Forbidding Sets and Normal Forms for Language Forbidding-Enforcing Systems

Daniela Genova

Department of Mathematics and Statistics

University of North Florida

Jacksonville, FL 32224, USA

6th International Conference on Language and Automata Theory and Applications

LATA 2012, A Coruña, Spain, March 5-9, 2012

Main topics

- Forbidding-enforcing systems (fe-systems)
- Fe-systems defining families of languages
- Fe-systems defining a single language
- Normal forms for single language fe-systems

Motivation for fe-systems

- Introduced by A. Ehrenfeucht and G. Rozenberg
- Forbidding conditions model the fact that the presence of certain groups of molecules in a molecular system causes the system to “die".
- Enforcing conditions model the fact that presence of certain molecules in the system triggers the presence of some other molecules in the system.

Forbidding sets, f-families

A *forbidding set* $\mathcal{F}$ over an alphabet A is a possibly infinite family of finite nonempty subsets of A^+; each element $F \in \mathcal{F}$ is called a *forbidder*.

A language L is said to be *consistent with a forbidder* F, denoted by $L \,\underline{con}\, F$, if and only if $F \not\subseteq \underline{sub}(L)$.

A language L is *consistent with a forbidding set* $\mathcal{F}$ ($L \,\underline{con}\, \mathcal{F}$), if and only if $L \,\underline{con}\, F$ for all $F \in \mathcal{F}$. Denote $\mathcal{L}(\mathcal{F}) = \{L \mid L \,\underline{con}\, \mathcal{F}\}$.

Let $A = \{a, b\}$. Consider the forbidding set $\mathcal{F} = \{\{aa, bb\}, \{ab, ba\}\}$.

There are four maximal languages in $\mathcal{L}(\mathcal{F})$:

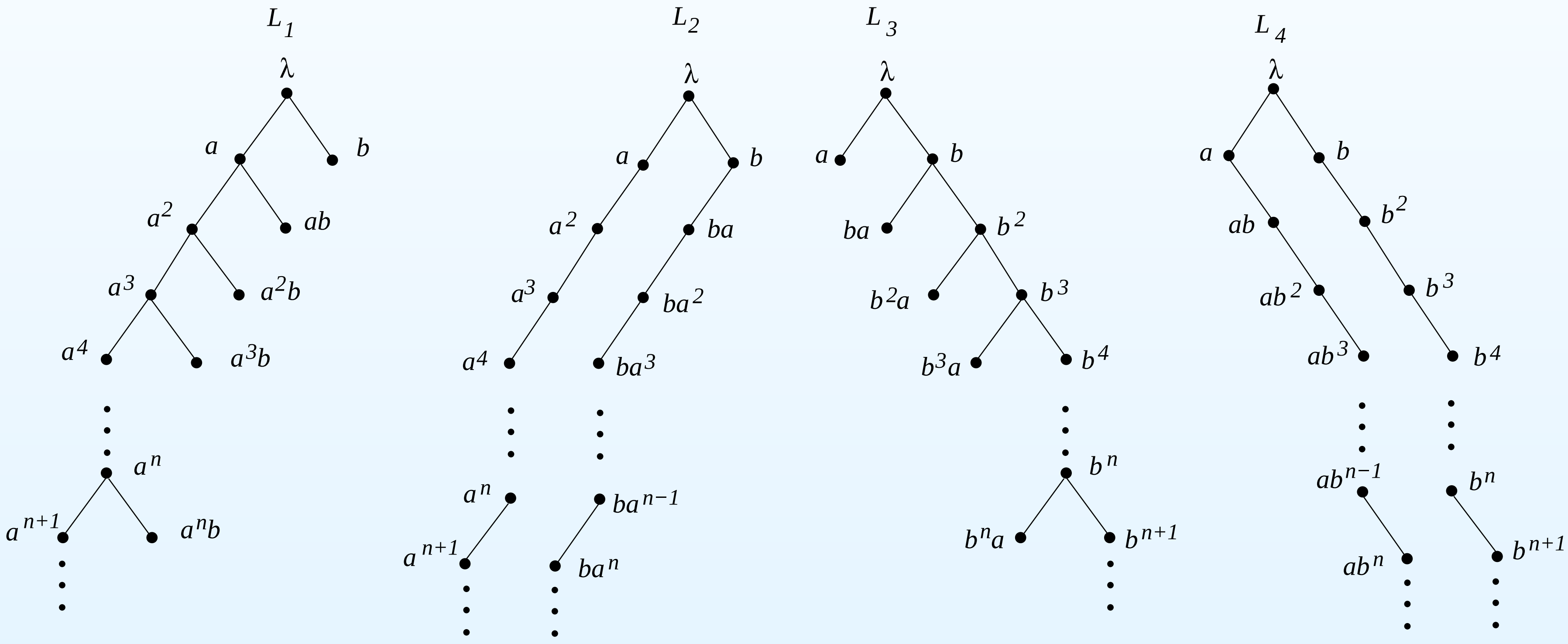
$L_1 = a^*b \cup a^*$, $L_2 = ba^* \cup a^*$, $L_3 = b^*a \cup b^*$, and $L_4 = ab^* \cup b^*$.

