The Isomorphism Problem for Catalan Families

Norbert Hungerbühler[1]

To the memory of Prof. Max Jeger

Abstract: Given two sets of configurations S and T (defined by some geometric or algebraic rules) with weightfunctions $\omega_S : S \to \mathbb{N}$ and $\omega_T : T \to \mathbb{N}$ respectively. Suppose S and T are equipotent, i.e. $|\omega_S^{-1}(n)| = |\omega_T^{-1}(n)|$ for all $n \in \mathbb{N}$, then we may ask for a weight-preserving "canonical isomorphism" between S and T. We consider different Catalan families as a model in order to study this question: we investigate how a concrete isomorphism at the combinatorial level may be constructed and how it generates an isomorphism at the level of species.

AMS subject classification codes: 05A15, 05C05, 06B05

1 Introduction

Many examples of Catalan families are known and one can find them in almost every classical work on combinatorics (see e.g. [1], [4], [11] or [12]). But here we are interested in a closed description of the *relations between* those families. It is the aim of this article to clear up these relations in a systematic way and to interpret the result in the language of species (see Theorem 2, Section 4 and Section 6). The reader who is familiar with the examples of the Catalan families may skip the first one or two sections.

The sequence $(C_n)_{n\in\mathbb{N}}$ of the *Catalan numbers* is defined by

$$C_n := \frac{1}{n+1}\binom{2n}{n}. \tag{1}$$

We will see later that the recursion formula

$$C_{n+1} = \sum_{k=0}^{n} C_k\, C_{n-k} \tag{2}$$

holds and that the generating series $f(x) := \sum_{n=0}^{\infty} C_n x^n$ formally satisfies the relation

$$x\, f(x)^2 - f(x) + 1 = 0\,. \tag{3}$$

Let M be a countable set of configurations with weightfunction $\omega : M \to \mathbb{N}$ and generating series

$$f(x) = \sum_{\phi\in M} x^{\omega(\phi)} = \sum_{n=0}^{\infty} |M_n| x^n$$

(where M_n is the set of all figures of M having weight n), then we call M a *Catalan family* provided $|M_n| = C_n$ holds.

Now we give some examples of Catalan families (of course, the list is far from being complete).

1.1 Euler triangulation: The first example of a Catalan family was given by Leonhard Euler. In 1751 he described the following problem in a letter to Goldbach:

A convex polygon of $n+2$ edges may be triangulated by $n-1$ nonintersecting diagonals. Let M_n be the set of those triangulations. Euler proved the formulas (1)–(3) and wrote to Goldbach: "Die Induktion, so ich gebraucht, war ziemlich mühsam..."

[1]Norbert Hungerbühler, Mathematik Departement, ETH Zürich, CH-8092 Zürich, Switzerland

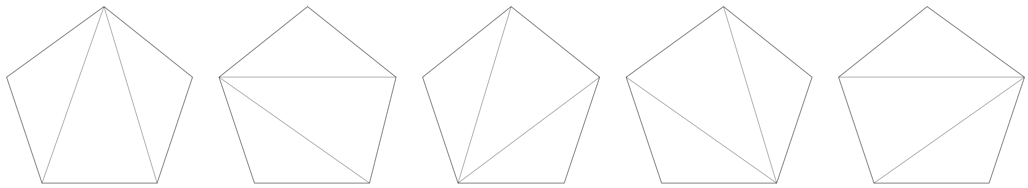

Figure 1: Euler triangulation M_3

1.2 Bracket figures: In 1838 Charles Catalan posed the following problem:

In how many ways one can bracket a term of $n+1$ variables? Let us denote those bracketings by M_n. Then M_3 looks like that:

$$M_3 = \Big\{ \big(x*(x*(x*x))\big), ((x*x)*(x*x)), \big(((x*x)*x)*x\big), \big((x*(x*x))*x\big), \big(x*((x*x)*x)\big) \Big\}$$

1.3 Free bracket figures: A variation of the previous problem is the following: How many syntactic bracketings one can build with n pairs of brackets? This is also a Catalan family and M_3 is

$$M_3 = \Big\{ ()()(), (())(), ()(()), (()()), ((())) \Big\}.$$

1.4 Trivalent plane rooted trees: A *trivalent plane rooted tree* is a circlefree connected graph which is embedded in the plane and whose vertices have degree 1 or 3. One of the vertices of degree 1 is marked as root r. Let M_n be the set of trivalent plane rooted trees with $n+1$ leaves. In Figure 2 the set M_3 is shown.

