THE HOPF ALGEBRA OF GRAPH INVARIANTS

NICOLAS BORIE

Abstract. We propose an algebraic study of the simple graph isomorphism problem. We define a Hopf algebra from an explicit realization of its elements as formal power series. We show that these series can be evaluated on graphs and count occurrences of subgraphs. We establish a criterion for the isomorphism test of two simple graphs by means of occurrence counting of subgraphs. This criterion is deduced from algebraic relations between elements of our algebra.

1. Introduction

Let G be a finite graph, drawn on a paper sheet. G is represented with dots for the vertices and segments between some pairs of vertices. If n counts the number of edges, if a counts the number of adjacent pairs of edges and d the number of pairs of disjoint edges, obviously, we have $a + d = \binom{n}{2}$. Now, we want to play the same game with more complicated patterns and investigate how these counting functions are related. These motivations are not new: similar approaches can be found for example in [9] or in [7] but these papers deal with finite cases where graphs have a fixed number of vertices.

We will take care of two difficult points for giving an algebraic sense to these counting functions. The first one is that one can describe (for example on a computer) only labeled graphs, so we will have to deal with the symmetries consisting in relabeling graphs. The second one is that we do not want to depend on the size of graphs. Counting the number of occurrences of a pattern inside a graph has a meaning whatever the size of the graph. To achieve this goal, we will make sense of these counting function by means of elements of a Hopf algebra $UGQSym$ (Unlabeled Graph Quasi Symmetric functions) which is close to the combinatorial Hopf algebra $GTSym$ introduced by Novelli, Thibon and Thiéry [8].

We will first realize functions counting occurrences of subgraph, as power series in an infinite number of variables. In Section 3, we investigate the Hopf structure of the algebra $UGQSym$. We show how this algebra is connected to invariants of graphs in Section 4 and give a sufficient criterion for two graphs to be isomorphic by means of subgraph occurrences counting. Finally, we apply our results to the reconstruction problem of finite graphs.

2. Subgraph enumeration functions as formal power series

Let $\mathbb{A}$ be an infinite set of variables x_{ij} indexed by pairs (i, j) of positive integers such that $i < j$. An adapted presentation of this alphabet can be made as a triangle as follows:

$$
\begin{array}{ccc|c|cccc}
1 & & & & & & & \\
x_{12} & 2 & & & & & & \\
x_{13} & x_{23} & 3 & & & & & \\
\hline
x_{14} & x_{24} & x_{34} & 4 & & & & \\
\hline
x_{15} & x_{25} & x_{35} & x_{45} & 5 & & & \\
x_{16} & x_{26} & x_{36} & x_{46} & x_{56} & 6 & & \\
x_{17} & x_{27} & x_{37} & x_{47} & x_{57} & x_{67} & 7 & \\
\vdots & & & \vdots & & & & \ddots
\end{array}
$$

Definition 2.1. *Let n be an integer and σ be a permutation of the symmetric group $\mathfrak{S}_n$. We define an action of $\mathfrak{S}_n$ over pairs (i, j) (variables x_{ij}) such that $1 \leqslant i < j \leqslant n$ as follows*

$$\sigma\cdot(i,j) := \begin{cases} (\sigma(i),\sigma(j)) & \text{if } \sigma(i) < \sigma(j), \\ (\sigma(j),\sigma(i)) & \text{otherwise.} \end{cases} \qquad \left(\sigma \cdot x_{ij} := \begin{cases} x_{\sigma(i)\sigma(j)} & \text{if } \sigma(i) < \sigma(j), \\ x_{\sigma(j)\sigma(i)} & \text{otherwise.} \end{cases} \right)$$

For finite graphs over n vertices, this action corresponds exactly to the relabeling action on graphs viewed on the lower triangular parts of their adjacency matrices.

Definition 2.2. *Let $n \geqslant 2$ be an integer and G be a graph over n vertices without isolated vertex. We denote by $((a_{ij}))_{1\leqslant i<j\leqslant n}$ the lower triangular part of the incidence matrix of G when one has chosen any labeling of the vertices of G. We thus define an* invariant power series *$\mathcal{M}_G$ as*

$$\mathcal{M}_G := \sum_{r_1<\cdots<r_n} \left(\sum_{\sigma\in Orb_{\mathfrak{S}_n}(G)} \left(\prod_{1\leqslant i<j\leqslant n} x_{r_ir_j}^{a_{\sigma\cdot(i,j)}} \right) \right),$$

where $Orb_{\mathfrak{S}_n}(G)$ is a set of permutations of $\mathfrak{S}_n$ required to deploy the orbit of the graph G under the relabeling action. For the unique graph over 1 node (loops are not allowed, this graph cannot contain any edge), we set $\mathcal{M}_{\bullet} := 1$.

