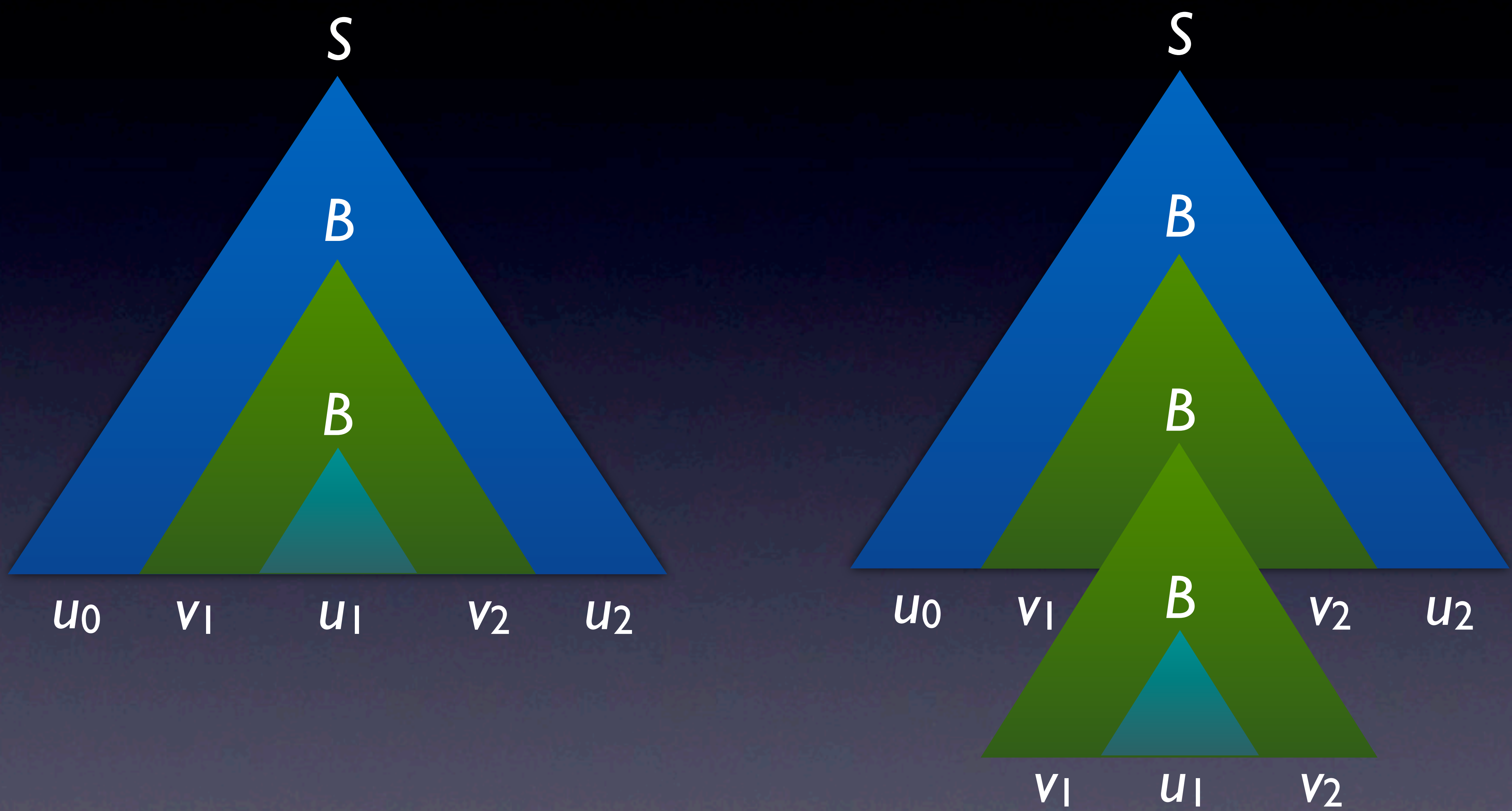
Pumping

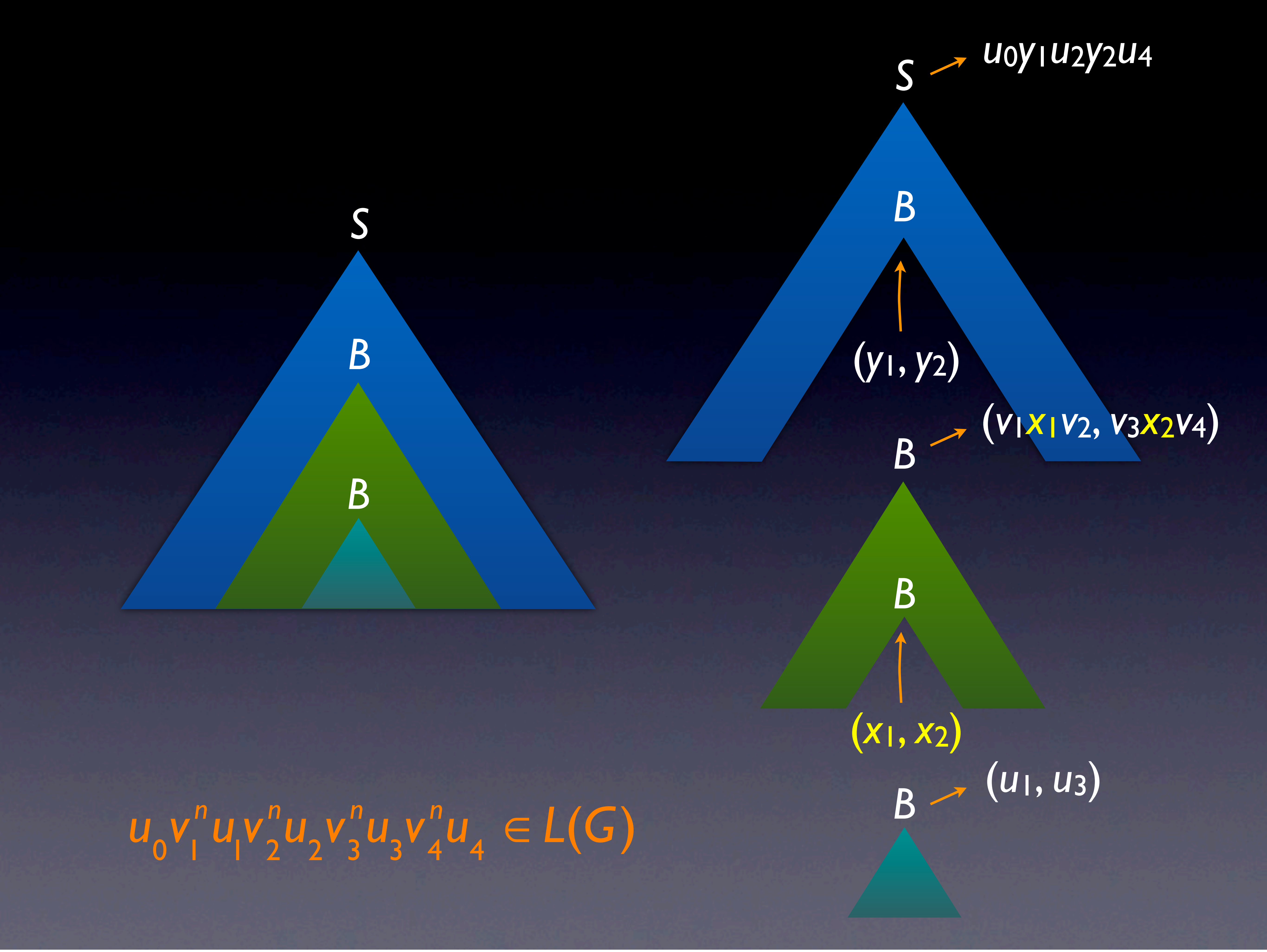
Makoto Kanazawa
National Institute of Informatics

