Gröbner bases and the automaton property of Hecke-Kiselman algebras

Arkadiusz Męcel

`Young Researchers Colloquium, IMPAN 24.05.2019 r.`

This talk concerns a certain class of algebraic non-commutative objects, called Hecke-Kiselman algebras, that are being under research in the Algebra Group at the University of Warsaw. On a very elementary level, non-commutative problems arise simply because the composition operation of two functions is surprisingly often not commutative (even if it can be done both ways). If those functions concern discrete objects as finite sets, graphs, posets, we tend to disguise them in two fundamental ways - as matrices or as words. In the first case an algebraist will often say that there is a matrix representation, or a linear representation of some class of objects, and in the second case we talk about the presentation for a class of objects. If there is an algebraic action involved, then a corresponding operation should appear in the representation or in the presentation of those objects. It would also be very nice if different objects were represented differently. This is what we look after.

The most elementary example that will lead us into Hecke-Kiselman monoid is the symmetric group, the set of bijective self-maps of $\{1, \ldots, n\}$. You all know that these permutations can be represented as 0-1 matrices of size $n \times n$ with exactly one non-zero element in each row and column. The matrix multiplication corresponds to the permutation composition. Of course, there is a question if you could somehow use smaller matrices, maybe with different entries, like complex numbers, or other fields. What would be the sizes of these matrices? What would those sizes mean? These are representation theory questions. From a combinatorial point of view you can argue differently. You introduce the set of elementary transpositions $s_1, \ldots, s_{n-1}$, where s_i swaps i and $i+1$ and everything else stays the same. Now its easy to prove that every permutation is a finite composition of such transpositions. We can represent such function as a word, for instance $s_2 s_3 s_{n-1} s_1$. When two words of this kind meet, there's a set of rules to multiply them. Here are those rules.

(i) $s_i^2 = 1$, where $1 \leqslant i \leqslant n$,

(ii) $s_i s_j = s_j s_i$, if $|i-j| > 1$,

(iii) $s_i s_j s_i = s_j s_i s_j$, if $|i-j| = 1$.

This choice of generators $s_1, \ldots, s_{n-1}$ and the relations above is called the presentation of the symmetric group. This is the second approach. Represent an algebraic objects a words and relations. This approach is very tricky. First, we need to know that one group can disguise itself in many different presentations. It is often almost impossible to say if two presentations define the same group or not. Moreover, while having a set presentation of an algebraic object we can have two words and no idea how to relate them to each other. Are they equal? This is a famous word problem. In the 60' it was solved by Tits not only in the case of permutation groups, but also in the case of the so-called Coxeter groups, and more general - in a context of groups with BN-pairs. For solving this problem we will use the terminology of normal words and Gröbner bases. Having said all of this I will now introduce the Hecke-Kiselman monoid.

1 Hecke-Kiselman monoids

Consider a finite graph Θ with n vertices $\{1,\dots,n\}$. We assume that these vertices can be connected by unoriented edges and also by oriented arrows, but for each pair of vertices there is at most one edge or arrow that connects them. This is called a finite simple digraph. For such graph Θ we define a monoid HK_Θ generated by n elements $x_1,\dots,x_n$ that are related by the following conditions.

(i) $x_i^2 = x_i$, where $1 \leqslant i \leqslant n$,

(ii) if the vertices i, j are not connected in Θ, then $x_ix_j = x_jx_i$,

(iii) if i, j are connected by an arrow $i \to j$ in Θ, then $x_ix_jx_i = x_jx_ix_j = x_ix_j$,

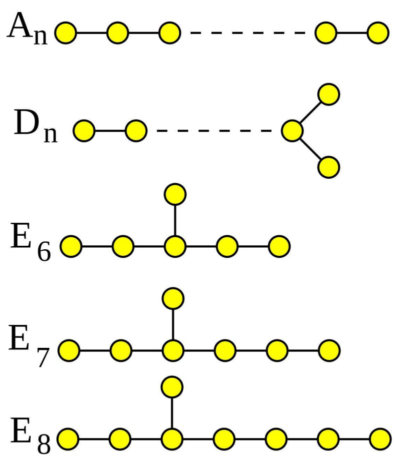
(iv) if i, j are connected by an (unoriented) edge in Θ, then $x_ix_jx_i = x_jx_ix_j$.