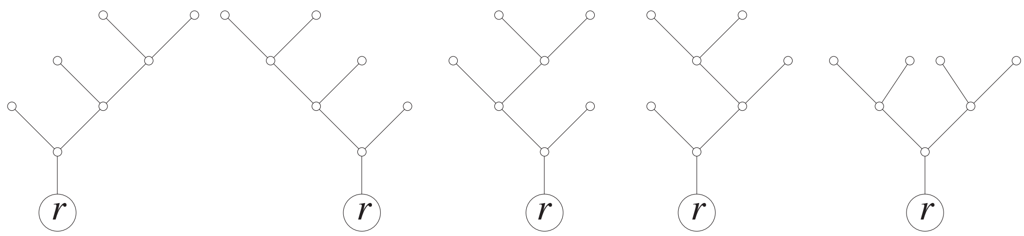

Figure 2: Trivalent plane rooted trees M_3

1.5 Catalan trees: Let M_n be the set of *planted* trees (i.e. trees with a root of degree 1) in the plane with $n+1$ edges. This is a Catalan family. Figure 3 shows M_3:

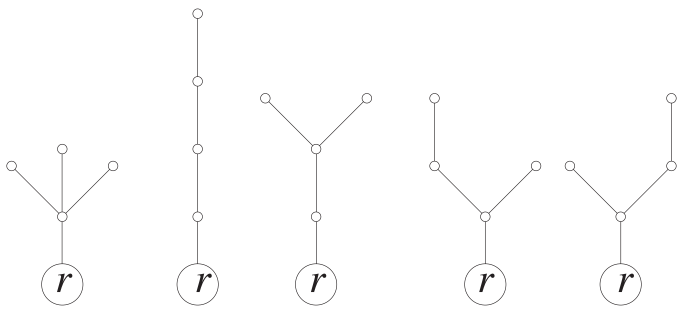

Figure 3: Plane planted trees M_3

1.6 Lattice paths: Consider the set of all paths in a lattice connecting $(0,0)$ and (n,n) such that the coordinate functions are nondecreasing and no point (x,y) with $x<y$ is visited. Let us denote this set by M_n. See Figure 4.

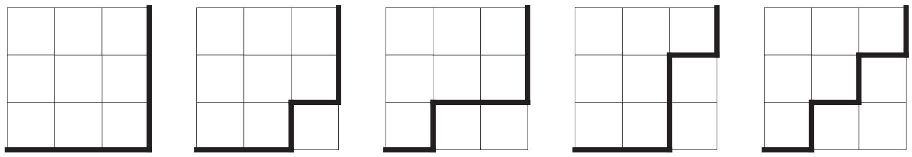

Figure 4: Lattice paths M_3

2 Definitions

Before we show that the configurations presented in the previous section are Catalan families and how they are related to each other, we give here the necessary definitions to make a structural approach to the problem (as a general reference for this concept see e.g. [3] and [2]).

Definition 1 *Let* $(M_n)_{n\in\mathbb{N}}$ *be a sequence of sets such that:*

a) $|M_0|=1$
b) For every $n\in\mathbb{N}$ *there is a bijective map*

$$f_{n+1}: M_{n+1} \mapsto \bigcup_{i=0}^{n} M_i \times M_{n-i}\,,$$

where $\cup$ *denotes the disjoint union. Let* $M := \cup_{n\in\mathbb{N}} M_n$ *and* $M^* := M\setminus M_0$. *We call the decomposition* $f: M^* \mapsto M\times M$ *with* $f|M_n = f_n$ Catalan mapping *and the pair* (M,f) special Catalan family.