$\mathcal{L}(\mathcal{F}) = \{L \mid L \subseteq L_i \text{ for } i = 1, \ldots, 4\}$

Maximal languages

The maximal languages for $\mathcal{F} = \{\{aa, bb\}, \{ab, ba\}\}$.

$L_1 = a^*b \cup a^*$, $L_2 = ba^* \cup a^*$, $L_3 = b^*a \cup b^*$, and $L_4 = ab^* \cup b^*$.



Forbidding systems, f-languages

Given a forbidding set $\mathcal{F}$ over A.

A word w is *consistent with a forbidder* F, denoted by $w \,\underline{con}\, F$, if and only if $F \not\subseteq \underline{sub}(w)$.

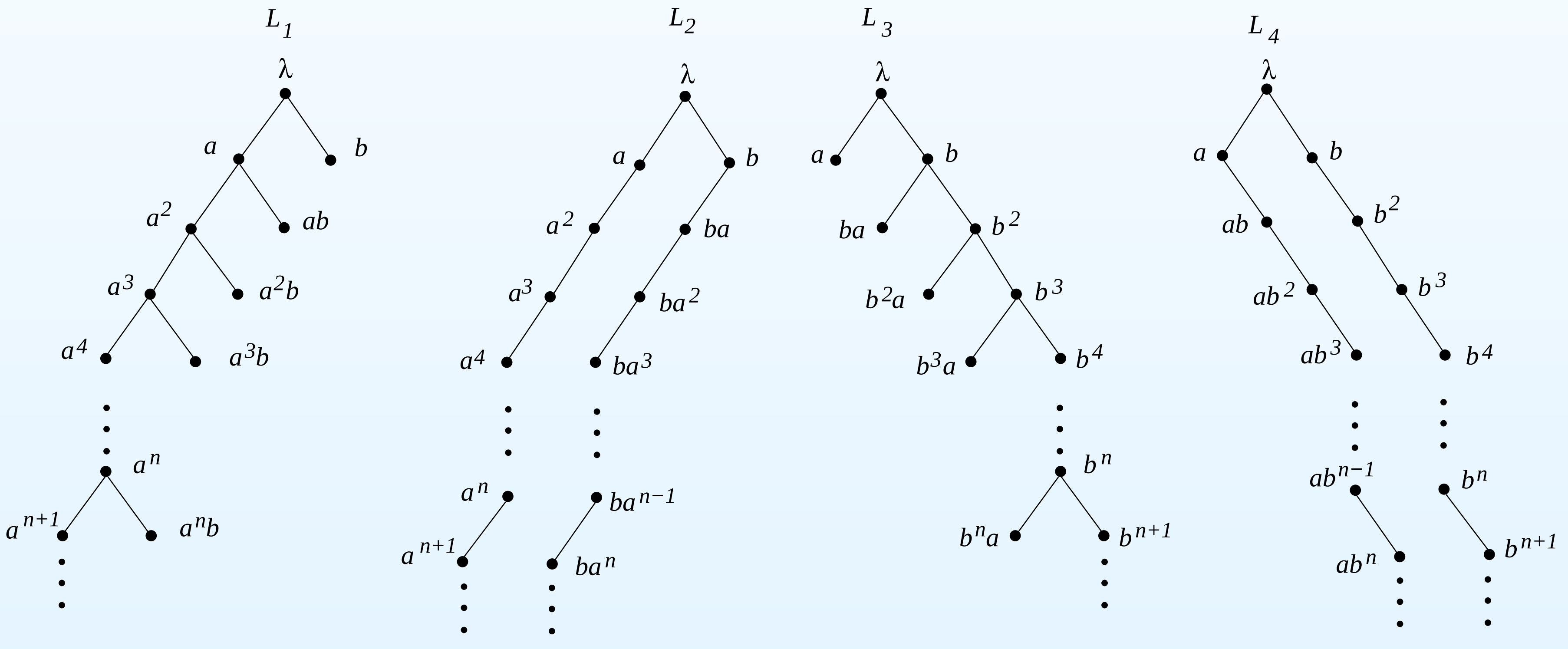
A word w is *consistent with a forbidding set* $\mathcal{F}$ denoted by $w \,\underline{con}\, \mathcal{F}$, if and only if $w \,\underline{con}\, F$ for all $F \in \mathcal{F}$. If w is not consistent with $\mathcal{F}$, the notation is $w \,\underline{ncon}\, \mathcal{F}$.