This presentation yields the Hecke-Kiselman monoid HK_Θ of a graph Θ. So we define the set words and some composition rules. Here there is an additional underlying structure of a graph. For some of us this should come as no surprise. When we talked about the symmetric group, there was also a graph involved. If we actually assumed Θ to be an unoriented chain of $n-1$ vertices and if we generated a group of $n-1$ generators that satisfy (ii), (iv), along with a modified relation (i) of form $x_i^2 = 1$, then we would obtain exactly a presentation of the symmetric group S_n. This is no coincidence, as for every simple unoriented graph Θ of vertices $\{1,2,\dots,n\}$ there exists a corresponding finitely generated Coxeter group W_Θ of presentation:

$$(1)\, x_i^2 = 1, \quad (2)\, (x_ix_j)^2 = 1, \quad (3)\, (x_ix_j)^3 = 1,$$

where the condition (1) is satisfied by all generators, (2) is satisfied by those generators which correspond to non-connected vertices Θ and (3) is satisfied for generators corresponding to connected vertices $i-j$. It is obvious, that via group manipulations of equalities (1)-(3) one can obtain a presentation that resembles that of a Hecke-Kiselman monoid in a way described above. These are called simply laced Coxeter groups, as such groups can be formed also for graphs with multiple connections between vertices of an unoriented graph. These groups describe symmetries in the real space and appear in many other algebraic settings.

To make the resemblance between the Coxeter group W_Θ and a Hecke-Kiselman monoid HK_Θ clear we should mention that the latter in the unoriented case is called a 0-Hecke monoid, or a Coxeter monoid. This name comes from the representation theory, and originates from the modular representation theory of Coxeter groups, in particular a symmetric group. In essence, this is a specialization of the so called Hecke-Iwahori algebra $H_0(W_\Theta, v)$, a unital algebra generated by generators T_i such that are related by braid relations (they correspond to the ones for W_Θ) and the relation of kind (i) is modified to the form $(T_i-q)(T_i+1) = 0$, where q is an indeterminate. Roughly speaking, when $q = 1$ this is just a group algebra of a Coxeter group, and when $q = 0$ this is exactly a semigroup algebra of our Hecke-Kiselman monoid. The last important bit of information is that the dimension of the Iwahori-Hecke algebra is the same as the number of elements in the Coxeter group, no matter the specialization of indeterminate q. Therefore, it can be proved, that there is a bijection between the Coxeter group, and the associated Coxeter monoid. It is a very strong one as the reduced forms of words in these two objects are completely the same. This yields, for instance, that the finiteness problem for Hecke-Kiselman monoids in the unoriented case is in fact equivalent to the finiteness problem for Coxeter groups, and it is thus solved years ago. The graph, for which this monoid is finite are exactly the simply laced Dynkin diagrams.



The notion of HK_Θ was introduced by Ganyushkin and Mazorchuk in 2011 as an attempt to generalize a class of Kiselman monoids and their quotients, which corresponds to the case when the graph Θ is oriented. I will not go into too many details, but I would to show you two examples of Hecke-Kiselman monoids in the oriented case.

Consider an oriented chain in $n-1$ vertices of form $1 \to 2 \to \ldots \to n-1$. The Hecke-Kiselman semigroup that arises is finite and it is the so-called Catalan C_n monoid of all order-preserving, weakly increasing self-maps f of $\{1,\ldots,n\}$. It is well known that the cardinality of C_n is the n-th Catalan number. This is a very interesting combinatorial object, but i wonder if it really easy to see the connection between the presentation and the combinatorial object? Of course not, and this is very typical.

It is, by no means completely trivial to say when in the oriented Θ case, is such semigroup HK_Θ finite. Even if we take an oriented cycle X_n of n-elements, which can be called an affine Catalan monoid, it does not seem immediately obvious from the defining relations how to prove that it is infinite. If this cycle is of form $x_1 \to x_2 \to \ldots \to x_n \to x_1$ then an obvious candidate for an infinite-order element is $x_1 \ldots x_n$, but it is not easy to prove this. If you can, tell me. I know two methods, one very tricky and one very tedious. This is what we try to accomplish here - the questions on the finiteness, the uniqueness of word presentations, and ultimately - on the internal structure of such monoids, their ideals, radicals of their algebras and so on.