Here are the first examples of functions $\mathcal{M}_G$.

$\mathcal{M}$  $= \sum_{0<i<j} x_{ij} = x_{12} + x_{13} + x_{23} + x_{14} + x_{24} + x_{34} + \dots$

$\mathcal{M}$  $= \sum_{0<i<j<k} x_{ij}x_{ik} + x_{ij}x_{jk} + x_{ik}x_{jk} = x_{12}x_{13} + x_{12}x_{23} + x_{13}x_{23} + \dots$

$\mathcal{M}$  $= \sum_{0<i<j<k<l} x_{ij}x_{kl} + x_{ik}x_{jl} + x_{il}x_{jk} = x_{12}x_{34} + x_{13}x_{24} + x_{14}x_{23} + \dots$

For each finite labeled graph g with positive integers, we associate a monomial $m(g)$ with g as follows

$$m(g) := \prod_{\substack{0<i<j \\ \text{i and j are linked}}} x_{ij}$$

Proposition 2.3. *For $n \geqslant 2$ an integer and G a graph over n vertices without isolated vertex, an alternative description of $\mathcal{M}_G$ is given by*

$$\mathcal{M}_G = \sum_{\substack{g:\text{labeling of } G \\ \text{using positive integers}}} m(g).$$

Proof. Monomials are in bijection with labeled graphs. The $\mathcal{M}_G$ are infinite sums over some monomials and we can set a canonical representative for each $\mathcal{M}_G$ using two commutative congruence relations over the commutative monoid $(x_{ij}|x_{ij}^2 - x_{ij})^*$. As described in details in [2], a canonical monomial is the maximum in its orbit under the relabeling action for the lexicographic order.

$$\begin{array}{ccc} g_{support} & \xrightarrow{pack} & g_{1\ldots n} \\ {\scriptstyle canonical}\downarrow & & \downarrow{\scriptstyle canonical} \\ G_{support} & \xrightarrow{pack} & G_{1\ldots n} \end{array}$$

For example, we take the monomial (or labeled graph) $x_{25}x_{57}$.

$$\begin{array}{ccc} x_{25}x_{57} & \xrightarrow[pack]{(2,5,7)\to(1,2,3)} & x_{12}x_{23} \\ {\scriptstyle (2,5,7)\to(5,2,7)}\downarrow{\scriptstyle canonical} & & {\scriptstyle canonical}\downarrow{\scriptstyle (1,2,3)\to(2,1,3)} \\ x_{25}x_{27} & \xrightarrow[pack]{(2,5,7)\to(1,2,3)} & x_{12}x_{13} \end{array}$$

□

Let H be a finite graph over n nodes. Let $(a_{ij})_{1\leqslant i<j\leqslant n}$ be the lower triangular part of its incidence matrix corresponding to any labeling of H with integers $\{1,\ldots,n\}$. Now complete the triangle a_{ij} with an infinite number of 0 for a_{ij} where $j>n$. The evaluation of $\mathcal{M}_G$ over the infinite sequence of boolean associated with H counts the number of embeddings of G inside H.

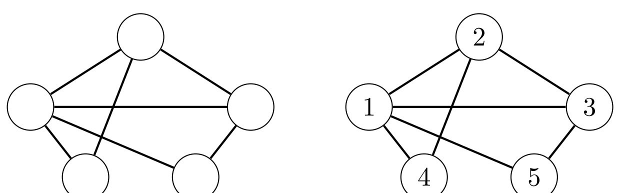
$$\mathcal{M}_G(H) := \mathcal{M}_G((a_{ij})) = \#\{G \hookrightarrow H\} \in \mathbb{N}$$
$$\mathcal{M}_G(G) = 1 \qquad \forall G \text{ graph without isolated vertex}$$

The main motivation justifying the definition of $\mathcal{M}_G$ stand in the following result.

Theorem 2.4. *$\mathcal{M}_G$ can be seen as a maps, defined on every finite graph, counting occurrences of subgraphs. For example, the function $\mathcal{M}_{\triangle}$ counts the number of triangles inside a graph.*

$\mathcal{M}_{\triangle} = x_{12}x_{13}x_{23} + x_{12}x_{14}x_{24} + x_{13}x_{14}x_{34} + x_{23}x_{24}x_{34} + x_{12}x_{15}x_{25} + x_{13}x_{15}x_{35} + x_{23}x_{25}x_{35} + x_{14}x_{15}x_{45} + x_{24}x_{25}x_{45} + x_{34}x_{35}x_{45} + \ldots$

If H is the following graph over 5 nodes, we give it labels with integers from 1 to 5



incidence matrix(H) :

