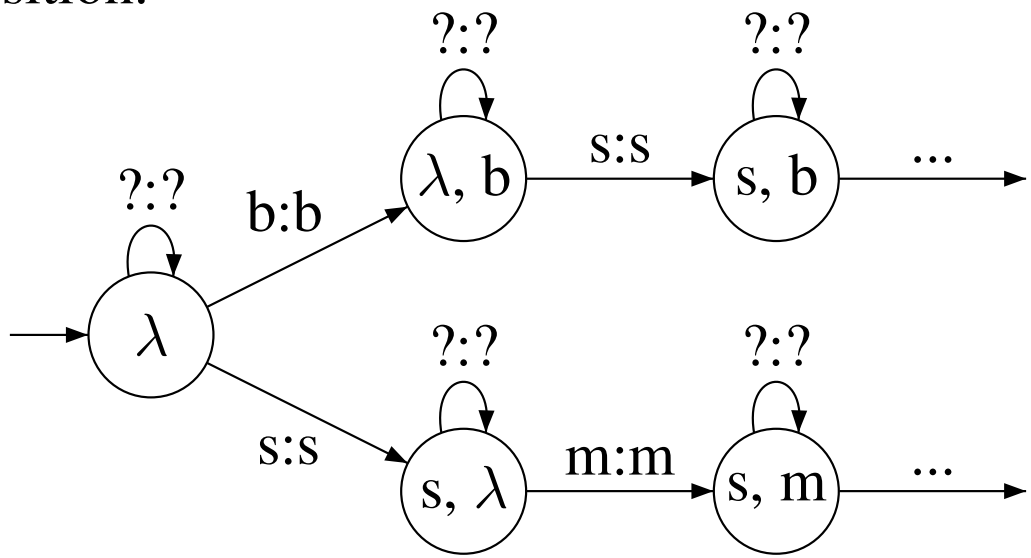


Multiple patterns

- What if more than one pattern exists in a language?
  - Different restrictions enforced on different tiers
- For phonotactics, conjunctions of TSL languages can help in some cases. . .
- . . . but not if one pattern overrides another.
- However, in extending the Tier-based approach to *maps*, standard functional composition is successful in several interesting cases

Co-existing long-distance maps

- If two patterns operate on disjoint tiers. . .
  - e.g., Tamashek Tuareg has unbounded regressive sibilant harmony and labial dissimilation (Heath, 2005; Bennett, 2013)
- . . . then these maps can be combined with standard functional composition.



Co-existing long-distance maps

- Distinct maps operating on the same (or overlapping) tiers?
- Hard to come by, but consider (cf. Bennett 2015, McMullin 2016):
  - liquid assimilation in a transvocalic context combined with unbounded liquid dissimilation
  - rVl $\mapsto$ rVr
  - rVVVr $\mapsto$ rVVVl
  - rVlVVVr $\mapsto$??
- Prediction: such patterns are possible to the extent that they can be described as the composition of two or more TSL functions.

References I

Bennett, W. (2013). *Dissimilation, Consonant Harmony, and Surface Correspondence*. PhD thesis, Rutgers.

Chandlee, J. (2014). *Strictly Local Phonological Processes*. PhD thesis, University of Delaware.

Chandlee, J., Eyraud, R., and Heinz, J. (2015a). Output strictly local functions. In *Proceedings of the 14th Meeting on the Mathematics of Language (MOL 2015)*.

Chandlee, J., Heinz, J., and Eyraud, R. (2014). Learning strictly local subsequential functions. *Transactions of the Association for Computational Linguistics*, 2:491–503.

Chandlee, J., Heinz, J., and Jardine, A. (2015b). Representing and learning opaque maps with strictly local functions. Workshop on The Implications of Computation and Learnability for Phonological Theory. GLOW 38. Paris, France, April 18.

References II

Fallon, P. D. (1993). Liquid dissimilation in georgian. In Kathol, A. and Bernstein, M., editors, *Proceedings of the 10th Eastern States Conference on Linguistics*, pages 105–116. Ithaca, NY: DMLL Publications.

Heath, J. (2005). *A grammar of Tamashek (Tuareg of Mali)*. Mouton de Gruyter, Berlin.

Heinz, J. (2010). Learning long-distance phonotactics. *Linguistic Inquiry*, 41(4):623–661.

Heinz, J., Rawal, C., and Tanner, H. G. (2011). Tier-based Strictly Local constraints for phonology. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, pages 58–64, Portland, Oregon, USA. Association for Computational Linguistics.

References III

Jardine, A., Chandlee, J., Eyraud, R., and Heinz, J. (2014). Very efficient learning of structured classes of subsequential functions from positive data. In *Proceedings of the 12th International Conference on Grammatical Inference (ICGI 2014)*, JMLR Workshop Proceedings, pages 94–108.

Jardine, A. and Heinz, J. (2016). Learning tier-based strictly 2-local languages. *Transactions of the Association for Computational Linguistics*, 4:87–98.

Jardine, A. and McMullin, K. (to appear). Efficient learning of tier-based strictly k-local languages. In Drewes, F., Martín-Vide, C., and Truthe, B., editors, *Language and Automata Theory and Applications, 11th International Conference*, Lecture Notes in Computer Science. Springer.

Jurgec, P. (2011). *Feature spreading 2.0: a unified theory of assimilation*. PhD thesis, University of Tromsø.

References IV

Lee, S. and Pater, J. (2008). Phonological inference and word recognition: Evidence from korean. Ms., Korea University and University of Massachusetts, Amherst.

McMullin, K. (2016). *Tier-based Locality in Long-distance Phonotactics: Learnability and Typology*. PhD thesis, University of British Columbia.

Meinof, C. (1932). *Introduction to the phonology of the Bantu languages*. Berlin: Dietrich Reimer/Ernst Vohsen. Trans. by N. J. van Warmelo.

Mielke, J. (2008). *The Emergence of Distinctive Features*. Oxford University Press, Oxford.

Odden, D. (1994). Adjacency parameters in phonology. *Language*, 70(2):289–330.

Onn, F. M. (1980). *Aspects of Malay Phonology and Morphology: A Generative Approach*. Kuala Lumpur: Universiti Kebangsaan Malaysia.

References V

Rose, S. and Walker, R. (2004). A typology of consonant agreement as correspondence. *Language*, 80:475–531.