## Multisets

multiset M over V — function from  $V^*$  into  ${\bf N}$ 

M(x) — multiplicity of  $x \in V^*$ 

A multiset M is called finite if there is a finite subset U of  $V^*$  such that M(x) = 0 for all  $x \notin U$ .

a finite multiset can be represented as  $[w_1, w_2, \dots, w_n]$ ,  $l(M) = |w_1 w_2 \dots w_n|$ ,  $\#_a(M) = \#_a(w_1 w_2 \dots w_n)$ 

# Multiset Splicing I

**Definition**: For multisets  $M = [w_1, w_2, \dots, w_n]$  and  $M' = [v_1, v_2, \dots, v_n]$  of words over V and a set P of splicing rules over V, we define

- a sequential derivation step  $M \Longrightarrow M'$  by  $[w_1, w_2] \Longrightarrow [v_1, v_2]$  and  $w_i = v_i$  for  $3 \le j \le n$  for some  $p \in P$  and some appropriate order of the elements in M and M',
- a maximally parallel derivation step  $M \Longrightarrow_{mp} M'$  by  $[w_{2i-1}, w_{2i}] \Longrightarrow_{p_i} [v_{2i-1}, v_{2i}]$  for  $1 \le i \le k \le \frac{n}{2}$  and  $w_i = v_i$  for  $2k + 1 \le j \le n$  for some  $p_i \in P$  and some appropriate order of the elements in M and M', and by the requirement that there is no multiset  $[w, w'] \subseteq [w_{2k+1}, w_{2k+2}, \dots, w_n]$  to which a splicing rule  $p \in P$  can be (successfully) applied,
- a strongly maximally parallel derivation step  $M \Longrightarrow_{smp} M'$ by  $M \Longrightarrow_{mp} M'$  for some k (as in the preceding item) and there is no M''with  $M \Longrightarrow_{mp} M''$  for some k' > k.

# Multiset Splicing Systems

**Definition**: A multiset splicing system is a triple G = (V, P, M) where

- -V is an alphabet,
- P is a finite set of splicing rules over V such that, for any rule  $r_1 \# r_2 \$ r_3 \# r_4 \in P$ ,  $r_i \neq \lambda$  for  $1 \leq i \leq 4$ , and
- M is a finite multiset over V.

**Definition** We define the sequential, maximally parallel and strongly maximally parallel multiset languages mL(G,s), mL(G,mp) and mL(G,smp) generated by G as

$$mL(G, s) = \{K \mid M \Longrightarrow_{s}^{*} K\},\$$

$$mL(G, mp) = \{K \mid M \Longrightarrow_{mp}^{*} K\},\$$

$$mL(G, smp) = \{K \mid M \Longrightarrow_{smp}^{*} K\}.$$

### **Some Notations**

for  $Y \in \{s, mp, smp\}$ , we denote by

- $m\mathcal{L}(Y)$  the family of all languages mL(G, Y) which can be generated by a multiset splicing system G in the derivation mode Y,
- $m\mathcal{L}_n(Y)$  the family of all languages mL(G, Y) which can be generated by a multiset splicing system G = (V, P, M)with #(M) = n in the derivation mode Y.

# Some Facts I

#### Lemma:

For any multiset splicing system  $G=(V,P,M),~a\in V$  ,  $Y\in\{s,mp,smp\}$  , and any  $K\in mL(G,Y)$  ,

$$\#(K) = \#(M), \ l(K) = l(M), \text{ and } \#_a(K) = \#_a(M).$$

#### Theorem:

For two integers n and m,  $m \neq n$ , and two derivation modes  $Y \in \{s, mp, smp\}$  and  $Y' \in \{s, mp, smp\}$ , the language families  $m\mathcal{L}_n(Y)$  and  $m\mathcal{L}_m(Y')$  are incomparable.

