Walkin' in free Inverse Monoids

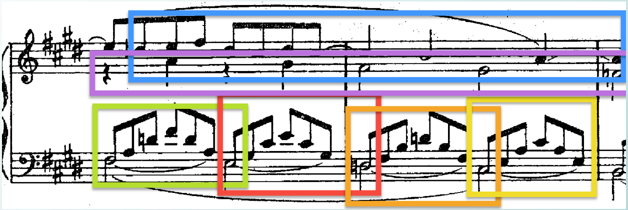
(how to bring coal to newcastle ?)

NBSAN meeting
York
Wednesday, the 21st of november 2012

David Janin,
LaBRI, Université de Bordeaux

Music modeling

The structure of music is complex with mixed sequential, parallel and hierarchical features.



A theory of overlapping structures is needed for computer music analysis and/or production.

Observation

Inverse semigroup theory provides almost everything we need for music analysis [Jan12c] or for music design and production [BJM12].

1. Playground

Within free inverse monoids

Bi-rooted trees

The free inverse monoid and its Rees' quotients

$$\begin{tikzcd} \text{Walks} & (A+\bar{A})^* \arrow[r, two heads, "\eta"] \arrow[d, two heads, "\theta"'] & (A+\bar{A})^*/\theta^{-1}(\perp) \arrow[d, two heads, "\theta"'] \\ \text{BTrees} & FIM(A) \arrow[r, two heads, "\eta"] & FIM(A)/\perp \end{tikzcd}$$

Examples

Typical models defined by choosing adequate ideals.

- *directed trees* generated by $\perp = \langle \{a\bar{b}\}_{a,b\in A, a\neq b}\rangle$.
- McAlister *tiles* generated by $\perp = \langle \{a\bar{b}, \bar{a}b\}_{a,b\in A, a\neq b}\rangle$.

Birooted F-terms

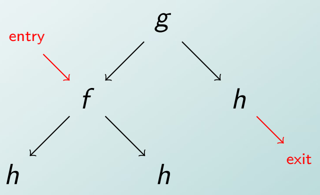
Signature

A finite alphabet of function names F and an arity mapping $\rho : F \to \mathcal{P}(E)$ with finite set E of argument names.

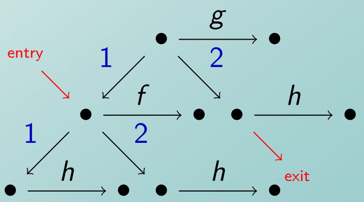
Example

$F = \{f, g, h\}$ with $\rho(f) = \{1, 2\}$, $\rho(g) = \{1, 2\}$ and $\rho(h) = \emptyset$.

F-tree $g(f(h, h), h)$



encoded as a bi-rooted tree



Birooted F-trees

Signature

A finite alphabet of function names F and an arity mapping $\rho : F \to \mathcal{P}(E)$.

Observation

Birooted F-trees can be embedded into $FIM(A)/\perp$ with alphabet $A = F + \{(f, e, g) \in F \times E \times F : e \in \rho(f)\}$ and $\perp$ the ideal of bad encodings, i.e. birooted trees that do not define a partial F-tree.

Observation

Complete (finite) birooted F-trees are minimal non zero elements in the natural order.

Examples

- $F = \{1\}$ with $\rho(1) = A$, essentially directed trees,
- $F = A$ with $\rho(a) = \{1\}$, essentially McAslister tiles.

2. Languages of birooted trees

Towards a birooted tree language theory

Three classical classes of languages

REC

Languages $L \subseteq FIM(A)$ recognizable by morphism, i.e. there is morphism $\varphi : FIM(A) \to S$ with finite S such that $L = \varphi^{-1}(\varphi(L))$.

RAT

Languages $L \subseteq FIM(A)$ definable by a rational expression, i.e. a finite combination of finite languages with sum $+$, product $\cdot$ and iterated product (Kleene star) $*$.

MSO

Languages $L \subseteq FIM(A)$ definable by an formula of Monadic Second Order logic (MSO), i.e. $L = \{x \in FIM(A) : x \models \varphi_L\}$ for some MSO definable characteristic property φ_L of L.

Separation result

Theorem (Buchi, Elgot)

Within A^* we have $REC = RAT = MSO$.

Theorem ([Jan13, DJ12])

Within $FIM(A)$ (or even M_A) we have $REC \subset RAT \subset MSO$ with strict inclusion.