	1	**2**	**3**	**4**	**5**
1	·				
2	1	·			
3	1	1	·		
4	1	1	0	·	
5	1	0	1	0	·

$\mathcal{M}_{\Delta}(H) = 1\cdot1\cdot1+1\cdot1\cdot1+1\cdot1\cdot0+1\cdot1\cdot0+1\cdot1\cdot0+1\cdot1\cdot1+1\cdot0\cdot1+1\cdot 1\cdot0+1\cdot0\cdot0+0\cdot1\cdot0+\dots$

$\mathcal{M}_{\Delta}(H) = 3$ as the only three monomials contributing to H are $x_{12}x_{13}x_{23}$, $x_{12}x_{14}x_{24}$ and $x_{13}x_{15}x_{35}$. This number of triangles is independent of the way chosen for labeling the graph since the functions $\mathcal{M}_G$ are invariant under the relabeling action.

3. The algebra *UGQSym*

We denote by $UGQSym$ the subspace of $\mathbb{K}[x_{ij}|x_{ij}^2 = x_{ij}, i<j]$ generated by the $\mathcal{M}_G$. We will call this space the *Unlabeled Graph Quasi Symmetric functions*.

As we focus on simple graphs (unoriented, simple-edged), the variables, which model edges, can take only two values: 0 if the edge is not present and 1 otherwise. These two numbers are the roots of the polynomial x^2-x and choosing formal variables x_{ij} satisfying $x_{ij}^2 = x_{ij}$ allow us to describe things more combinatorically.

Theorem 3.1. *$UGQSym$ is a subalgebra of $\mathbb{K}[x_{ij}|i<j : x_{ij}^2 = x_{ij}]$. Precisely, there exist non negative integers $c_{G',G''}^{G}$ such that*

$$\mathcal{M}_{G'}\cdot\mathcal{M}_{G''} = \sum_G c_{G',G''}^{G}\mathcal{M}_G.$$

The $c_{G',G''}^{G}$ counts precisely the number of ways of embedding simultaneously G' and G'' onto G.

Here are some examples of products obtained with Sage [4] using tools for computing canonicals under relabeling action [2].

$$\mathcal{M}_{\text{I}}\cdot\mathcal{M}_{\text{I}} = \mathcal{M}_{\text{I}} + 2\mathcal{M}_{\Lambda} + 2\mathcal{M}_{\text{II}}$$

$$\mathcal{M}_{\Lambda}\cdot\mathcal{M}_{\text{I}} = 2\mathcal{M}_{\Lambda} + 3\mathcal{M}_{\Delta} + 3\mathcal{M}_{\text{Y}} + 2\mathcal{M}_{\text{N}} + \mathcal{M}_{\Lambda\text{I}} \tag{1}$$

$$\mathcal{M}_{\text{II}}\cdot\mathcal{M}_{\text{I}} = 2\mathcal{M}_{\text{II}} + \mathcal{M}_{\text{N}} + 2\mathcal{M}_{\Lambda\text{I}} + 3\mathcal{M}_{\text{III}}$$

$$\mathcal{M}_{\Delta}\cdot\mathcal{M}_{\Delta} = \mathcal{M}_{\Delta} + 2\mathcal{M}_{\Delta\!\!\nabla} + 2\mathcal{M}_{\Delta\Delta} + 2\mathcal{M}_{\Delta\nabla}$$

As an exercise, the reader can check that the definition of functions $\mathcal{M}_G$ realizes the simple combinatorial fact: inside a graph, pairs of edges are adjacent or disjoint (use $x_{ij}^2 = x_{ij}$).

$$\binom{\mathcal{M}_{\text{I}}}{2} := \frac{(\mathcal{M}_{\text{I}})(\mathcal{M}_{\text{I}}-1)}{2!} = \mathcal{M}_{\Lambda} + \mathcal{M}_{\text{II}}$$

More generally, it is possible to show that for any non negative integer n, we have

$$\binom{\mathcal{M}_{\text{I}}}{n} := \frac{(\mathcal{M}_{\text{I}})(\mathcal{M}_{\text{I}}-1)\dots(\mathcal{M}_{\text{I}}-(n-1))}{n!} = \sum_{G \text{ has } n \text{ edges}} \mathcal{M}_G$$

Proposition 3.2. *The set of functions $\{\mathcal{M}_G\}$ for a graph G without isolated vertex forms a linear basis of the algebra UGQSym.*

Proof. By definition, the family generates *UGQSym*. Defining leading terms as canonical labeled graphs over packed support, we immediately see that elements of this family are linearly independent. □

The products of functions $\{\mathcal{M}_G\}$ establish algebraic relations between functions counting occurrences of subgraphs. Finally, we can show that only connected patterns are important.