Pumping



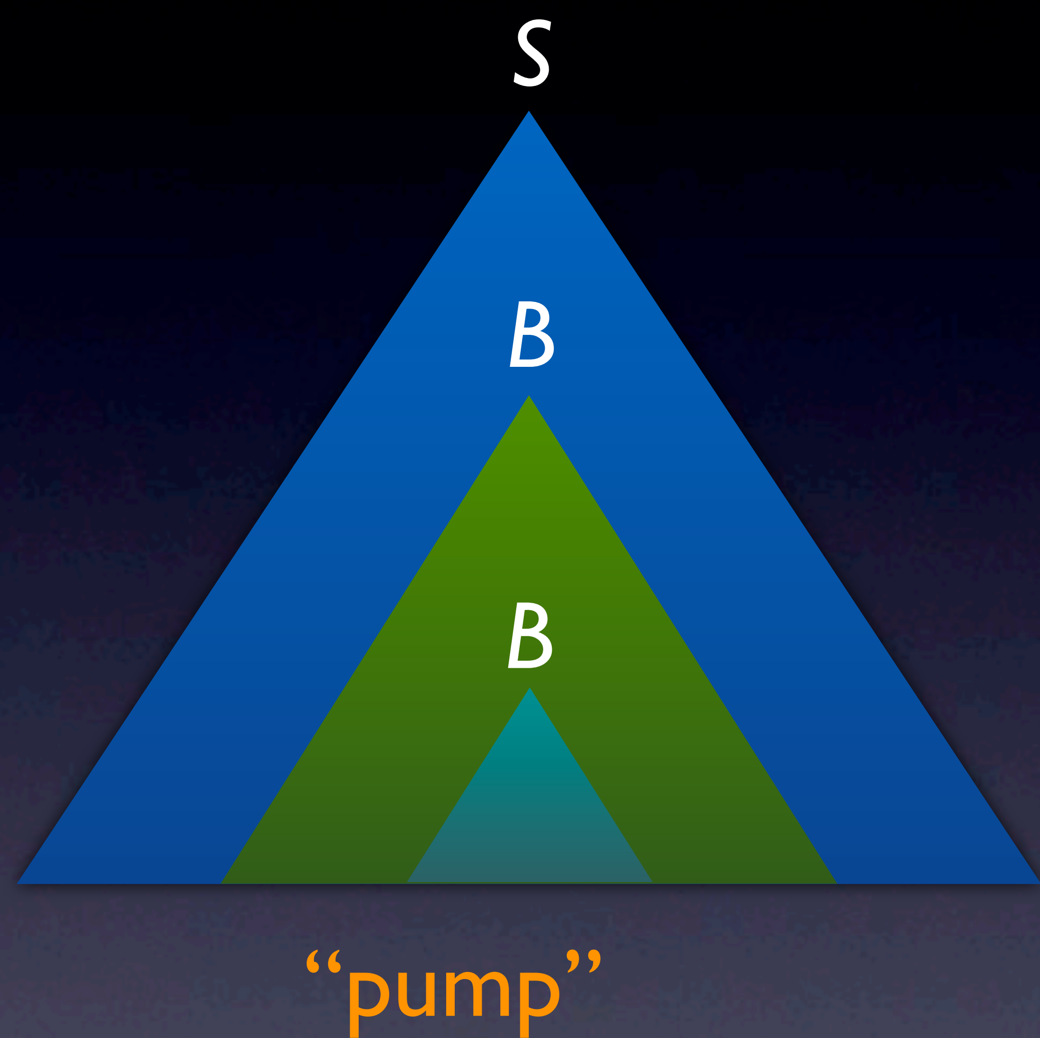
$u_0 v_1^n u_1 v_2^n u_2 \in L(G)$ for all $n \geq 0$

The case of CFG.



The case of 2-MCFGs.
Is this the general picture?

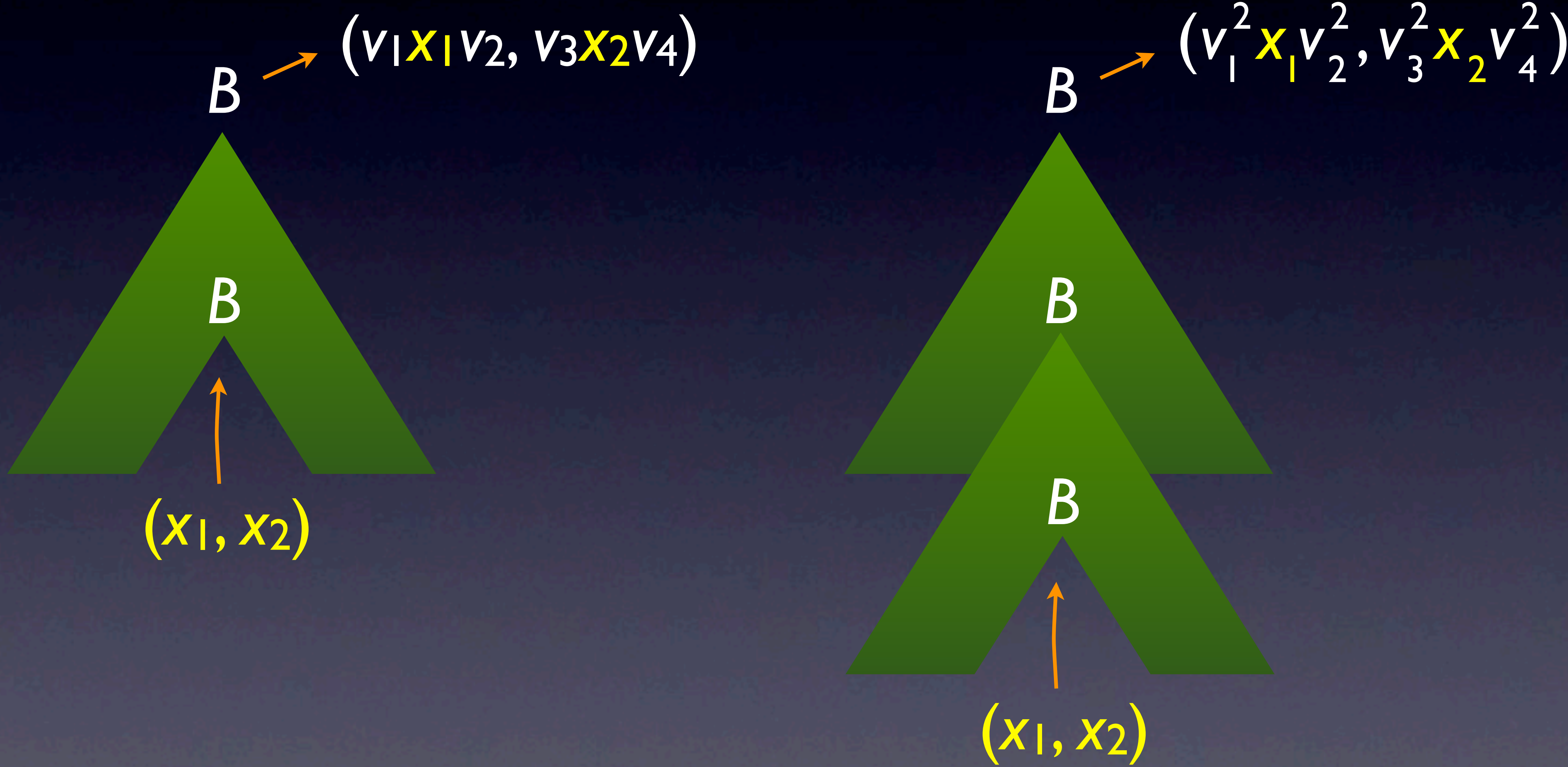
Difficulty with Pumping



All but finitely many derivation trees contain a pump.