A map $h: M\to N$ *between two special Catalan families* (M,f) *and* (N,g) *is called* Catalan isomorphism *if* $h(M_0)=N_0$ *and if the following diagram commutes:*

$$\begin{array}{ccc} M^* & \xrightarrow{h} & N^* \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle g} \\ M\times M & \xrightarrow{h\times h} & N\times N \end{array}$$

Remark 1: Of course, every *special* Catalan family in fact *is* a Catalan family: to verify that $|M_n| = C_n$ observe that a Catalan mapping induces a bijection: $M\to M_1\times M\times M\cup M_0$. Since the weight n is additive we may apply Polya's theorem (see [5]) to get equation (3) in Section 1 for the generating function $f(x)$. This is a quadratic equation in the ring of formal power series $\mathbb{R}[\![x]\!]$ with solution

$$f(x) = \frac{1}{2x}\left(1 \pm (1-4x)^{1/2}\right)\,.$$

Expansion in a binomial series leads to

$$f(x) = \frac{1}{2x}\left(1 \pm \sum_{n\in\mathbb{N}} \binom{1/2}{n}(-4x)^n\right) = \frac{1}{2}\sum_{n\in\mathbb{N}}\binom{1/2}{n+1}(-1)^n 4^{n+1}x^n = \sum_{n\in\mathbb{N}} C_n x^n\,.$$

Note that only the positive sign respects $|M_0| = 1$.

Obviously we have $|M_{n+1}| = \sum_{i=0}^{n} |M_i|\,|M_{n-i}|$, and hence also the recursion formula (2) is established. □

Theorem 2 *Let (M, f) and (N, g) be special Catalan families. Then there exists a unique Catalan isomorphism $h : M \to N$.*

Proof: We define $h_k : M_k \to N_k$ recursively.

$$\begin{aligned} h_0 : m_0 &\mapsto n_0 \quad \text{where } M_0 = \{m_0\} \text{ and } N_0 = \{n_0\} \\ h_k : m &\mapsto n \quad \text{where } f_k(m) = (x, y) \in M_i \times M_{k-i-1},\ n = g_k^{-1}(h_i(x), h_{k-i-1}(y)). \end{aligned}$$

Then $h : M \to N$ with $h|M_k = h_k$ is a Catalan isomorphism. The proof of the uniqueness is similar. □

Remark 2:

(i) Up to isomorphism there exists exactly one special Catalan family.

(ii) If we find on an arbitrarily defined family $\{K_n\}_{n\in\mathbb{N}}$ of sets a Catalan mapping, then it is a Catalan family by Remark 1.

(iii) If we find a Catalan mapping on a Catalan family, the proof of Theorem 2 gives us an isomorphism to any other special Catalan family.

(iv) If we find a bijective mapping $h : M \to N$ of a special Catalan family (M, f) to an arbitrary family N, then $g := (h \times h) \circ f \circ h^{-1}$ is a Catalan mapping on N which makes N to a special Catalan family and the diagram in Definition 1 commutes.

3 Examples of Catalan mappings

We apply Remark 2(ii) to show that Examples 1.1 to 1.6 are in fact Catalan families. In each Example we have to show the existence of a Catalan mapping.

3.1 Euler triangulation: Let $m \in M_n$ be a triangulated convex polygon of $n + 2$ edges. We may think of one side to be marked. This side is the side of a triangle which decomposes the polygon into two convex polygons m_1 on the left and m_2 on the right ("left" and "right" with respect to a given orientation). We choose the two sides touching the triangle as the marked sides of the two new polygons (see Figure 5). It is easy to verify that this decomposition is a Catalan mapping.

3.2 Bracket figures: Let $m \in M_n$ be a bracket figure. Then there exist bracket figures m_1 and m_2 with $m = (m_1 * m_2)$ and $m \mapsto (m_1, m_2)$ defines a Catalan Mapping.