Proposition 3.3. *The functions $\{\mathcal{M}_G\}$ for G connected graph generates the algebra UGQSym.*

Proof. By induction on the number of connected components. Any product $\mathcal{M}_{G_1} \times \mathcal{M}_{G_2}$ with respectively n_1 and n_2 connected components in G_1 and G_2 takes the form

$$\mathcal{M}_{G_1} \times \mathcal{M}_{G_2} = \mathcal{M}_{G_1 \sqcup G_2} + \sum_{\substack{G \text{ has } n < n_1 + n_2 \\ \text{connected components}}} \mathcal{M}_G$$

□

Filtered by number of nodes, the dimensions of "pseudo" homogeneous components is counted by $graph\_node(n)$, the number of graphs over n nodes, Sequence A000088 of the OEIS [11]. We recall the first values $1, 1, 2, 4, 11, 34, 156, 1044, \dots$. These graphs can be seen as monomials over connected graphs, for that, we take connected components as variables and exponents are respectively the multiplicities of isomorphic connected components appearing in the graph. Setting $connected\_node(n)$ as the number of connected graphs over n nodes beginning by $1, 1, 1, 2, 6, 21, 112, 853, \dots$ also sequence A001349 of the OEIS [11]. We have the following relation with power series:

$$\prod_{n>0} \frac{1}{(1-q^n)^{connected\_node(n)}} = \sum_{n \geqslant 0} graph\_node(n) q^n.$$

Therefore, the $\{\mathcal{M}_G\}$ for G connected graphs are algebraically independent.

We can play the same game as before, counting, this time, graphs by number of edges. The number $graph\_edge(n)$ of graphs with n edges (and without isolated vertex) is A000664 of the OEIS and begins with $1, 1, 2, 5, 11, 26, 68, 177, 497$ and the number of connected graphs $connected\_edge(n)$ with n edges is A002905 of the OEIS: $1, 1, 1, 3, 5, 12, 30, 79, 227, 710, \dots$. Although the OEIS does not mention that the first one is the Euler transform of the second, with the same argument as before, we also have:

$$\prod_{n>0} \frac{1}{(1-q^n)^{connected\_edge(n)}} = \sum_{n \geqslant 0} graph\_edge(n) q^n.$$

We can realize a coproduct by means of *doubling alphabet trick* of [5]. Noting $\mathbb{A}$ our infinite triangle of variables x_{ij} with $0 < i < j$, we introduce another copy of this alphabet, $\mathbb{B}$ composed of variables $x_{i'j'}$ with i' and j' satisfying the same constraints as i and j. We define $\mathbb{A} \oplus \mathbb{B}$ like $\mathbb{A} \sqcup \mathbb{B}$ since we want to forbid variables of the form $x_{ij'}$ or $x_{j'i}$.

$$\begin{aligned}
\mathcal{M}_G(\mathbb{A}\oplus\mathbb{B}) &:= \sum_{\substack{g:\text{labeling of } G\\ \text{using normal and prime letters}}} m(g)\\
&= \sum_{\substack{g_1,g_2:\text{labeling of } G_1\sqcup G_2=G\\ G_1 \text{ labeled with normal letters, } G_2 \text{ labeled with prime letters}}} m(g_1\sqcup g_2)\\
&= \sum_{G_1\sqcup G_2=G}\Big(\sum_{\substack{g_1:\text{labeling of } G_1\\ G_1 \text{ labeled with normal letters}}} m(g_1)\Big)\cdot\Big(\sum_{\substack{g_2:\text{labeling of } G_2\\ G_2 \text{ labeled with prime letters}}} m(g_2)\Big)\\
\mathcal{M}_G(\mathbb{A}\oplus\mathbb{B}) &= \sum_{G_1\sqcup G_2=G}\mathcal{M}_{G_1}(\mathbb{A})\otimes\mathbb{M}_{G_2}(\mathbb{B})
\end{aligned}$$

Definition 3.4. *Let G be a finite graph without isolated vertex. We define a coproduct Δ on basis elements $\mathcal{M}_G$ of UGQSym as follows:*

$$\Delta(\mathcal{M}_G) := \sum_{G'\sqcup G''=G} M_{G'}\otimes M_{G''}$$

The sum runs over ordered pairs (G', G'') of graphs without isolated vertex such that G becomes the disjoint union of G' and G''. Recall that the graph reduced to a single vertex (without any edge) has no isolated vertex.