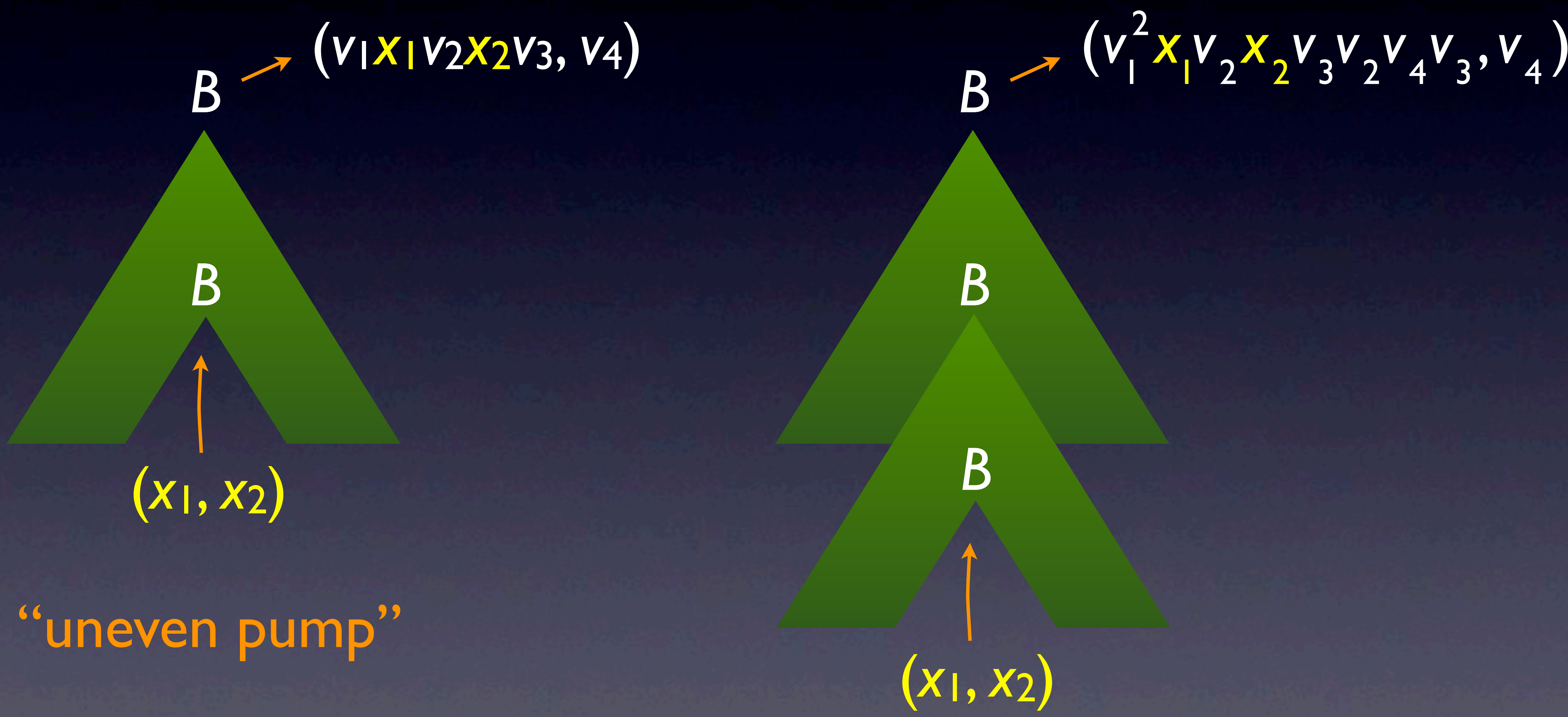
All sufficiently large derivation trees contain a part that can be iterated.

Difficulty with Pumping

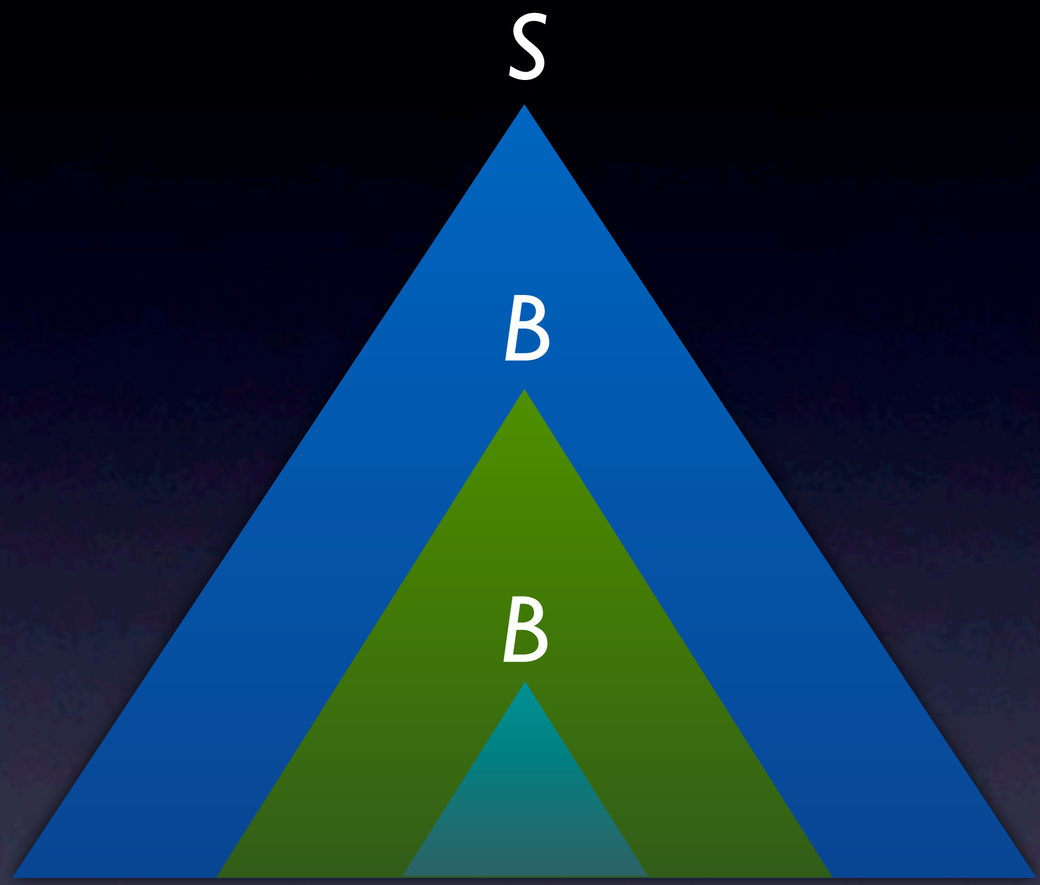


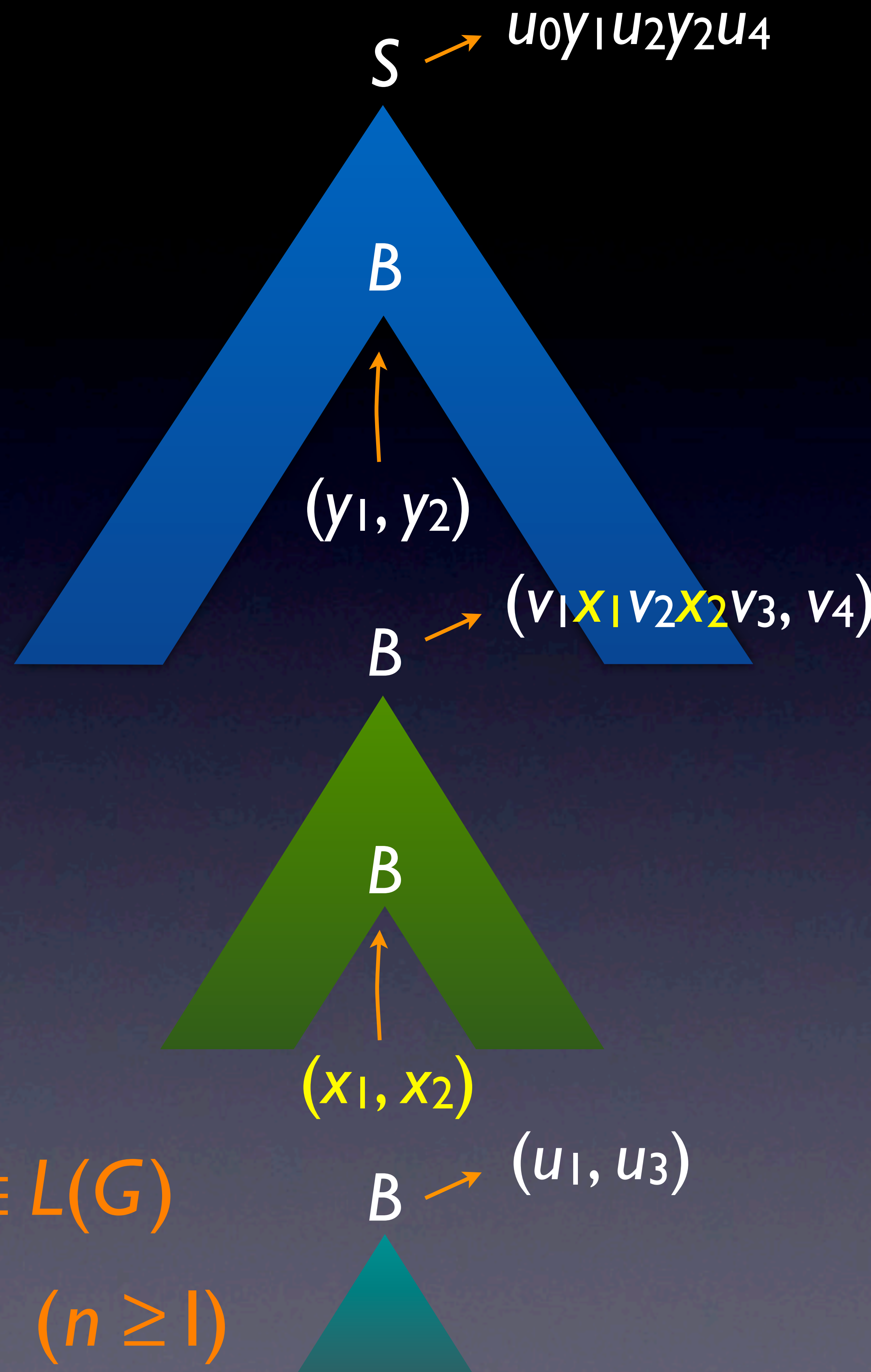
A derivation tree containing this pump yields a 4-pumpable string.