2 Gröbner bases

The questions on presentation, the word problems for instances, can be attacked via the method of Gröbner bases. To use it, we will expand our Hecke-Kiselman structure, from the monoid structure, to the structure of an semigroup algebra. To put it very briefly - instead of considering the HK_Θ monoid whose elements are represented (perhaps not uniquely) by words from the alphabet $x_1,\ldots,x_n$, we consider the linear span of those words - the non-commutative polynomials in $x_1,\ldots,x_n$. They can be added the same way polynomials are always added, but you multiply using the presentation rules for multiplying monomials that represent elements of the Hecke-Kiselman monoids.

The use of Gröbner bases is a standard method of checking if a certain polynomial belongs to an ideal generated by a finite numbers of polynomials. The method is to transform the polynomial generators to a set of „nicer generators" by introducing some kind of ordering on the monomials and by some kind of division algorithm. As a result: if the leading elements of a probed polynomial divides a leading polynomial of those "nice generators" this means that our candidate indeed belongs to the ideal.

You have been doing this since you first lean the Gauss algorithm of row reducing the matrix of a set of linear equations. For example if $f_1 = x_1+x_2-1$ and $f_2 = x_1-x_2+2$, then Gaussian elimination uses term x_1 in f_1 as a pivot, and replaces f_2 with $f_2 := f_2 - f_1 = -2x_2+3$, and back substitution uses the term $-2x_2$ in the new f_2 as a pivot to remove the x_2 term form f_1. One can obtain the original polynomials as linear combinations of $x_1 + \frac{1}{2}$ and $-2x_2+3$. The similar thing happens when you consider $f_1 = x^3 - x^2 - 2x$, $f_2 = x^2 - 3x + 2$. The Euclidean algorithm attempts to uncover new lead terms by canceling lead terms. We quickly realize that the ideal generated by f_1 and f_2 is in fact the ideal generated by $f = x - 2$. So the system $f_1 = f_2 = 0$ has a unique solution. If we stay in the commutative case, but go from the one variable-case to the mutltivariable case, we get another algorithm called the Buchberger algorithm, which again uses some kind of ordering on the monomials and the reductions process. In the examples above we had two orders on monomials: $x_2 > x_1 > 1$ and $x^3 > x^2 > x > 1$. This generalizes further. In the commutative setting similar procedures end in finitely many steps and computers can be used. In the non-commutative settings things do not work that smooth as the procedure for obtaining the so-called Gröbner basis of an ideal may not terminate after finitely many steps. But if we obtain a basis somehow (finite or not), there is a way to retrieve a particularly elegant basis of a quotient algebra. The method of obtaining such basis requires the use of the so-called diamond lemma. Let us recall some details.

Let F denote the free monoid on the set X of $n \geqslant 3$ free generators $x_1, \ldots, x_n$. Let k be a field and let $k[F] = k\langle x_1, \ldots, x_n \rangle$ denote the corresponding free algebra over k. Assume that a well order $<$ is fixed on X and consider the induced degree-lexicographical order on F (also denoted by $<$). Let A be a finitely generated algebra over k with a set of generators $r_1, \ldots, r_n$ and let $\pi : k[F] \to A$ be the natural homomorphism of k-algebras with $\pi(x_i) = r_i$. We will assume that $\ker(\pi)$ is spanned by elements of the form $w - v$, where $w, v \in F$ (in other words, A is a semigroup algebra). Let I be the ideal of F consisting of all leading monomials of $\ker(\pi)$. The set of normal words corresponding to the chosen presentation for A and to the chosen order on F is defined by $N(A) = F \setminus I$.

A subset G of an ideal $J = \ker(\pi)$ is called a Gröbner basis of J (or of $A = k[F]/J$) if $0 \notin G$, J is generated by G as an ideal and for every nonzero $f \in J$ there exists $g \in G$ such that the leading monomial $\overline{g} \in F$ of g is a factor of the leading monomial $\overline{f}$ of f. A word $w \in F$ is normal if and only if w has no factors that are leading monomials in $g \in G$.