Denote $L(\mathcal{F}) = \{w \mid w \,\underline{con}\, \mathcal{F}\}$.

Let $A = \{a, b\}$ and $\mathcal{F} = \{\{aa, bb\}, \{ab, ba\}\}$. Then $L(\mathcal{F}) = \{a^n, b^n, ab^n, a^n b, ba^n, b^n a \mid n \geq 0\}$.

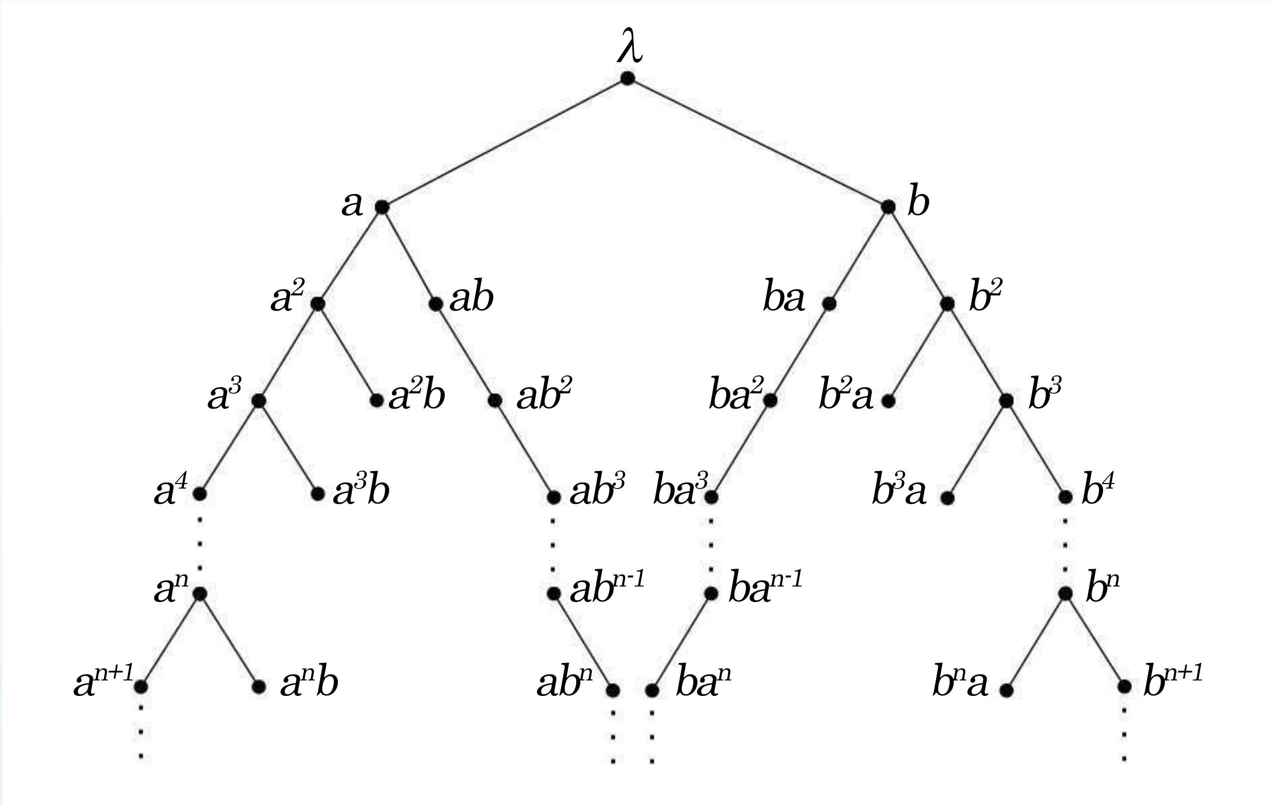
Relationship between f-families and f-languages