Difficulty with Pumping



Rather complex pattern.





$$u_1 u_2 u_2 u_3 u_4 \in L(G)$$

$$u_0 v_1^n u_1 v_2 u_3 v_3 (v_2 v_4 v_3)^{n-1} u_2 v_4 u_4 \in L(G)$$

$$(n \geq 1)$$

The original string (n=1) cannot be pumped, but the string obtained by iterating the pump twice is 2-pumpable.
Cf. Vijay-Shanker 1987.

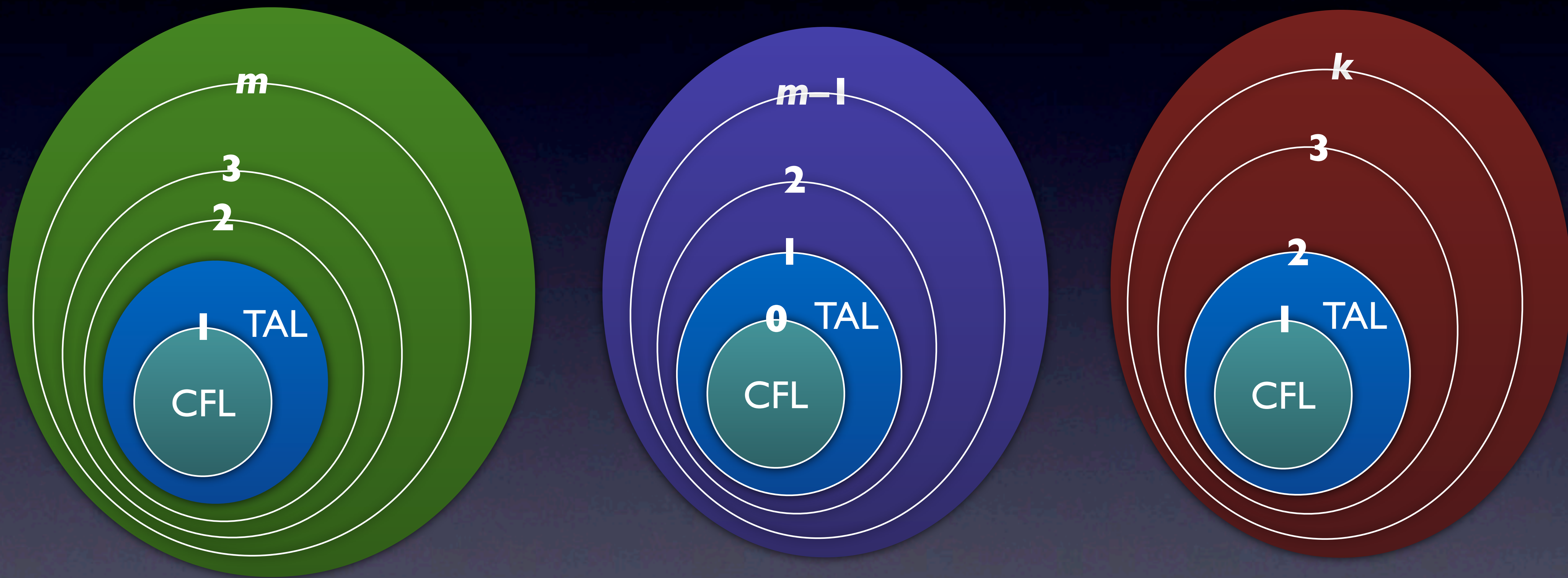
Theorem (Seki et al. 1991)**.**
$L \in m$-MCFL $\Rightarrow L$ is weakly $2m$-iterative.

Myth. $L \in m$-MCFL $\Rightarrow L$ is $2m$-iterative.

Radzinski 1991, Groenink 1997, Kracht 2003

L is weakly k-iterative if it contains an infite subset that is k-iterative.
Many people erroneously believed that Seki et al. proved a stronger result.

Three Infinite Hierarchies

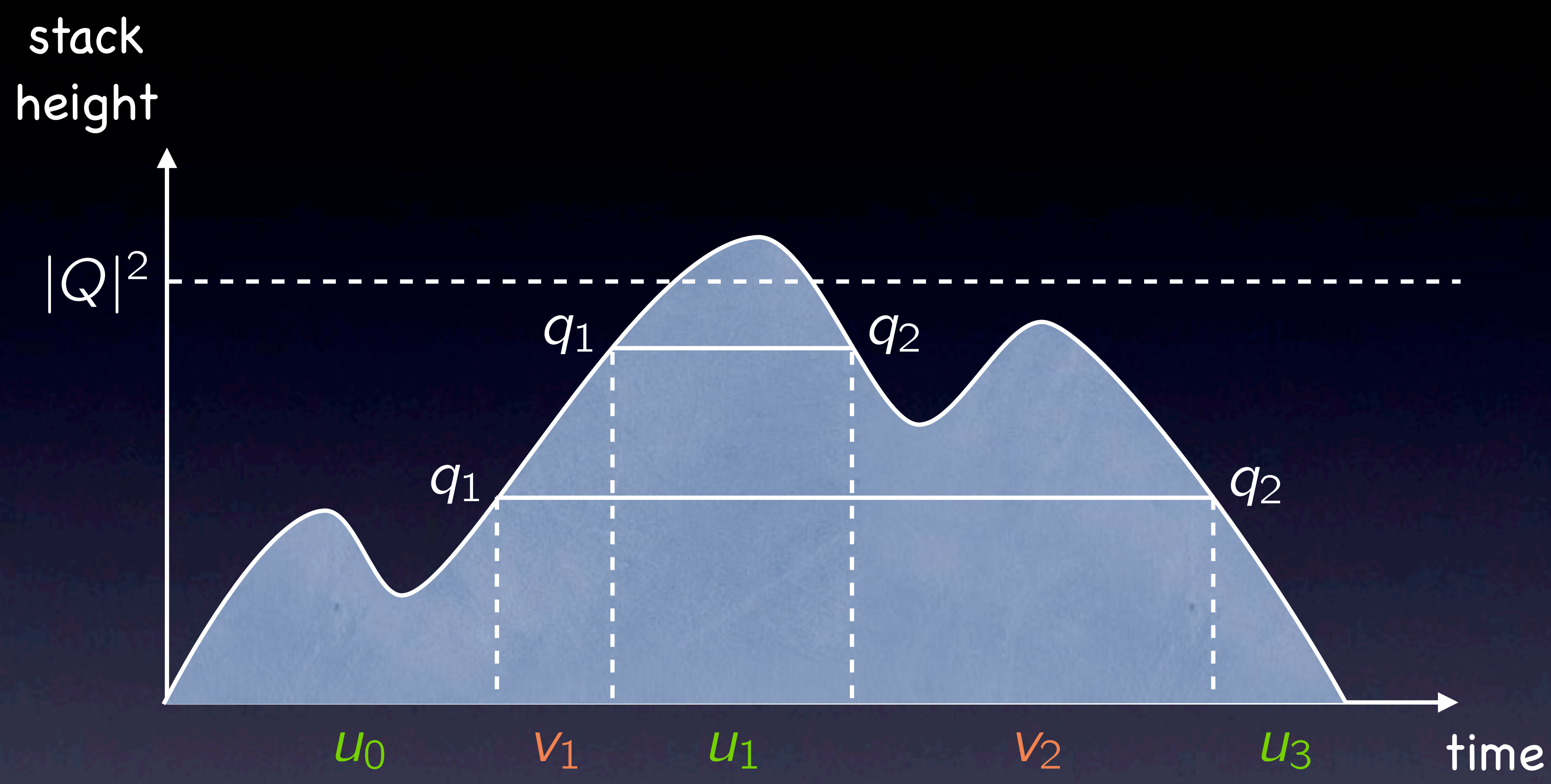


$$\mathrm{MCFL} = \bigcup_{m\geq 1} m\text{-MCFL} \qquad y\mathrm{CFT}_{\mathrm{sp}} = \bigcup_{m\geq 1} y\mathrm{CFT}_{\mathrm{sp}}(m-1) \qquad \mathbf{C} = \bigcup_{k\geq 1} \mathbf{C}_k$$

$$\mathrm{MCFL}_{\mathrm{wn}} = \bigcup_{m\geq 1} m\text{-MCFL}_{\mathrm{wn}}$$

Pumping lemmas in the usual form hold for the two subhierarchies of the MCFLs.