3. Tile languages

McAlister monoid

Positive tiles encoded as triples $(u, v, w) \in A^* \times A^* \times A^*$

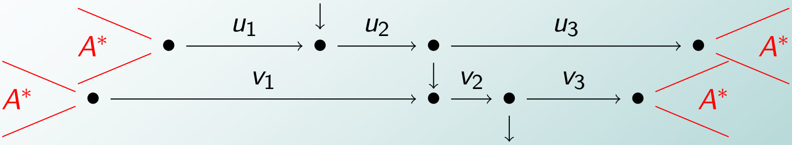
$$\bullet \xrightarrow{\quad u \quad} \overset{\downarrow}{\bullet} \xrightarrow{\quad v \quad} \underset{\downarrow}{\bullet} \xrightarrow{\qquad\qquad w \qquad\qquad} \bullet$$

and negative tiles encoded as $(uv, \bar{v}, vw) \in A^* \times \bar{A}^* \times A^*$

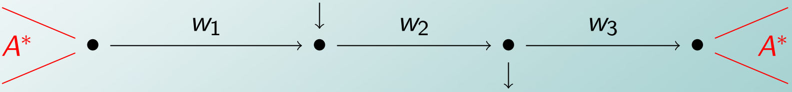
$$\bullet \xrightarrow{\quad u \quad} \underset{\downarrow}{\bullet} \xrightarrow{\quad v \quad} \overset{\downarrow}{\bullet} \xrightarrow{\qquad\qquad w \qquad\qquad} \bullet$$

Tiles product

Given two tiles encoded as $u = (u_1, u_2, u_3)$ and $v = (v_1, v_2, v_3)$,



there is at most one tile $w = (w_1, w_2, w_3)$



- left match : $A^* w_1 = A^* u_1 \cap A^* v_1 \bar{u}_2$,
- right match : $w_3 A^* = v_3 A^* \cap \bar{v}_2 u_3 A^*$.

In that case we take $u \cdot v = w$ and otherwise we take $u \cdot v = 0$.
The resulting tile monoid, McAlister monoid, is denoted by M_A.

MSO

Operators on languages

- Sum: $X + Y = X \cup Y$,
- Product: $X \cdot Y = \{xy : x \in X, y \in Y\}$,
- Iterated product (star): $X^* = \sum_{k \in \mathbb{N}} X^k$,
- Idempotent proj.: $X^E = \{x \in X : xx = x\} = X \cap E(FIM(A))$.
- Inverse: $X^{-1} = \{x^{-1} : x \in X\}$.

Theorem (Robustness [Jan13], [Jan12d])

The class MSO of languages of tiles is closed under complement, sum, product, iterated product (star), inverses, idempotent projections.

MSO

Theorem (Simplicity [Jan13])

For every MSO language L, given L^+ (resp. L^-) the set of positive (resp. negative) tiles in L, we have:

$$L^+ = \sum_{k \in I} L_k \times C_k \times R_k \text{ and, resp. } L^- = \sum_{k \in J} (L_k \times C_k \times R_k)^{-1}$$

for finite I and J and regular word languages L_k, C_k and $R_k \subseteq A^*$.

Proof.

An MSO definable language of positive tiles is just an MSO definable language of words in $A_p^* A^* A_s^*$ with A_p and A_s two disjoint copies of A. □

MSO

Definition

Let E-RAT be the class of languages definable by means of sum, product, star and idempotent projection.

Theorem (MSO = E-RAT [DJ12])

Language $L \subseteq M_A$ is MSO if and only if it is definable by sum, product, star and idempotent projection of finite languages.

Proof.

E-RAT is closed under inverse operator and, for every regular L, C and $R \subseteq A^*$, $L \times C \times R = (1 \times L \times 1)^L \cdot (1 \times C \times 1) \cdot (1 \times R \times 1)^R$ with $X^L = \{x^{-1}x : x \in X\}$, $X^R = \{xx^{-1} : x \in X\}$ and that fact that $X^L = (X^{-1}X)^E = (X^{-1})^R$. $\square$

RAT

Fact

There is an ideal $\perp \subseteq (A+\bar{A})^*$ such that:

$$\begin{tikzcd} (A+\bar{A})^* \arrow[r, two heads, "/\perp"] \arrow[d, two heads, "\theta"] & (A+\bar{A})^*/\perp \arrow[d, two heads, "\theta"] \\ FIM(A) \arrow[r, two heads, "/\perp"] & M_A \end{tikzcd}$$

(walks)

(structures)

Theorem ([DJ12])

Language $L \subseteq M_A$ is RAT if and only if $L = \theta(W)$ for some regular language $W \subseteq (A+\bar{A})^*$.

RAT

Corollary

Language $L \subseteq M_A$ is RAT if and only if L recognizable by a finite walking automaton.

Proof.