We directly see that Δ has for possible coefficients 0 and 1, that it is coassociative and cocomutative, and that its primitive elements are connected graphs. Here are some examples:

$\Delta(\mathcal{M}$  $) = \mathcal{M}$  $\otimes 1 + 1\otimes\mathcal{M}$ 

$\Delta(\mathcal{M}$  $) = \mathcal{M}$  $\otimes 1 + \mathcal{M}$  $\otimes\mathcal{M}$  $+ 1\otimes\mathcal{M}$ 

$\Delta(\mathcal{M}$  $) = \mathcal{M}$  $\otimes 1 + \mathcal{M}$  $\otimes\mathcal{M}$  $+\mathcal{M}$  $\otimes +\mathcal{M}$  $+ 1\otimes\mathcal{M}$ 

Proposition 3.5. *$(UGQSym, \cdot, \Delta)$ is a Hopf algebra. Moreover, it is filtered by the number of edges or by the number of nodes.*

Proof. The product and coproduct are compatible by induction on the number of connected components of the operands. For G_1 and G_2 two non isomorphic connected graphs, the product $\mathcal{M}_{G_1}\cdot\mathcal{M}_{G_2}$ takes the form

$$\mathcal{M}_{G_1}\cdot\mathcal{M}_{G_2} = \mathcal{M}_{G_1\sqcup G_2} + \sum c^H_{G_1,G_2}\mathcal{M}_H, \tag{2}$$

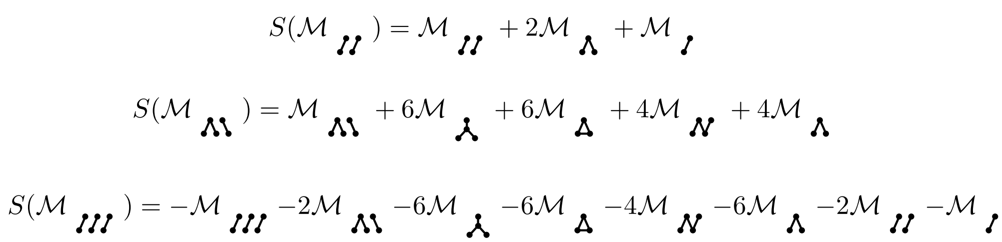
where the sum runs over some connected graphs H (note that if G_1 and G_2 are isomorphic, the coefficient of $\mathcal{M}_{G_1\sqcup G_1}$ is 2 instead of 1). From this, as the coproduct stays simple for connected graphs, we easily check that $\Delta(f\cdot g) = \Delta(f)(\cdot\otimes\cdot)\Delta(g)$.

Generally, Formula (2) shows that any product $\mathcal{M}_{G_1}\cdot\mathcal{M}_{G_2}$ contains a leading term whose associated graph has as number of nodes the sum of the number of nodes of the operands, as number of edges the sum of the number of edges of the operands and as number of connected components the sum of the number of connected components of the operands. Other terms of the product are associated with graphs which are non trivial merges of G_1 and G_2 and for which at least one edge has been joined (and thus at least two vertices from G_1 and G_2 has been merged). □

Being commutative, cocommutative and filtered, the antipode S is uniquely determined by $S(1) = 1$ together with the induction formula:

$$S(\mathcal{M}_G) = -\sum_{\substack{(G',G'')\in\Delta(\mathcal{M}_G)\\ G''\neq G}} \mathcal{M}_{G'} \cdot S(\mathcal{M}_{G''})$$

Moreover S satisfies $S^2 = Id$. We have naturally $S(\mathcal{M}_G) = -\mathcal{M}_G$ for any G connected graph. Here are some non trivial examples.



4. Connection with algebraic invariant theory

The definition of the $\mathcal{M}_G$ obviously shows that the algebra $UGQSym$ is composed of invariant functions under the relabeling action of the vertices. We now show that $UGQSym$ contains a sufficient number of invariants to separate finite graphs.

The next experiment consists in keeping only variables indexed by small indices. For a fixed integer n, in our triangle of variables, we will send to 0 all variables x_{ij} for $i > n$ (we thus keep only an upper triangular part). Applying this restriction on the $\mathcal{M}_G$, we define polynomials $\mathcal{P}_{n,G}$ which are sums of monomials whose support contains only $\binom{n}{2}$ different variables and inside which variables can appear with degree 1. These polynomials are the central objects of [7].

In the sequel, we insist on the fact that $\mathbb{K}$ is of characteristic 0 and that we do not keep relations $x_{ij}^2 = x_{ij}$ on remaining variables.

Let n be a positive integer. The group acting on edges of graphs over n nodes is a permutation group (isomorphic to the symmetric group $\mathfrak{S}_n$), a subgroup of the symmetric group $\mathfrak{S}_{\binom{n}{2}}$. The following theorem exploits the combinatorial structure of rings of invariants under the action of a permutation group.