# Some Facts II

#### Theorem:

- i) For  $n \in \{1, 2, 3\}$ ,  $m\mathcal{L}_n(s) = m\mathcal{L}_n(mp) = m\mathcal{L}_n(smp)$ .
- ii) For  $n \ge 4$ ,  $m\mathcal{L}_n(mp)$  and  $m\mathcal{L}_n(smp)$  are both incomparable to  $m\mathcal{L}_n(s)$ .
- iii) For  $n \geq 5$ ,  $m\mathcal{L}_n(mp)$  is not contained in  $m\mathcal{L}_n(smp)$ .
- iv) For  $n \ge 6$ , the classes  $m\mathcal{L}_n(s)$ ,  $m\mathcal{L}_n(mp)$ , and  $m\mathcal{L}_n(smp)$  are pairwise incomparable.

# **Sticking Operation – Prolongation to the Rigth**

For 
$$x \in W_{\varrho}(V)$$
 mit  $x = x_1 x_2 x_3$ ,  $y \in W_{\varrho}(V)$ , we define  $\mu_r(x, y)$  by  
1.  $x_1 x_2 \begin{bmatrix} u \\ v \end{bmatrix} y'$ , if  $x_3 = \begin{pmatrix} u \\ \lambda \end{pmatrix}$ ,  $y = \begin{pmatrix} \lambda \\ v \end{pmatrix} y'$   $(u, v \in V^*, y' \in R_{\varrho}(V))$ ,  
2.  $x_1 x_2 \begin{bmatrix} u \\ v \end{bmatrix} y'$ , if  $x_3 = \begin{pmatrix} \lambda \\ v \end{pmatrix}$ ,  $y = \begin{pmatrix} u \\ \lambda \end{pmatrix} y'$   $(u, v \in V^*, y' \in R_{\varrho}(V))$ ,  
3.  $x_1 x_2 \begin{bmatrix} u \\ v \end{bmatrix} \begin{pmatrix} u' \\ \lambda \end{pmatrix}$ , if  $x_3 = \begin{pmatrix} uu' \\ \lambda \end{pmatrix}$ ,  $y = \begin{pmatrix} \lambda \\ vv' \end{pmatrix}$   $(u, v, u' \in V^*, y' \in R_{\varrho}(V))$ ,  
4.  $x_1 x_2 \begin{bmatrix} u \\ v \end{bmatrix} \begin{pmatrix} \lambda \\ v' \end{pmatrix}$ , if  $x_3 = \begin{pmatrix} u \\ \lambda \end{pmatrix}$ ,  $y = \begin{pmatrix} \lambda \\ vv' \end{pmatrix}$   $(u, v, v' \in V^*, y' \in R_{\varrho}(V))$ ,  
5.  $x_1 x_2 \begin{pmatrix} uv \\ \lambda \end{pmatrix}$ , if  $x_3 = \begin{pmatrix} u \\ \lambda \end{pmatrix}$ ,  $y = \begin{pmatrix} v \\ \lambda \end{pmatrix}$   $(u, v \in V^*)$ ,  
6.  $x_1 x_2 \begin{bmatrix} v \\ u \end{bmatrix} \begin{pmatrix} \lambda \\ u' \end{pmatrix}$ , if  $x_3 = \begin{pmatrix} \lambda \\ uu' \end{pmatrix}$ ,  $y = \begin{pmatrix} v \\ \lambda \end{pmatrix}$   $(u, v, u' \in V^*)$ ,  
7.  $x_1 x_2 \begin{bmatrix} v \\ u \end{bmatrix} \begin{pmatrix} v' \\ \lambda \end{pmatrix}$ , if  $x_3 = \begin{pmatrix} \lambda \\ u \end{pmatrix}$ ,  $y = \begin{pmatrix} vv' \\ \lambda \end{pmatrix}$   $(u, v, v' \in V^*)$ ,  
8.  $x_1 x_2 \begin{pmatrix} uv \\ \lambda \end{pmatrix}$ , if  $x_3 = \begin{pmatrix} \lambda \\ u \end{pmatrix}$ ,  $y = \begin{pmatrix} \lambda \\ v \end{pmatrix}$   $(u, v \in V^*)$ .