Take the one way automaton on alphabet $A + \bar{A}$ that recognizes $W \subseteq (A + \bar{A})^*$ with $L = \theta(W)$.
Interpret it as a two-way automaton on tiles that recognizes $\theta(W) \subseteq M_A$. □

Corollary

The inclusion RAT $\subset$ MSO is strict as witnessed by $L = E(M_A)$ and a simple pumping argument (on the underlying walking automaton).

RAT

Corollary

If language $L \subseteq M_A$ is RAT then $L = \psi^{-1}(\psi(L))$ for some finite monoid S and relational morphism $\psi : M_A \to S$.

$$
\begin{tikzcd}
 & (A+\bar{A})^* \arrow[dl, "\theta"'] \arrow[dr, "\varphi"] & \\
M_A \arrow[rr, "\psi = \varphi \circ \theta^{-1}"] & & S
\end{tikzcd}
$$

Question

Does this lead to an interesting characterization of RAT ?

REC

Lemma

The inclusion REC $\subset$ RAT is strict as witnessed by $L = 1 \times ba^* \times 1$ that has a syntactic congruence of infinite index.

Theorem ([Jan13])

For every morphism $\varphi : M_A \to S$, every $s \in S - 0$, there are x and $y \in A^*$ such that: $\varphi^{-1}(s)$ is essentially a co-finite subset of tiles of the form (u, v, w) with ${}^{\omega}(xy) \geq_s u$, $v \in x(yx)^*$, $w \leq_p (yx)^{\omega}$.

Proof.

Let $\varphi : M_A \to S$ for some monoid S (even infinite). Let $s \in S - 0$. Then $\varphi^{-1}(s)$ is totally ordered both by left and right Green's preorder.

...and some combinatorics to conclude... $\square$

4. Walking in FIM(A)

Walk languages vs tree languages

Observation

Reading words of $(A+\bar{A})^*$ amount to walking on some underlying birooted trees.

Walk languages

Language $W \subseteq (A+\bar{A})^*$ is a walk language of the tree language $L \subseteq FIM(A)/\bot$ when $L + 0 = \eta \circ \theta(W) + 0$.

$$\begin{tikzcd}
(A+\bar{A})^* \arrow[rr, two heads, "\eta \circ \theta"] \arrow[dr, two heads, "\theta"'] & & FIM(A)/\bot \\
& FIM(A) \arrow[ur, two heads, "\eta"'] &
\end{tikzcd}$$

Question

How classes of tree languages in $FIM(A)/\bot$ are related with classes of the underlying walk languages in $(A+\bar{A})^*$?

Walking automata

Theorem ([Jan12d, DJ12])

- *REC = Strongly deterministic finite state walking Automata,*
- *RAT = Finite state walking automata,*
- *MSO = Many-Pebble finite state walking automata.*

Fact

$REC \neq RAT$ witnessed by ba^*.
$RAT \neq MSO$ witness by $E(FIM(A))$.

MSO and the pebble hierarchy

Idempotent projection

For every language X, let $X^E = \{x \in X : xx = x\}$.

k-rational languages

Language L is k-rational when either L is rational or $k > 0$ and L is a finite rational combination of languages of the form X or X^E with $X \in RAT^{k-1}$.

Fact

RAT^k is closed under inverses for every $k \in \mathbb{N}$.

Theorem ([Jan12d])

$REC \subset RAT = RAT^0 \subset RAT^1 \subseteq RAT^2 \subseteq \cdots \bigcup_k RAT^k \subseteq MSO$

probably with strict inclusions.

Theorem ([Jan13, DJ12])

Over tiles $REC \subset RAT \subset RAT^1 = MSO$.

5. Quasi-recognizability

A newcomer question

Fact

Within FIM(A), the class REC collapses.

Question

How to relax the notion REC into some (notion of) quasi-REC (QREC) in such a way MSO = QREC (in relevant case) ?

Ideas

1. relax morphism condition $\varphi(xy) = \varphi(x)\varphi(y)$ into premorphism condition $\varphi(xy) \leq \varphi(x)\varphi(y)$.
2. restrict to an adequate class of finite (ordered) monoid and premorphism in such a way that pre-images remain MSO definable.

Adhoc candidates for QREC

QREC points

Stable ordered monoid $(S, \leq)$ such that:

- $U(S) = \{x \leq 1\} \subseteq E(S)$, i.e. subunits are idempotents,
- for all $x \in S$, both $x_R = \bigwedge\{e \in U(S) : ex = x\}$ and $x_L = \{f \in U(S) : xf = x\}$ exist in $U(S)$,
- for all x and $y \in S$, if $x = x_R y x_L$ then $x \leq y$. .