For the forbidding set $\mathcal{F} = \{\{aa, bb\}, \{ab, ba\}\}$ over $A = \{a, b\}$, $\mathcal{L}(\mathcal{F}) = \{L \mid L \subseteq L_i \text{ for } i = 1, \ldots, 4\}$, where $L_1 = a^*b \cup a^*$, $L_2 = ba^* \cup a^*$, $L_3 = b^*a \cup b^*$, and $L_4 = ab^* \cup b^*$.



Relationship between f-families and f-languages

Let $\mathcal{F} = \{\{aa, bb\}, \{ab, ba\}\}$. Then
$L(\mathcal{F}) = \{a^n, b^n, ab^n, a^nb, ba^n, b^na \mid n \geq 0\}$.



Relationship between f-families and f-languages

Theorem 1 *Let $\mathcal{F}$ be a forbidding set. Then,* $L(\mathcal{F}) = \cup_{L \in \mathcal{L}(\mathcal{F})} L$.

Subword free and subword incomparable, f-families

A forbidding set $\mathcal{F}$ is called *subword free* if all of its forbidders are subword free and *subword incomparable* if for any two forbidders $F_1, F_2 \in \mathcal{F}$ with $F_1 \neq F_2$, it holds that $\underline{sub}(F_1) \not\subseteq \underline{sub}(F_2)$ and $\underline{sub}(F_2) \not\subseteq \underline{sub}(F_1)$.

A. Ehrenfeucht, H. J. Hoogeboom, G. Rozenberg, N. van Vugt proved that for f-families the subword free and subword incomparable normal form is minimal and unique.

Given $\mathcal{F}$ there exists a unique $\mathcal{F}'$ in minimal normal form, such that $\mathcal{L}(\mathcal{F}) = \mathcal{L}(\mathcal{F}')$.

Subword free and subword incomparable, f-languages

Theorem 2 *Every normal form for forbidding sets for fe-families is also a normal form for forbidding sets for fe-languages.*

Corollary 3 *For every language forbidding set there is an equivalent subword free and subword incomparable forbidding set.*

Subword free and subword incomparable example

Let $A = \{a, b\}$ and $\mathcal{F} = \{\{aabb\}, \{bbaa\}, \{bbabaa\}, \{aa, bb, abab\}\}$.

For f-families, this $\mathcal{F}$ is minimal:

- Remove a forbidder, e.g. $\{aabb\}$ and consider $\mathcal{F}' = \{\{bbaa\}, \{bbabaa\}, \{aa, bb, abab\}\}$. Then, for $L = \{aabb\}$, $L \notin \mathcal{L}(\mathcal{F})$ and $L \in \mathcal{L}(\mathcal{F}')$. Hence, $\mathcal{L}(\mathcal{F}) \neq \mathcal{L}(\mathcal{F}')$.
- Remove a word from a forbidder, e.g. $abab$ and consider $\mathcal{F}'' = \{\{aabb\}, \{bbaa\}, \{bbabaa\}, \{aa, bb\}\}$. Then, for $L = \{aa, bb\}$, $L \in \mathcal{F}$ and $L \notin \mathcal{L}(\mathcal{F}'')$. Hence, $\mathcal{L}(\mathcal{F}) \neq \mathcal{L}(\mathcal{F}'')$.