Pumping Lemma for PDA

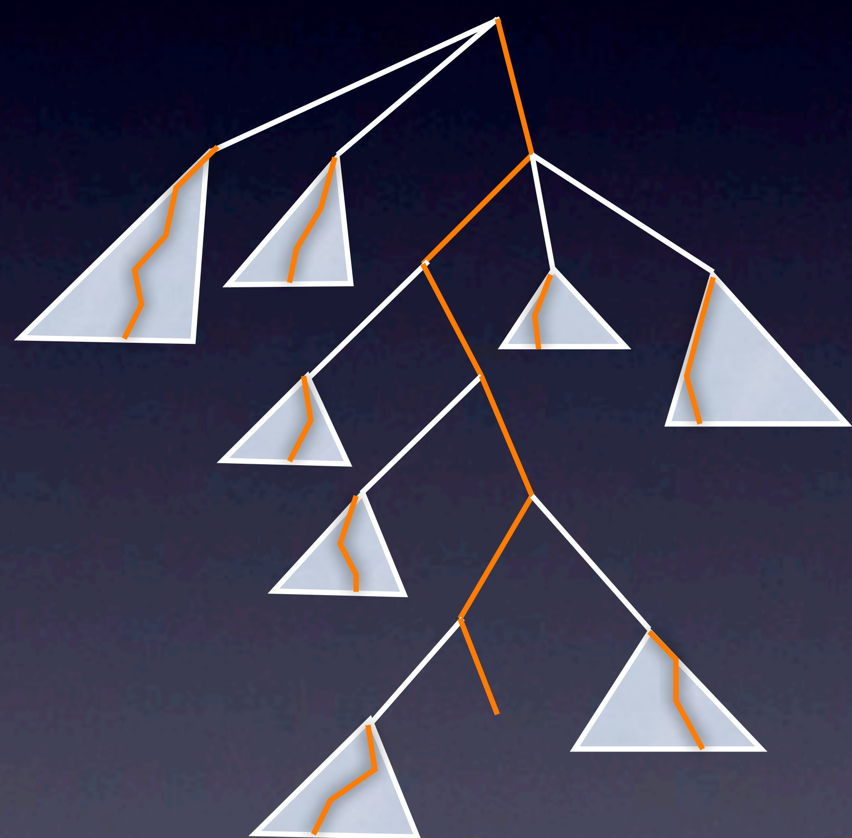


- ¬(All but finitely many accepting computations reach stack height $|Q|^2$)
- { w | w has an accepting computation that doesn't reach stack height $|Q|^2$ } is regular

The proof in each case is somewhat similar to the proof of the pumping lemma for CFLs using PDA, rather than CFG.

Pumping Lemma for **C**$_k$

$$\mathrm{CT}(L_1, L_2) = L_1 \cap (\widehat{L_2})^{*,c}$$

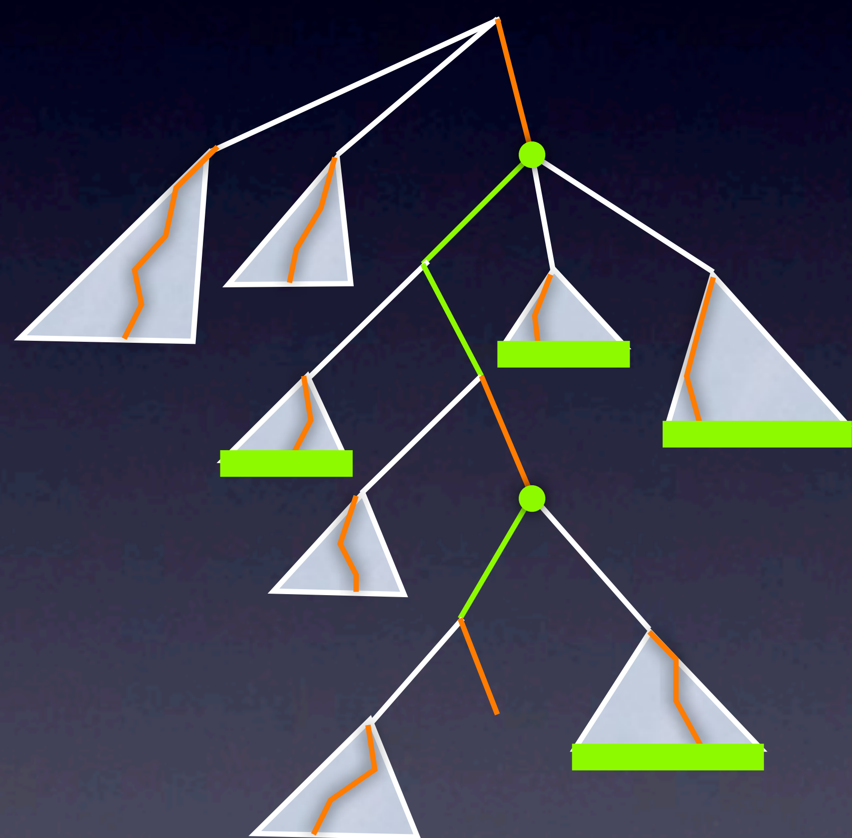


- No long spine ⇒ element of a regular set

The set of trees without long spines are the Kleene star of a finite set.

Pumping Lemma for $\mathbf{C}_k$

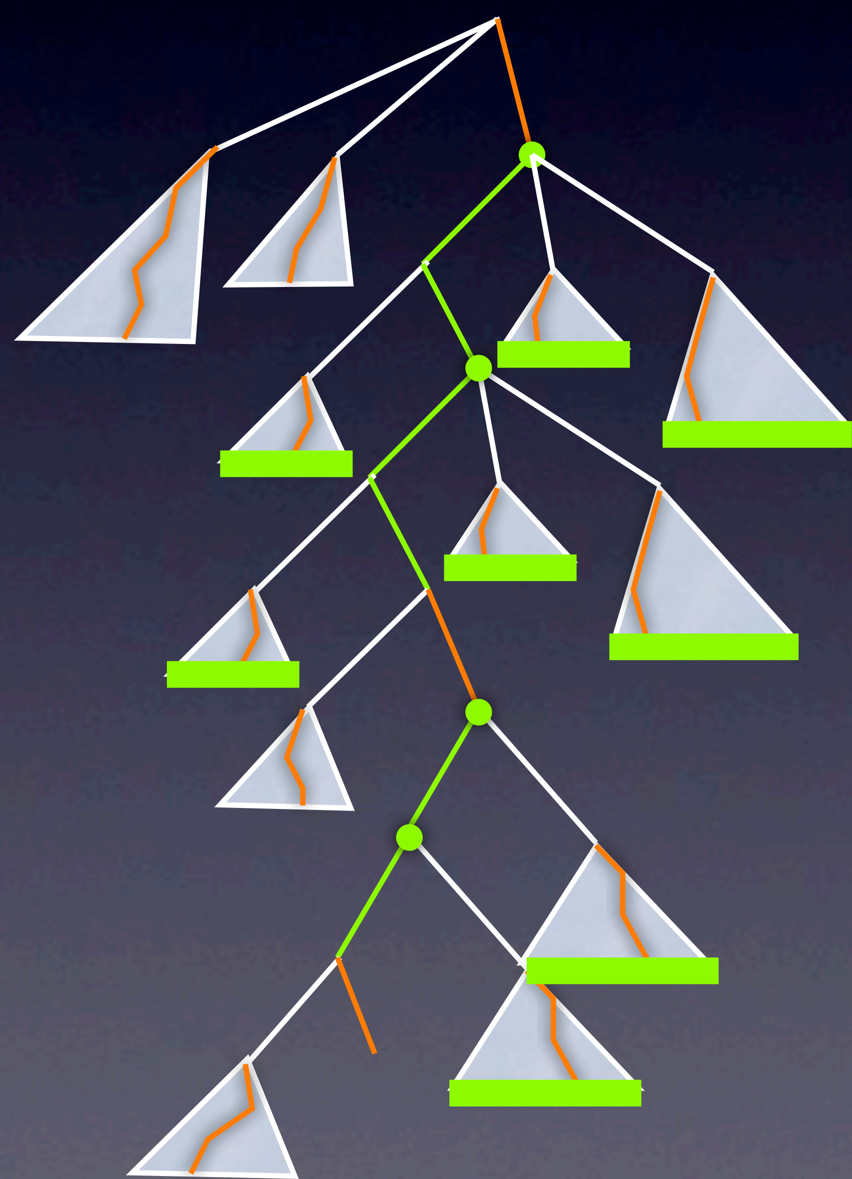
$$\mathrm{CT}(L_1, L_2) = L_1 \cap (\widehat{L_2})^{*,c}$$



When some spine is long enough to be pumpable, …

Pumping Lemma for $\mathbf{C}_k$

$$\mathrm{CT}(L_1, L_2) = L_1 \cap (\widehat{L_2})^{*,c}$$



52

When a tree has a spine that is m-pumpable, the yield of the tree is 2m-pumpable.