The so-called diamond lemma, or as others say: a composition-lemma is often used in this context. By a reduction in $k[F]$ determined by a pair $(w, w') \in F^2$, where $w' < w$ (the deg-lex order of F), we mean any operation of replacing a factor w in a word $f \in F$ by the factor w'. For a set $T \subseteq F^2$ of such pairs (these pairs will be called reductions as well) we say that the word $f \in F$ is T-reduced if no factor of f is the leading term w of a reduction (w, w') from the set T. The deg-lex order on F satisfies the descending chain condition, which means there is no infinite decreasing chain of elements in F. This means that a T-reduced form of a word $w \in F$ can always be obtained in a finite series of steps. The linear space spanned by T-reduced monomials in $k[F]$ is denoted by $R(T)$.

The diamond lemma gives necessary and sufficient conditions for the set $N(A)$ of normal words to coincide with the set of T-reduced words in F. The key tool is the notion of ambiguity. Let $\sigma = (w_\sigma, v_\sigma)$, $\tau = (w_\tau, v_\tau)$ be reductions in T. By an overlap ambiguity we mean a quintuple (σ, τ, l, w, r), where $1 \neq l, w, r \in F$ are such that $w_\sigma = wr$ and $w_\tau = lw$. A quintuple (σ, τ, l, w, r) is called an inclusive ambiguity if $w_\sigma = w$ and $w_\tau = lwr$. For brevity we will denote these ambiguities as $l(wr) = (lw)r$ and $l(w)r = (lwr)$, respectively. We will also say that they are of type σ-τ. We say that the overlap (inclusive, respectively) ambiguity is resolvable if $v_\tau r$ and lv_σ (v_τ and $lv_\sigma r$, respectively) have equal T-reduced forms. Recall the following simplified version of Bergman's diamond lemma.

Lemma 1. *Let T be a reduction set in the free algebra $k[F]$ over a field k, with a fixed deg-lex order in the free monoid F over X. Then the following conditions are equivalent:*

- *all ambiguities on T are resolvable,*
- *each monomial $f \in F$ can be uniquely T-reduced,*
- *if $I(T)$ denotes the ideal of $k[F]$ generated by $\{w - v : (w, v) \in T\}$ then $k[F] = I(T) \oplus R(T)$ as vector spaces.*

Moreover if the conditions above are satisfied then the k-algebra $A = k[F]/I(T)$ can be identified with $R(T)$ equipped with a k-algebra structure with $f \cdot g$ defined as the T-reduced form of fg, for $f, g \in R(T)$. In this case, $\{w - v : (w, v) \in T\}$ is a Gröbner basis of A.

3 Some results and open questions

The use of diamond-lemma involves somehow-guessing the appropriate reduction set for an algebra, even if it can be derived algorithmically (indeed, we use some packages of GAP to do certain computations) This was successful in the case of Hecke-Kiselman algebras of oriented graphs, as we were able to prove:

Theorem 2 (Okniński, M. (2018))**.** *Let Θ be a finite simple oriented graph with vertices $V(\Theta) = \{x_1, x_2, \ldots, x_n\}$. Extend the natural ordering $x_1 < x_2 < \cdots < x_n$ on the set $V(\Theta)$ to the deg-lex order on the free monoid $F = \langle V(\Theta)\rangle$. Consider the following set T of reductions on the algebra $k[F]$:*

(i) (twt, tw), for any $t \in V(\Theta)$ and $w \in F$ such that $w \nrightarrow t$,

(ii) (twt, wt), for any $t \in V(\Theta)$ and $w \in F$ such that $t \nrightarrow w$,

(iii) (t_1wt_2, t_2t_1w), for any $t_1, t_2 \in V(\Theta)$ and $w \in F$ such that $t_1 > t_2$ and $t_2 \nleftrightarrow t_1w$.

where for $t \in V(\Theta)$ and $w \in F = \langle V(\Theta)\rangle$ we write $w \nrightarrow t$ if $t \notin \mathrm{supp}(w)$ and there are no $x \in \mathrm{supp}(w)$ such that $x \to t$ in Θ. Similarly, we define $t \nrightarrow w$: again we assume that $t \notin \mathrm{supp}(w)$ and there is no arrow $t \to y$, where $y \in \mathrm{supp}(w)$. Moreover, when $t \nrightarrow w$ and $w \nrightarrow t$, we write $t \nleftrightarrow w$. Then the set $\{w - v,$ where $(w, v) \in T\}$ forms a Gröbner basis of the algebra $k[\mathrm{HK}_\Theta]$.