Subword free and subword incomparable example

Let $A = \{a, b\}$ and $\mathcal{F} = \{\{aabb\}, \{bbaa\}, \{bbabaa\}, \{aa, bb, abab\}\}$.

For f-languages, this $\mathcal{F}$ is not minimal.

Remove $abab$ and consider
$\mathcal{F}' = \{\{aabb\}, \{bbaa\}, \{bbabaa\}, \{aa, bb\}\}$.

For every w such that $abab \notin \underline{sub}(w)$ and $\{aa, bb\} \subseteq \underline{sub}(w)$,
either $aabb \in \underline{sub}(w)$ or $bbaa \in \underline{sub}(w)$ or $bbabaa \in \underline{sub}(w)$.
Thus, $L(\mathcal{F}) = L(\mathcal{F}')$.

Since $\mathcal{F}'$ is not subword incomparable, it can be reduced to
$\mathcal{F}'' = \{\{aa, bb\}\}$.

Connecting words

Given F, x is a *connecting word of* F if and only if $F \subseteq \underline{sub}(x)$.

$C(F) = \{x \mid F \subseteq \underline{sub}(x)\}$

$C_{min}(F)$ is the set of minimal connecting words.

Remark: Given F and $w \in A^*$, either $w \, \underline{con} \, F$ or $w \in C(F)$.

Example: For $A = \{a, b\}$ and $F = \{aa, bb\}$,

$aabbabb \in C(F)$, $aabbabb \notin C_{min}(F)$

$aababb \in C_{min}(F)$

$C_{min}(F) = \{aa(ba)^i bb, bb(ab)^i aa \mid i \geq 0\}$.

Connecting free

$\mathcal{F}$ is *connecting free* if for every $F \in \mathcal{F}$ with $|F| \geq 2$ and every $x \in F$ there exists $w \in C(F \backslash \{x\})$ with $x \notin \underline{sub}(w)$ such that $w \, \underline{con} \, G$ for every $G \in \mathcal{F}$, $G \neq F$.

$\mathcal{F} = \{\{aabb\}, \{bbaa\}, \{bbabaa\}, \{aa, bb, abab\}\}$ is not connecting free.

Proposition 4 *Every connecting free forbidding set is subword free.*

Lemma 5 *Let $\mathcal{F}$ be subword incomparable and connecting free. Then, for every $F \in \mathcal{F}$ and every $x \in F$, $L(\mathcal{F}') \subset L(\mathcal{F})$, where $\mathcal{F}' = (\mathcal{F} \backslash \{F\}) \cup \{F'\}$ such that $F' = F \backslash \{x\}$.*

Connecting reduced

$\mathcal{F}$ is *connecting reduced* if for every $F \in \mathcal{F}$ there exists $w \in C(F)$, such that $w \notin C(G)$ for every $G \in \mathcal{F}$, $G \neq F$.

Let $A = \{a, b\}$, $\mathcal{F} = \{\{ab\}, \{ba\}, \{aa, bb\}\}$, and $F = \{aa, bb\}$. This $\mathcal{F}$ is not connecting reduced.

$C_{min}(F) = \{aa(ba)^i bb, bb(ab)^i aa \mid i \geq 0\}$

Proposition 6 *Every connecting reduced forbidding set is subword incomparable.*

Lemma 7 *Let $\mathcal{F}$ be a connecting reduced forbidding set. Then, $L(\mathcal{F}) \subset L(\mathcal{F} \setminus \{F\})$ for any $F \in \mathcal{F}$.*

Reduced

$\mathcal{F}$ is *reduced* if it is both connecting free and connecting reduced.

Theorem 8 *Every reduced forbidding set is minimal.*

$\mathcal{F}_1 = \{\{ab\}, \{ba\}\}$ is reduced.

$\mathcal{F}_2 = \{\{ab, ba\}, \{aa, bb\}\}$ is also reduced.