Example: $m = \Big(x * \big(x * (x * x)\big)\Big) \;\mapsto\; m_1 = x,\ m_2 = \big(x * (x * x)\big)$.

3.3 Free bracket figures: Let $m \in M_n$ be a free bracket figure. Then there exist free bracket figures m_1 and m_2 with $m = (m_1)m_2$. It is easy to see, that $m \mapsto (m_1, m_2)$ defines a Catalan mapping.

Example: $m = \mathbf{(}(()(()()))\mathbf{)}(()()) \;\mapsto\; m_1 = (()(()())),\ m_2 = (()())$.

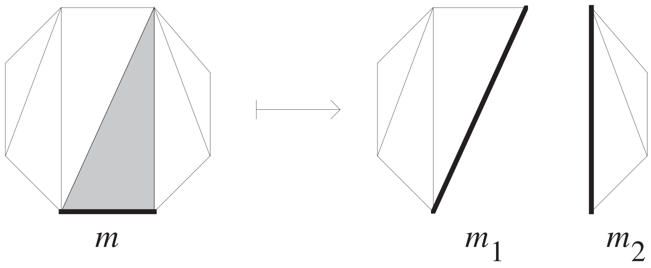

Figure 5: Decomposition of an Euler triangulation

3.4 Trivalent plane rooted trees: Let m be a trivalent plane tree with root r. We take the vertex s next to r as the root of the two trivalent plane subtrees with root s—the left one as m_1, the right one as m_2 (right means here the first branch that we meet walking anticlockwise round s starting in r). This decomposition is a Catalan mapping.

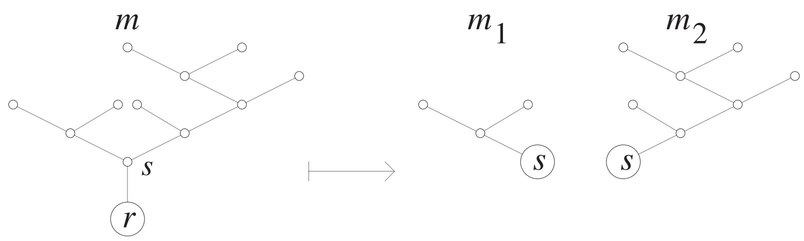

Figure 6: Decomposition of a trivalent plane rooted tree

3.5 Catalan trees: Let m be a Catalan tree with root r and s the vertex next to r. We take the most left branch starting in s as subtree m_1 with root s. If we denote the rest of m by m_2, we get a Catalan mapping $m \mapsto (m_1, m_2)$. See Figure 7.

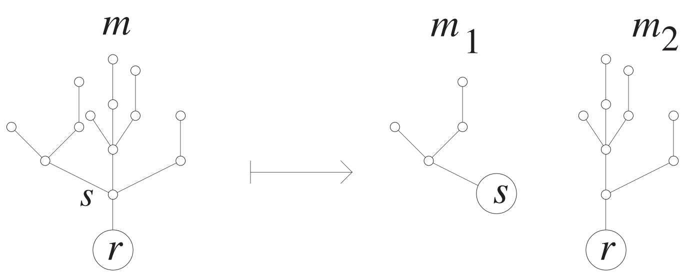

Figure 7: Decomposition of a Catalan tree

3.6 Lattice paths: Let m be a path as described in Section 1.6. Starting in the origin we go the first point (p, p) of the path lying on the diagonal. We denote the path from $(1, 0)$ to $(p, p-1)$ by m_1 and the path from (p, p) to (n, n) by m_2. This decomposition defines a Catalan mapping. See Figure 8.