Theorem 4.1. *Let n be a positive integer and G be the permutation group acting on bi-indexed variables $\mathbf{x} := (x_{12}, x_{13}, \ldots, x_{n-1\ n})$ as the relabeling action on graphs over n nodes. There exists a finite family of $\frac{\binom{n}{2}!}{n!}$ polynomials η_{λ_i} which are linear combinations of higher Specht polynomials such that the invariant ring $\mathbb{K}[\mathbf{x}]^G$ under the action of G can be decomposed as*

$$\mathbb{K}[\mathbf{x}]^G = \bigoplus_{\lambda\vdash\binom{n}{2}} \bigoplus_i \eta_{\lambda_i}\mathbb{K}[e_1, e_2, \ldots e_{\binom{n}{2}}], \tag{3}$$

where the e_k are the elementary symmetric polynomials in $\binom{n}{2}$ variables and each η_{λ_i} is a linear combination of some higher Specht polynomials F_T^S with S ant T standard Young tableaux of shape the partition λ.

Proof. Details are available in [3]. The algorithm presented in this paper shows how to compute the secondary invariants η_{λ_i} inside irreducible representations of the ambient symmetric group of degree $\binom{n}{2}$. The slicing in Equation (3) is finer than the one presented in classical Hironaka decomposition because the first direct sum runs over partitions and not over possible degrees. As we can have several G-stable subspaces at a given degree, secondary invariants are better partitioned here. □

Remark 4.2. *The family formed by* $\{e_1, e_2, \dots e_{\binom{n}{2}}\} \cup \{\eta_{\lambda_i}\}_{\lambda \vdash \binom{n}{2}, i}$ *can separate all pairs of orbits of* G *(relabeling action) when acting on* $\mathbb{C}^{\binom{n}{2}}$ *(vectors of* $\binom{n}{2}$ *complex numbers).*

The family generates the whole ring of invariants under the action of G and there always exists an invariant separating two different orbits. Here, this huge family separates non oriented multi-graphs with a complex number labeling each edge (also symmetric complex matrices of size n with zeros on the diagonal).

Now, we go back to simple graphs in which, for each pair of nodes, either we have an edge (when the label of the edge is 1) or either we do not have the link (the label is then 0). Our goal is now to reduce the number of elements of the huge separating family since simple graphs are vectors of $\binom{n}{2}$ booleans.

Proposition 4.3. *For any partition* $\lambda \vdash n$ *having at least three parts and any pair* (S, T) *of standard Young tableaux of shape* λ*, the evaluation of higher Specht polynomials* F_T^S *is zero on vectors of the type* $\{0, 1\}^n$*.*

Proof. Higher Specht polynomials are built with Young symmetrizers which introduce some anti-symmetries. If the tableau has at least three boxes in the first column, the associated higher Specht will be at least divisible by a Vandermonde type factor over at least three variables. But a Vandermonde factor in k variables needs at least k different values to be non zero ($(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)$ is non zero if x_1, x_2 and x_3 are pairwise different). □

Remark 4.4. *Let* v *be a vector of booleans of length* n *and* $i(v)$ *the number of* 1 *in* v*. Then the evaluations of the elementary symmetric polynomials* e_k *on* v *depend only on* $e_1(v) = i(v)$*. Precisely, we have:*

$$\forall\ 1 \leqslant k \leqslant n, \qquad e_k(v) = \binom{i(v)}{k} = \binom{e_1(v)}{k}.$$

We deduce from Proposition 4.3 and Remark 4.4 that our separating family can be largely reduced for simple graphs. We can keep only the first elementary symmetric polynomial (which counts the number of edges) and combinations of higher Specht polynomials for the relabeling action associated with partitions having two parts.

Moreover, as we just want to separate orbits (and not generate the whole ring of invariants), we can use the fact that the action of a permutation σ over an higher Specht polynomial F_T^S depends only on the irreducible representation (tableau T) and not on the degree scaling (tableau S for the multiplicity), on such polynomials, we have $\sigma \cdot F_T^S = F_{\sigma \cdot T}^S$ [1]. Therefore, Specht polynomials are sufficient and higher Spechts bring only more copies of isomorphic G-stable subspaces which contribute to generate the ring of invariants but do not contribute to separate orbits. More details are available in [1, 3].

Finally, noticing that a tableau of size $\binom{n}{2}$ composed of two rows has at most $\lfloor \frac{\binom{n}{2}}{2} \rfloor$ boxes at height 2, we have:

Theorem 4.5. *Let H_1 and H_2 two graphs over n nodes. The following statements are equivalents:*

(i) H_1 and H_2 are isomorphic.