QREC arrows

Premorphism $\varphi : FIM(A)/\bot \to (S, \leq)$, i.e. $\varphi(1) = 1$ and, for every x and y, if $x \leq y$ then $\varphi(x) \leq \varphi(y)$ and $\varphi(xy) \leq \varphi(x)\varphi(y)$, such that:

- for every disjoint product $x \cdot y$, we have $\varphi(x \cdot y) = \varphi(x)\varphi(y)$,
- for every x, we have $\varphi(x_L) = (\varphi(x))_L$ and $\varphi(x_R) = (\varphi(x))_R$.

with $x_L = x^{-1}x$ and $x_R = xx^{-1}$ in $FIM(A)/\bot$.

QREC vs MSO

Let $M_A^+ = 0 + A^* \times A^* \times A^*$ be the submonoid of M_A of *positive tiles*.

Theorem ([Jan12b])

If $L \subseteq M_A^+$ is QREC then L is MSO.

Theorem ([Jan12b])

If $L \subseteq M_A^+$ is MSO and if tiles of L are plugged, i.e. with tiles of the form $(\#u, v, w\#)$ for some marker $\#$, then L is QREC.

$\mathcal{Q}$-expansion

Monoid $\mathcal{Q}$-expansion

Let S be a monoid. Let $\mathcal{Q}(S) = 0 + \mathcal{L}_S \times S \times \mathcal{R}_S$ with

$$(L, s, R) \cdot (M, t, N) = (L \cap (M)s^{-1}, st, t^{-1}(R) \cap N)$$

when compatible, and 0 otherwise.

Theorem ([Jan12a])

For every monoid S, monoid $\mathcal{Q}(S)$ ordered by $(L, s, R) \leq (M, t, N)$ when $L \subseteq M$, $s = t$ and $R \subseteq N$ is a stable U-semiadequate monoid.

Theorem

There is an embedding $\iota_A : M_A^+ \to \mathcal{Q}(A^*)$.

$\mathcal{Q}$-expansion

Morphism $\mathcal{Q}$-expansion

Let $\varphi : S \to T$. Let $\mathcal{Q}(\varphi) : \mathcal{Q}(S) \to \mathcal{Q}(T)$ defined, on every non zero positive tile (L, s, R), by

$$\mathcal{Q}(\varphi)(L, s, R) = (S\varphi(L), \varphi(s), \varphi(R)S)$$

and let $\eta_S : \mathcal{Q}(S) \to S^0$ defined by $\eta_S((L, s, R)) = s$.

Theorem ([Jan12a])

For every morphism $\varphi : S \to T$, mapping $\mathcal{Q}(\varphi)$ is a well-behaved premorphism (i.e. QREC arrows) and the following diagram commutes.

$$\begin{tikzcd}
\mathcal{Q}(S) \arrow[r, "\mathcal{Q}(\varphi)"] \arrow[d, "\eta_S"'] & \mathcal{Q}(T) \arrow[d, "\eta_T"] \\
S^0 \arrow[r, "\varphi"] & T^0
\end{tikzcd}$$

Theorem

For every plugged language $L \subseteq M^{+}_{A+\#}$, if L is MSO then, given $\varphi : A^* \to S$ "recognizing" $L \subseteq \#A^* \times A^* \times A^*\#$, then L is QREC by $\mathcal{Q}(\varphi) : M_A \to \mathcal{Q}(S)$.

6. Conclusion

Work in progress

Extending/developing QREC towards:

- languages of positive and negative tiles,
- languages of finite directed trees,
- languages of finite and infinite trees (completing $FIM(A)$ with infinite many rooted trees).

F. Berthaut, D. Janin, and B. Martin.
Advanced synchronization of audio or symbolic musical patterns.
In *Sixth IEEE International Conference on Semantic Computing*, pages 202–209. IEEE Society Press, 2012.

A. Dicky and D. Janin.
Two-way automata and regular languages of overlapping tiles.
Technical Report RR-1463-12, LaBRI, Université de Bordeaux, 2012.

D. Janin.
Quasi-inverse monoids (and premorphisms).
Technical Report RR-1459-12, LaBRI, Université de Bordeaux, 2012.

D. Janin.
Quasi-recognizable vs MSO definable languages of one-dimentionnal overlaping tiles.

In *Mathematical Foundations of computer Science (MFCS)*, volume 7464 of *LNCS*, pages 516–528, 2012.

D. Janin.
Vers une modélisation combinatoire des structures rythmiques simples de la musique.
Revue Francophone d'Informatique Musicale (RFIM), 2, 2012.

D. Janin.
Walking automata in the free inverse monoid.
Technical Report RR-1464-12, LaBRI, Université de Bordeaux, 2012.

D. Janin.
On languages of one-dimensional overlapping tiles.
In *International Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM)*, volume (to appear) of *LNCS*, 2013.