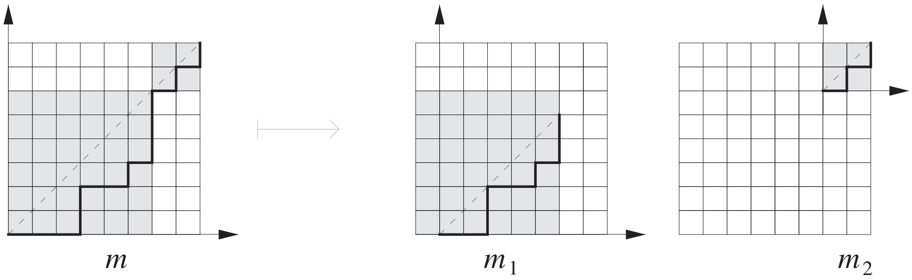

Figure 8: Decomposition of a lattice path

4 Catalan Isomorphisms

Now we regard the examples given above as special Catalan families. By Theorem 2 we can find isomorphisms between them by considering corresponding decompositions. It will turn out that all these isomorphisms, although they are given recursively, admit a nonrecursive interpretation. Variants of most of the following isomorphisms (but not their derivation) may be found in the previously mentioned literature.

4.1 Free bracket figures versus plane Catalan trees: Let $m = (a)b$, a and b be free bracket figures. The Catalan decomposition is $m = (a)b \mapsto m_1 = a$, $m_2 = b$. The corresponding Catalan tree has according to Theorem 2 the corresponding decomposition (see Figure 9). The subtrees A with root r and B with root s correspond to a and b

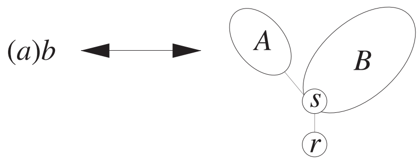

Figure 9: Corresponding decomposition of a free bracket figure and a Catalan tree

respectively. Hence we may describe the isomorphism like this: Every edge corresponds to a pair of brackets which contains the bracket figure that corresponds to the part of the tree which continues the edge. Hence, we can “read” the bracket figure as the thin line in Figure 10 indicates:

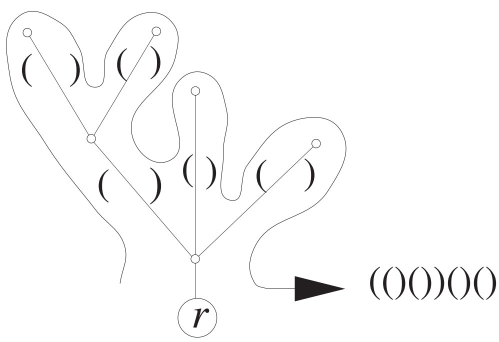

Figure 10: Isomorphism between free bracket figures and Catalan trees

4.2 Trivalent plane rooted trees versus Catalan trees: Let t be a trivalent rooted tree. The corresponding Catalan tree T has the corresponding decomposition: See Figure 11. The subtrees a with root v and b with root v correspond to the subtrees A with

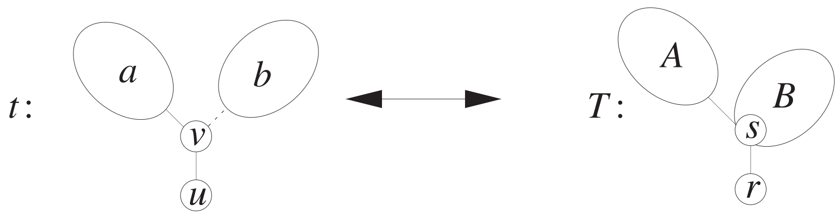

Figure 11: Corresponding decomposition of a trivalent plane tree and a Catalan tree

root s and B with root r respectively. Now it is easy to see how we get from t to T: We contract the dashed edge in t and continue in the same way in the subtrees. Thus we may describe the isomorphism like this: First we bring the trivalent rooted plane tree in a special position: all edges run from left to right or from bottom to top. Then we *contract* the horizontal edges (See Figure 12).