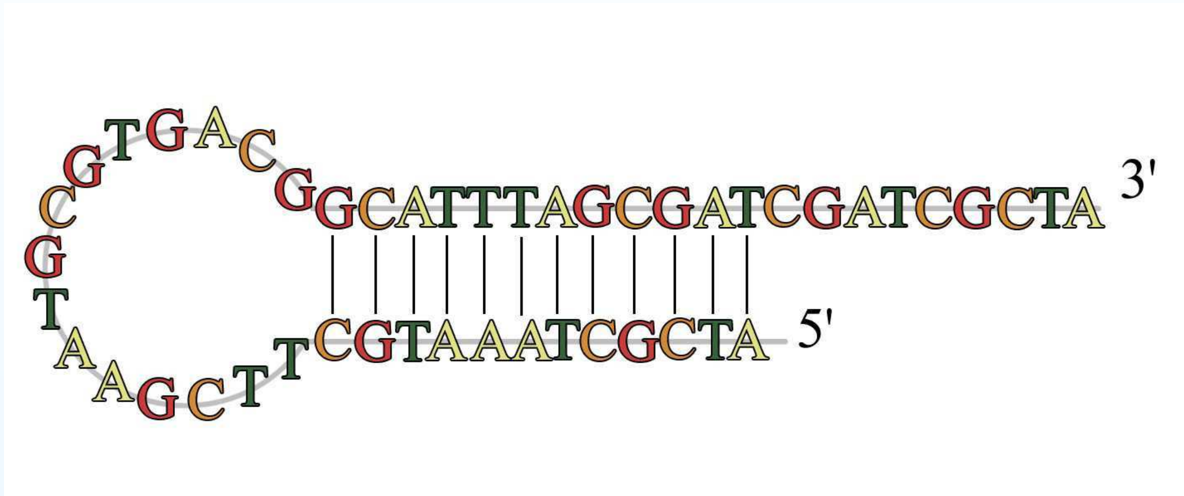
Strict forbidding sets

Theorem 9 *For every forbidding set there exists an equivalent strict forbidding set.*

Construction: Given $\mathcal{F}$, for every $F \in \mathcal{F}$ construct $\mathcal{F}_F = \{\{s\} \mid s \in C_{min}(F)\}$ and consider $\mathcal{F}' = \cup_{F \in \mathcal{F}} \mathcal{F}_F$.
Then, $L(\mathcal{F}') = L(\mathcal{F})$.

Theorem 10 *For every forbidding set there exists an equivalent unique minimal strict forbidding set.*

Defining DNA code words



Given alphabet Σ, let $\theta : \Sigma^* \to \Sigma^*$ be a morphic or antimorphic involution.

The set X is called *θ-subword-k-code* if for all $u \in \Sigma^*$ such that $|u| = k$ we have $\Sigma^* u \Sigma^i \theta(u) \Sigma^* \cap X = \emptyset$ for all $i \geq 1$.

Let $k \geq 1$ be an integer. Consider $\mathcal{F} = \{\{u, \theta(u)\} \mid u \in \Sigma^k\}$. Then for every $X \subseteq (L(\mathcal{F}) \setminus \{\lambda\})$, X is a θ-subword-k-code.

Acknowledgement

This work has been supported in part by a UNF Faculty Development Scholarship Grant.

References

A. Ehrenfeucht, G. Rozenberg, *Forbidding-enforcing systems*. Theoretical Computer Science **292** (2003) 611–638.

A. Ehrenfeucht, H.J. Hoogeboom, G. Rozenberg, N. van Vugt, *Forbidding and enforcing*, In: E. Winfree, D.K. Gifford (eds.) DNA Based Computers V. AMS DIMACS **54**, Providence, RI, (2000) 195-206.

D. Genova, *Defining Languages by Forbidding-Enforcing Systems*, In: B. Lowe, D. Normann (eds.) Models of Computation in Context, Computability in Europe CiE 2011, Lecture Notes in Computer Science, **6735** (2011) 92-101.

D. Genova, N. Jonoska, *Topological Properties of Forbidding-Enforcing Systems*, Journal of Automata, Languages and Combinatorics **11** No. 4 (2006) 375-397.

Thank you!