## **Sticker Systems**

**Definition**: i) A sticker system is a quadruple  $G = (V, \rho, A, D)$  where

- -V is an alphabet,
- $\varrho \subset V \times V$  is a symmetric relation on V,
- A is a finite subset of  $LR_{\varrho}(V)$ , and
- D is a finite subset of  $W_{\varrho}(V) \times W_{\varrho}(V)$ .
- ii) We say that  $y \in LR_{\varrho}(V)$  is derived by  $x \in LR_{\varrho}(V)$  in one step (written as  $x \Longrightarrow y$ ) iff  $y = \mu_l(\mu_r(x, y_2), y_1)$  for some  $(y_1, y_2) \in D$ .
- By  $\stackrel{*}{\Longrightarrow}$  we denote the reflexive and transitive closure of  $\Longrightarrow$ .

iii) The molecule language ML(G) and the word language wL(G) generated by G are defined by

$$ML(G) = \{ z \mid x \Longrightarrow s * z, \ x \in A, \ z \in \begin{bmatrix} V \\ V \end{bmatrix}_{\varrho}^+ \}$$

 $\mathsf{and}$ 

$$wL(G) = \{w \mid {w \brack v} \in ML(G) \text{ for some } v \in V^+\},$$
tively.

respectively.

Formal Languages and Biological Processes

### Special Sticker Systems I

#### **Definition**:

A sticker system  $G = (V, \varrho, A, D)$  is called

• one-sided if, for each pair  $(u, v) \in D$ ,  $u = \begin{pmatrix} \lambda \\ \lambda \end{pmatrix}$  or  $v = \begin{pmatrix} \lambda \\ \lambda \end{pmatrix}$  hold,

• regular if, for each pair  $(u, v) \in D$ ,  $u = \begin{pmatrix} \lambda \\ \lambda \end{pmatrix}$  holds,

• simple if, for each pair  $(u, v) \in D$ ,  $uv \in {\binom{V^*}{\lambda}}$  or  $uv \in {\binom{\lambda}{V^*}}$  hold.

## Special Sticker Systems II

#### **Definition**:

i) For a sticker system  $G = (V, \varrho, A, D)$  and a natural number  $d \ge 1$ , we define the language  $ML_d(G)$  as the set of all molecules which have a derivation

$$x = x_0 \Longrightarrow x_1 \Longrightarrow \ldots \Longrightarrow x_k$$
 with  $x_k \in \begin{bmatrix} V \\ V \end{bmatrix}_{\varrho}^*$  and  $d(x_i) \le d$  for  $0 \le i \le k$ .

ii) We say that a molecule language  $L \subset \begin{bmatrix} V \\ V \end{bmatrix}_{\varrho}^*$  or a word language  $L' \subset V^*$  can be generated with *bounded delay*, if there are a sticker system  $G = (V, \varrho, A, D)$  and a natural number  $d \geq 1$  such that  $ML(G) = ML_d(G)$  and L = ML(G) and L' = wL(G), respectively, are valid.

#### Sticker System – Results I

**Lemma**: For  $X \in \{A, O, R, SA, SO, SR\}$ ,  $XSL(b) \subseteq XSL$ .

**Lemma**: For  $y \in \{(b), \lambda\}$ , the following diagram holds.



#### Sticker Systems – Results II

**Lemma**:  $ASL \subseteq \mathcal{L}(CS)$ .

Lemma:  $OSL \subseteq \mathcal{L}(REG)$ .

**Lemma**: SOSL(b) = SOSL und SRSL(b) = SRSL.

**Lemma**:  $\mathcal{L}(REG) \subseteq RSL(b)$ .

**Lemma**:  $ASL(b) = \mathcal{L}(LIN)$ .

**Lemma**: There is a regular language which is not in SOSL.