In other words, if a word in F does not contain a factor (a block of letters) of form twt, as in (i) or (ii) or of form t_1wt_2, then it is reduced. And all elements of HK_Θ can be obtained in that way, as T-reduced elements.

This result, allows a number of corollaries, some of which were obtained in the earlier works. Here are some of them:

Theorem 3 (Okninski, M. (2017))**.** *Assume that Θ is a finite oriented simple graph. The following conditions are equivalent.*

(1) *Θ does not contain two different cycles connected by an oriented path of length $\geqslant 0$,*

(2) *A_Θ is a PI-algebra,*

(3) $\mathrm{GKdim}(A_\Theta) < \infty$,

(4) *the monoid HK_Θ does not contain a free submonoid of rank 2.*

Theorem 4 (Okninski, Wiertel (2018))**.** *Assume that Θ is a finite oriented simple graph. The following conditions are equivalent.*

(1) *A_Θ is right Noetherian,*

(2) *A_Θ is left Noetherian,*

(3) *each of the connected components of Θ is either an oriented cycle or an acyclic graph.*

Corollary 5. *The Hecke-Kiselman algebra of an oriented cycle embedds into the matrix algebra over a field.*

These results involved a careful considerations on the oriented cycle graph monoid HK_{X_n}, mentioned before. Its semigroup algebra has a lot of interesting properties. It can be proved that its growth is linear, that is, its Gelfand-Kirillov dimension is 1. Moreover, its Gröbner basis is finite, which is a non-trivial corollary from the description of the Gröbner basis in the general oriented case. With great computational effort of Okniński and Wiertel it was proved, that the cycle monoid allows a chain of ideals whose factors are co-finite with the generalized semigroups of matrix types, known in the classical theory of representations of semigroups. There is no place to comment on how are the ideals of the Hecke-Kilesman algebras studied, but some methods require representation ideas, as the one used before in the proof of the finiteness of the acyclic graph monoid, some require the use of graph automorphisms that yield some word-invariant operations in the monoid.

Another result that followed from those main theorem concerned the automaton property of Hecke-Kiselman monoids - an especially elegant property for a finitely generated algebra.

Theorem 6 (Okninski, M. (2018))**.** *Assume that Θ is a finite oriented simple graph. Then the algebra A_Θ is automaton, that is: the set $N(A)$ of normal words of A is obtained from a finite subset of F by applying a finite sequence of operations of union, multiplication and operation $*$ defined by $T^* = \bigcup_{i\geqslant 1} T^i$, for $T \subseteq F$.*

In other words, and this is famous Kleene theorem, the set of normal words of A is determined by a finite deterministic automaton.

The class of automaton algebras was introduced by Ufnarovskii in the late 80'. The main motivation was to study a class of finitely generated algebras that generalizes the class of algebras that admit a finite Gröbner basis with respect to some choice of generators and an ordering on monomials. The difficulty here lies in the fact that there are infinitely many generating sets as well as infinitely many admissible orderings on monomials to deal with. There are examples of algebras with finite Gröbner bases with respect to one ordering, and infinite bases with respect to the other. Up until recently it was not known whether for any of known examples of automaton algebras with infinite Gröbner bases with respect to certain orderings one could find a better ordering that would yield a finite Gröbner basis. First counterexamples were found by Iyudu and Shkarin.

There are many results indicating that the class of automaton algebras not only has better computational properties but also several structural properties that are better than in the class of arbitrary finitely generated algebras. In our setting there are many examples of algebras with infinite Gröbner basis. Attach even one outgoing arrow to a cycle, and you will obtain an algebra of Gelfand-Kirillov dimension 2, with an infinite Gröbner basis. As there are still many open question on the structure of Hecke-Kiselman algebras, even the ones obtained from the oriented cycle, for instance: what are the prime algebras, semiprime, semiprimitive, are the Hecke-Kiselman algebras cellular?

There is also a context of mixed graphs Θ and the corresponding Hecke-Kiselman algebras. Here all the questions raised in the beginning are still unanswered. Especially in the mixed graphs case, the problem of finiteness seems to be almost impossible to approach.

* * *

Many thanks to Adam Adams and Joachim Jelisiejew for making this talk possible.