(ii) For all graphs G over at most n nodes and having at most $\lfloor \frac{\binom{n}{2}}{2} \rfloor$ edges:

$$\mathcal{P}_{n,G}(H_1) = \mathcal{P}_{n,G}(H_2)$$

(iii) For all graphs G over at most n nodes and having at most $\lfloor \frac{\binom{n}{2}}{2} \rfloor$ edges:

$$\mathcal{M}_G(H_1) = \mathcal{M}_G(H_2)$$

Proof. We take all monomials associated with tableaux with two rows. They correspond to graphs over n nodes having as edges the labels in the upper row. Applying the orbit sum operator (Reynold's operator up to a scalar) over these monomials, we obtain the required polynomials $\mathcal{P}_{n,G}$, which are more than needed to separate all orbits of booleans vectors. □

Using the multiplicative structure of the algebra $UGQSym$, we can formulate:

Corollary 4.6. *Let H_1 and H_2 be two graphs over n nodes. H_1 and H_2 are isomorphic if and only if $\mathcal{M}_G(H_1) = \mathcal{M}_G(H_2)$ for all connected graph G over n nodes having at most $\lfloor \frac{\binom{n}{2}}{2} \rfloor$ edges.*

5. Application to graph reconstruction

The reconstruction of finite graphs is an old problem [6, 12]. Given a finite graph H over $n \geqslant 3$ nodes, we can build the multi-set of vertex deleted subgraphs of H: $\{H_1, H_2, \ldots, H_n\}$ which is formed by all induced subgraphs of H by deleting exactly one vertex. This process forms a map. The reconstruction conjecture investigates if this map is injective, therefore finite graphs may be determined by their subgraphs and finite graphs would be reconstructibles.

Proposition 5.1 (Kelly's lemma)**.** *Let n a positive integer, H a graph over n nodes and G a graph over $r < n$ nodes. Let $\{H_1, H_2, \ldots, H_n\}$ the (possibly multi-)set of graphs obtained from H when deleting a single vertex. Thus, the number of ways of embedding G in H can be deduced from the number of embedding G in each H_i:*

$$\mathcal{P}_{n,G}(H) = (n-r) \sum_{i=1}^{n} \mathcal{P}_{n,G}(H_i)$$

Proposition 5.2. *Finite graphs over n nodes are reconstructible if the values of $\mathcal{P}_{n,G}$ for G finite connected graph over n nodes with at most $\lfloor \frac{\binom{n}{2}}{2} \rfloor$ edges can be deduced from the values of $\mathcal{P}_{n,H}$ for H finite graph over at most $n-1$ nodes.*

Proof. Use Corollary 4.6 and Kelly's lemma 5.1. □

Case $n = 3$**:** Graphs over 3 nodes are isomorphic if and only if evaluations coincide on the single series $\mathcal{M}_{/}$ since $\lfloor \frac{\binom{3}{2}}{2} \rfloor = \lfloor \frac{3}{2} \rfloor = 1$. Notice that the following relations:

$$\binom{\mathcal{M}_{/}}{2} = \mathcal{M}_{\Lambda} + \mathcal{M}_{//}$$

$$\binom{\mathcal{M}_{/}}{3} = \mathcal{M}_{\Delta} + \mathcal{M}_{\curlywedge} + \mathcal{M}_{N} + \mathcal{M}_{\Lambda\backslash} + \mathcal{M}_{///} \tag{4}$$

become for the finite case for three nodes:

$$\binom{\mathcal{P}_{/}}{2} = \mathcal{P}_{\Lambda} \qquad\qquad \binom{\mathcal{P}_{/}}{3} = \mathcal{P}_{\Delta}$$

Since other patterns have as support more than 3 nodes, they vanish in this finite case.

Case $n = 4$**:** As $\lfloor \frac{\binom{4}{2}}{2} \rfloor = 3$, graphs over 4 nodes are reconstructibles if $\mathcal{P}_{\curlywedge}$ and $\mathcal{P}_{N}$ can be deduced from counting function of graphs over at most 3 nodes. These two functions correspond to the only two connected graphs over 4 nodes having at most 3 edges.

Equation (4) and the product (1) give for graphs over four nodes the identities:

$$\begin{cases} \binom{\mathcal{P}_{/}}{3} - \mathcal{P}_{\Delta} & = \quad \mathcal{P}_{\curlywedge} + \mathcal{P}_{N} \\ \mathcal{P}_{\Lambda} \cdot (\mathcal{P}_{/} - 2) - 3\,\mathcal{P}_{\Delta} & = \quad 3\,\mathcal{P}_{\curlywedge} + 2\,\mathcal{P}_{N} \end{cases}$$

Conjecture 5.3. *Let $n \geqslant 3$ and G a connected graph over n nodes having at most $\lfloor \frac{\binom{n}{2}}{2} \rfloor$ edges, then $\mathcal{P}_{n,G}$ is a polynomial over some $\mathcal{P}_{n,H}$ where H are graphs over at most $n-1$ nodes.*

If this conjecture is true, it would imply Ulam's conjecture. Note also that this conjecture will not give information about the complexity of graph isomorphism problem.