Pumping Lemma for $\mathbf{C}_k$

$$\mathrm{CT}(L_1, L_2) = L_1 \cap (\widehat{L_2})^{*,c}$$

$L_1 \in \mathrm{LOC}$ and L_2 is m-iterative
$\Rightarrow y\mathrm{CT}(L_1, L_2)$ is $2m$-iterative

Theorem (Palis and Shende 1995).
$L \in \mathbf{C}_k \Rightarrow L$ is 2^k-iterative.

Pumping Lemma for m-$\mathrm{MCFL}_{\mathrm{wn}}$

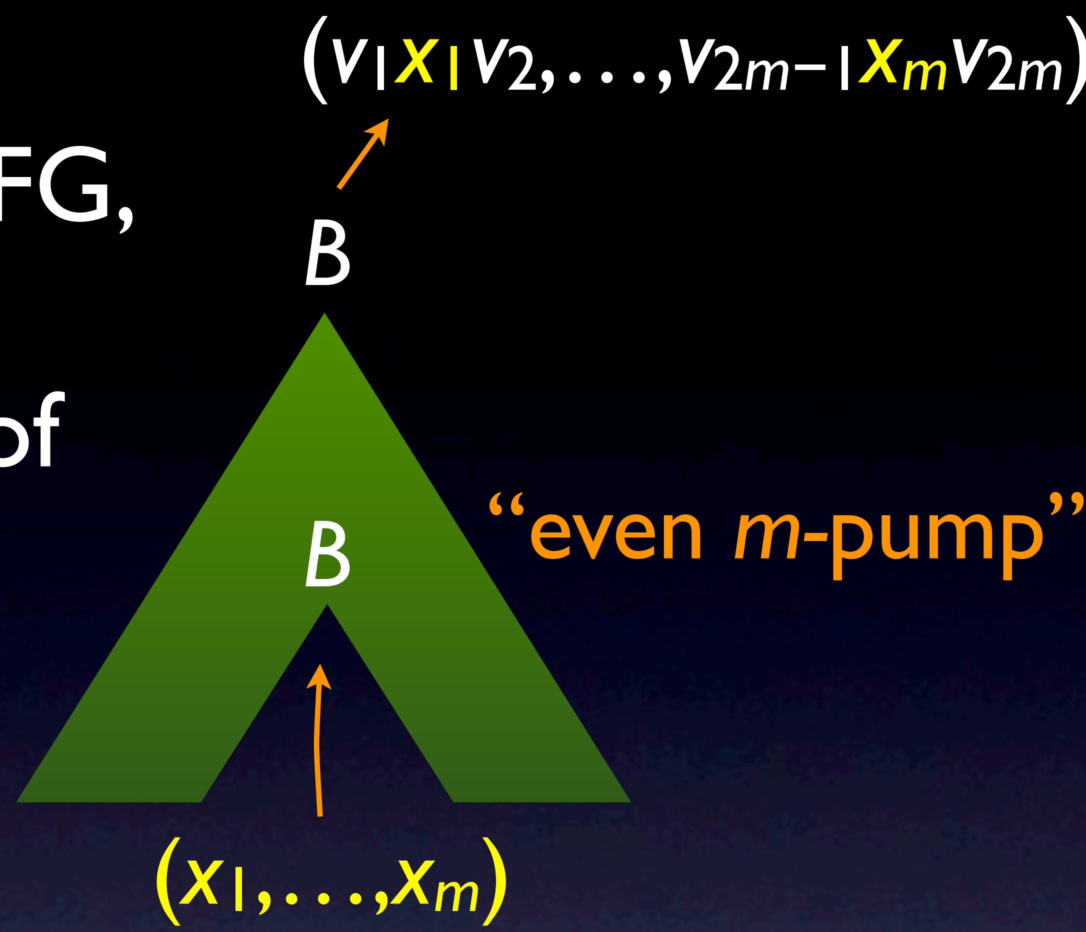
Theorem (Kanazawa 2009)**.**
$L \in m\text{-}\mathrm{MCFL}_{\mathrm{wn}} \Rightarrow L$ is $2m$-iterative.

The proof of the Pumping Lemma for m-$\mathrm{MCFL}_{\mathrm{wn}}$ is more complex.

- If G is a well-nested m-MCFG,

  { T | T is a derivation tree of G without even m-pumps }

  may not be finite.

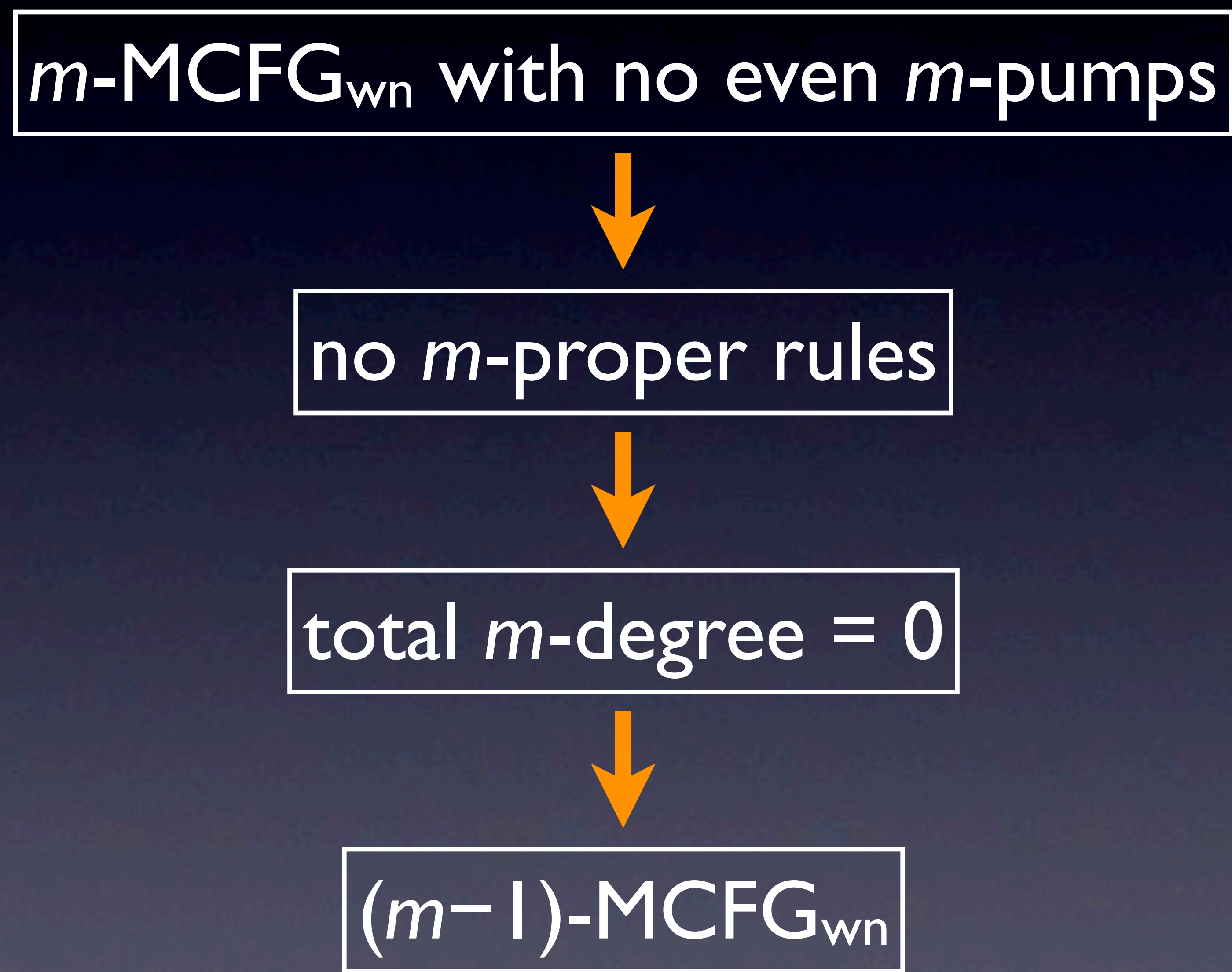


- But there is a well-nested $(m-1)$-MCFG generating

  { yield(T) | T is a derivation tree of G without even m-pumps }.

If the derivation tree contains an even m-pump, the string is 2m-pumpable.
Otherwise, the string is in the language of some w.n. (m-1)-MCFG, and therefore is 2(m-1)-pumpable (disregarding finitely many exceptions).
Proof by induction on m.

Program Transformation



The proof of this claim is by successive transformations on the grammar.

π_1: $S(\mathbf{x}_1\mathbf{x}_2) \leftarrow B(\mathbf{x}_1, \mathbf{x}_2)$

π_2: $B(a\mathbf{x}_1 b, c\mathbf{x}_2 d) \leftarrow A(\mathbf{x}_1, \mathbf{x}_2)$ *m*-proper rule

π_3: $A(a\mathbf{x}_1 b\mathbf{x}_2 c, d) \leftarrow A(\mathbf{x}_1, \mathbf{x}_2)$

π_4: $A(\varepsilon, \varepsilon) \leftarrow$

↓ unfolding

π_1: $S(\mathbf{x}_1\mathbf{x}_2) \leftarrow B(\mathbf{x}_1, \mathbf{x}_2)$

$\pi_2 \circ \pi_3$: $B(aa\mathbf{x}_1 b\mathbf{x}_2 cb, cdd) \leftarrow A(\mathbf{x}_1, \mathbf{x}_2)$

$\pi_2 \circ \pi_4$: $B(ab, cd) \leftarrow$

π_3: $A(a\mathbf{x}_1 b\mathbf{x}_2 c, d) \leftarrow A(\mathbf{x}_1, \mathbf{x}_2)$

π_4: $A(\varepsilon, \varepsilon) \leftarrow$

A rule is m-proper if the head nonterminal is m-ary and there is an m-ary nonterminal on the right-hand side, each of whose arguments appear in the corresponding argument of the head nonterminal.
Unfold until there is no m-proper rule. This procedure terminates because the grammar does not allow an even m-pump.

π_1: $S(\mathbf{x}_1\mathbf{x}_2) \leftarrow B(\mathbf{x}_1, \mathbf{x}_2)$

π_5: $B(aa\mathbf{x}_1 b\mathbf{x}_2 cb, cdd) \leftarrow A(\mathbf{x}_1, \mathbf{x}_2)$ — m-degree = 1

π_6: $B(ab, cd) \leftarrow$

π_3: $A(a\mathbf{x}_1 b\mathbf{x}_2 c, d) \leftarrow A(\mathbf{x}_1, \mathbf{x}_2)$ — m-degree = 1

π_4: $A(\varepsilon, \varepsilon) \leftarrow$

↓ unfolding^{-1}

π_1: $S(\mathbf{x}_1\mathbf{x}_2) \leftarrow B(\mathbf{x}_1, \mathbf{x}_2)$

$\pi_{5.1}$: $B(aa\mathbf{x}cb, cdd) \leftarrow C(\mathbf{x})$

$\pi_{5.2}$: $C(\mathbf{x}_1 b\mathbf{x}_2) \leftarrow A(\mathbf{x}_1, \mathbf{x}_2)$

$\pi_5 = \pi_{5.1} \circ \pi_{5.2}$

π_2: $B(ab, cd) \leftarrow$

$\pi_{3.1}$: $A(a\mathbf{x}c, d) \leftarrow D(\mathbf{x})$

$\pi_{3.2}$: $D(\mathbf{x}_1 b\mathbf{x}_2) \leftarrow A(\mathbf{x}_1, \mathbf{x}_2)$

$\pi_3 = \pi_{3.1} \circ \pi_{3.2}$

π_4: $A(\varepsilon, \varepsilon) \leftarrow$

The m-degree of a rule is 0 if the arity of the head nonterminal is < m; otherwise it's the number of m-ary nonterminals on the right-hand side.
Do the converse of unfolding.

π_1: $S(\mathbf{x}_1\mathbf{x}_2) \leftarrow B(\mathbf{x}_1, \mathbf{x}_2)$

$\pi_{5.1}$: $B(aa\mathbf{x}cb, cdd) \leftarrow C(\mathbf{x})$

$\pi_{5.2}$: $C(\mathbf{x}_1 b \mathbf{x}_2) \leftarrow A(\mathbf{x}_1, \mathbf{x}_2)$

π_2: $B(ab, cd) \leftarrow$

$\pi_{3.1}$: $A(a\mathbf{x}c, d) \leftarrow D(\mathbf{x})$

$\pi_{3.2}$: $D(\mathbf{x}_1 b \mathbf{x}_2) \leftarrow A(\mathbf{x}_1, \mathbf{x}_2)$

π_4: $A(\varepsilon, \varepsilon) \leftarrow$

$\downarrow$ unfolding

$\pi_1 \circ \pi_{5.1}$: $S(aa\mathbf{x}cbcdd) \leftarrow C(\mathbf{x})$

$\pi_1 \circ \pi_2$: $S(abcd) \leftarrow$

$\pi_{5.2} \circ \pi_{3.1}$: $C(a\mathbf{x}cbd) \leftarrow D(\mathbf{x})$

$\pi_{5.2} \circ \pi_4$: $C(b) \leftarrow$

$\pi_{3.2} \circ \pi_{3.1}$: $D(a\mathbf{x}cbd) \leftarrow D(\mathbf{x})$

$\pi_{3.2} \circ \pi_4$: $D(b) \leftarrow$

Now each rule contains m-ary nonterminals only on one side of the rule, if any. Unfolding eliminates all m-ary nonterminals.

Program Transformation

m-MCFG$_{wn}$ with no even m-pumps

↓ unfolding

no m-proper rules

↓ unfolding^{-1}

total m-degree = 0

↓ unfolding

$(m-1)$-MCFG$_{wn}$

Reduction of *m*-degrees

$$B(\mathbf{x}_1\mathbf{y}_1\mathbf{z}_1, \mathbf{z}_2\mathbf{y}_2 a\mathbf{y}_3 b, c\mathbf{x}_2 d) \leftarrow A(\mathbf{x}_1, \mathbf{x}_2), B(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3), C(\mathbf{z}_1, \mathbf{z}_2)$$

↓ unfolding^{-1}

$$B(\mathbf{x}_1\mathbf{w}_1, \mathbf{w}_2 b, c\mathbf{x}_2 d) \leftarrow A(\mathbf{x}_1, \mathbf{x}_2), D(\mathbf{w}_1, \mathbf{w}_2)$$

$$D(\mathbf{y}_1\mathbf{z}_1, \mathbf{z}_2\mathbf{y}_2 a\mathbf{y}_3) \leftarrow B(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3), C(\mathbf{z}_1, \mathbf{z}_2)$$