Here is the matrix $\mathcal{M}_G(H)$ for (G, H) the 23 smallest graphs without isolated nodes.

	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	2	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	3	3	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	2	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	3	3	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	3	2	.	1	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	4	5	1	1	1	2	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	4	4	.	2	.	4	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	5	8	2	2	2	6	4	1	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	6	12	4	3	4	12	12	3	6	1	.	.	.	.	.	.	.	.	.	.	.	.	.
	3	1	.	2	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.
	4	3	1	3	.	.	.	.	.	.	3	1	.	.	.	.	.	.	.	.	.	.	.
	4	6	.	.	4	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.
	4	4	.	2	1	2	.	.	.	.	1	.	.	1	.	.	.	.	.	.	.	.	.
	4	3	.	3	.	2	.	.	.	.	2	.	.	.	1	.	.	.	.	.	.	.	.
	5	8	1	2	4	4	2	.	.	.	1	.	1	2	.	1	.	.	.	.	.	.	.
	5	7	1	3	2	5	2	.	.	.	2	.	.	2	1	.	1	.	.	.	.	.	.
	5	6	1	4	1	4	1	.	.	.	4	1	.	1	2	.	.	1	.	.	.	.	.
	5	6	.	4	1	6	.	1	.	.	3	.	.	2	2	.	.	.	1	.	.	.	.
	5	5	.	5	.	5	.	.	.	.	5	.	.	.	5	.	.	.	.	1	.	.	.
	6	11	2	4	5	10	6	1	1	.	3	.	1	5	2	2	2	.	1	.	1	.	.
	6	**10**	**2**	**5**	3	10	5	1	1	.	5	1	.	4	4	.	2	2	1	.	.	1	.
	6	**10**	**2**	**5**	4	8	4	.	.	.	6	2	1	4	4	2	.	4	.	.	.	.	1

Even if completed with the remaining graphs over five nodes, the last two graphs are the algebraically closest pair of graphs over at most five nodes.

Acknowledgements

This research was driven by computer exploration using the open-source mathematical software `Sage` [4] and its algebraic combinatorics features developed by the `Sage-Combinat` community [10], and in particular the support for Hopf algebras developed by Nicolas M. Thiéry et al. and the enumeration tools modulo the action of a permutation group developed by author.

References

[1] Susumu Ariki, Tomohide Terasoma, and Hiro-Fumi Yamada. Higher Specht polynomials. *Hiroshima Math. J.*, 27(1):177–188, 1997.

[2] Nicolas Borie. Generating tuples of integers modulo the action of a permutation group and applications. In *25th International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2013)*, Discrete Math. Theor. Comput. Sci. Proc., AS, pages 767–778. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2013.

[3] Nicolas Borie. Effective invariant theory of permutation groups using representation theory. *Algebraic Informatics, L.N.C.S.*, pages 58–69, 2015.

[4] The Sage Developers. *Sage Mathematics Software (Version 6.4)*, 2014. `http://www.sagemath.org`.

[5] Gérard Duchamp, Florent Hivert, and Jean-Yves Thibon. Noncommutative symmetric functions. VI. Free quasi-symmetric functions and related algebras. *Internat. J. Algebra Comput.*, 12(5):671–717, 2002.

[6] Paul J. Kelly. A congruence theorem for trees. *Pacific J. Math.*, 7:961–968, 1957.

[7] Tomi Mikkonen and Xavier Buchwalder. The algebra of graph invariants - lower and upper bounds for minimal generators. *arXiv:0712.0142*, 2008.

[8] Jean-Christophe Novelli, Jean-Yves Thibon, and Nicolas M. Thiéry. Algèbres de Hopf de graphes. *C. R. Math. Acad. Sci. Paris*, 339(9):607–610, 2004.

[9] Maurice Pouzet and Nicolas M. Thiéry. Invariants algébriques de graphes et reconstruction. *C. R. Acad. Sci. Paris Sér. I Math.*, 333(9):821–826, 2001.

[10] The Sage-Combinat community. Sage-Combinat: enhancing Sage as a toolbox for computer exploration in algebraic combinatorics, 2008. `http://combinat.sagemath.org`.

[11] Neil J. A. Sloane. The on-line encyclopedia of integer sequences. *Notices Amer. Math. Soc.*, 50(8):912–915, 2003.

[12] Stanisław M. Ulam. *A collection of mathematical problems.* Interscience Tracts in Pure and Applied Mathematics, no. 8. Interscience Publishers, New York-London, 1960.

Univ. Paris Est Marne-La-Vallée, Laboratoire d'Informatique Gaspard Monge, 5, bd Descartes Champs sur Marne 77454 Marne-la-Vallée Cedex 2, France