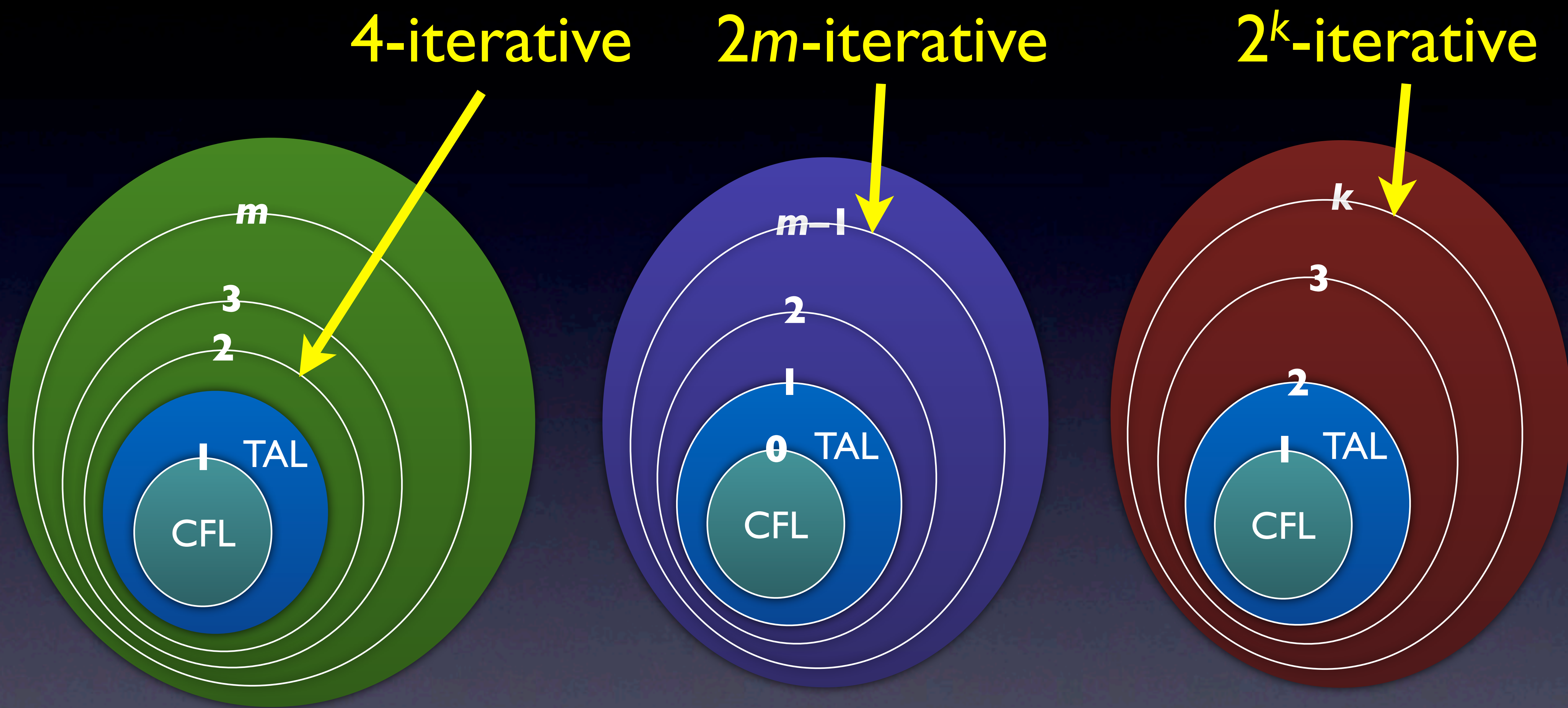
The well-nestedness assumption is necessary in the second step.
Here’s a case of a well-nested rule.

Reduction of *m*-degrees

$$B(\mathbf{x}_1\mathbf{y}_1\mathbf{x}_2,\ \mathbf{z}_1\mathbf{y}_2 a\mathbf{y}_3 b,\ c\mathbf{z}_2 d) \leftarrow A(\mathbf{x}_1, \mathbf{x}_2),\ B(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3),\ C(\mathbf{z}_1, \mathbf{z}_2)$$

$\downarrow$ unfolding^{-1}

If a rule is non-well-nested, the procedure does not work.



$$\mathrm{MCFL} = \bigcup_{m \geq 1} m\text{-}\mathrm{MCFL} \qquad y\mathrm{CFT}_{sp} = \bigcup_{m \geq 1} y\mathrm{CFT}_{sp}(m-1) \qquad \mathbf{C} = \bigcup_{k \geq 1} \mathbf{C}_k$$

$$\mathrm{MCFL}_{wn} = \bigcup_{m \geq 1} m\text{-}\mathrm{MCFL}_{wn}$$

The proof shows that a 2-MCFL is 4-iterative.

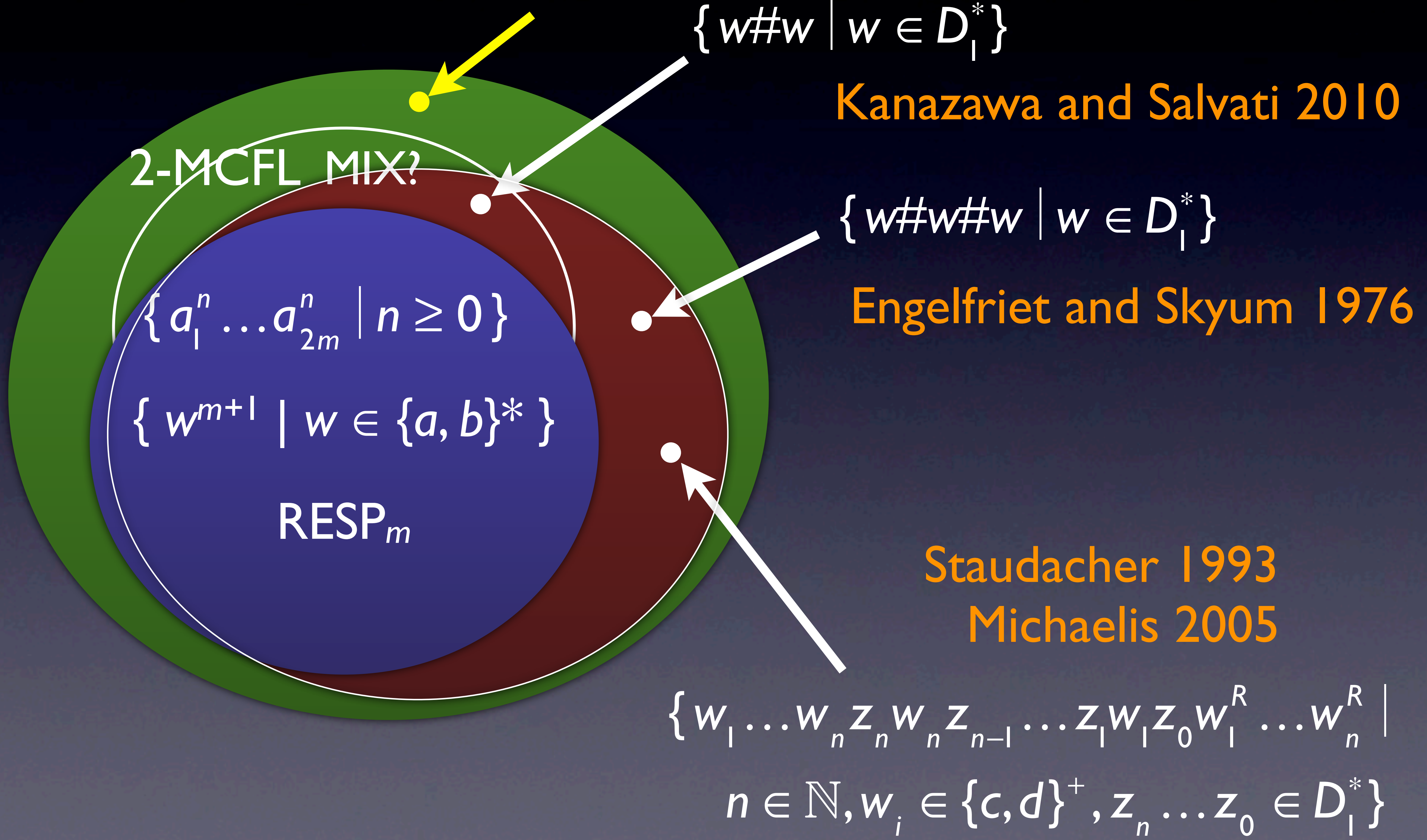
$$H(\mathbf{x}_2) \leftarrow G(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$$
$$G(a\mathbf{x}_1, \mathbf{y}_1 c \mathbf{x}_2 \overline{c} d \mathbf{y}_2 \overline{d} \mathbf{x}_3, \mathbf{y}_3 b) \leftarrow G(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3), G(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3)$$
$$G(a, \varepsilon, b) \leftarrow$$

$$a^{n+1} c \ldots b^m \; \overline{c} \; d \; a^n \ldots \overline{d} \; b^{m+1}$$

$$v_0 = \varepsilon$$
$$v_{n+1} = a^{n+1} c \, v_n \, \overline{c} \, d \, v_n \, \overline{d} \, b^{n+1}$$

Pumping fails for m-MCFLs for (m > 2).
Here's an example of a 3-MCFL that is not k-iterative for any k.

MCFL vs. MCFL_{wn} vs. **C**



Since every language in C is k-iterative for some k, this language separates MCFL from C.