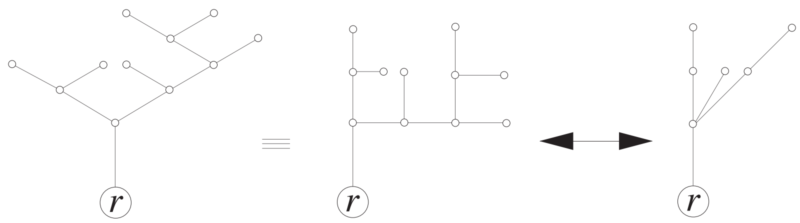

Figure 12: Isomorphism between trivalent plane trees and Catalan trees

Now we leave it to the reader to carry out explicitly the construction of the following isomorphisms and only give some of the results:

4.3 Lattice paths versus bracket figures: Consider a lattice path as in Section 1.6. For every horizontal step write "(" and for every vertical step write "$*$". Then the missing ")" and the variables are easily completed.

4.4 Bracket figures versus free bracket figures: Consider a bracket figure as in Section 1.2. Remove all variables and write ")" for every "$*$" to get the corresponding free bracket figure.

4.5 Bracket figures versus Euler triangulation: Figure 13 should make clear this isomorphism.

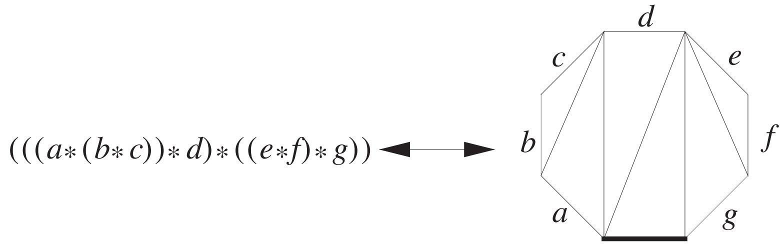

Figure 13: Isomorphism between bracket figures and Euler triangulation

4.6 Trivalent trees versus Euler triangulation: This isomorphism is also used in order to formulate the problem of coloring a map in the language of graph theory. See Figure 14.

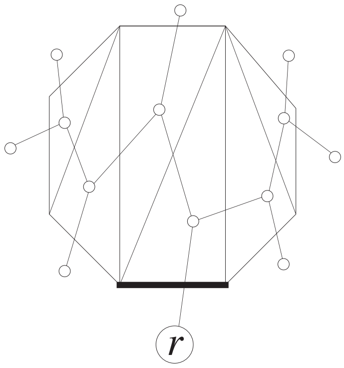

Figure 14: Isomorphism between trivalent plane rooted trees and Euler triangulation

Remark 3: This last isomorphism may be used to count the number of trivalent plane trees (no more rooted) with $n+2$ leaves: We have to count the Euler triangulations of a regular polygon modulo the dihedral group. Burnside's Lemma leads to the following formula for n even

$$|K_n| = \frac{1}{2n+4}\left(C_n + \frac{3n+6}{2}C_{\frac{n}{2}} + \frac{2n+4}{3}C_{\frac{n-1}{3}}\right)$$

and for n odd

$$|K_n| = \frac{1}{2n+4}\left(C_n + (n+2)C_{\frac{n}{2}} + \frac{2n+4}{3}C_{\frac{n-1}{3}}\right).$$

where we used the convention $C_n := 0$ for $n \notin \mathbb{N}$.

Remark 4: The same methods as for Catalan families apply to find isomorphisms for other combinatorial families, e.g. the generalized Catalan families described in [4].

5 Algorithmic remarks

In this section we adapt some ideas of [7] and [10] to define a ranking on special Catalan families and show the connection to the isomorphism problem.

Theorem 3 *For every special Catalan family (M, f) there exists a unique total order on M (usually called a ranking) compatible to f in the following sense: For $x \in M_u$, $y \in M_v$ with $f(x) = (x_1, x_2) \in M_i \times M_j$ and $f(y) = (y_1, y_2) \in M_k \times M_m$ there holds*

$$x < y \iff \begin{cases} u < v \text{ or} \\ u = v \text{ and } x_1 < y_1 \text{ or} \\ u = v \text{ and } x_1 = y_1 \text{ and } x_2 < y_2 \end{cases}$$

Proof: On M_0 we have the trivial total order relation. Let $<$ be defined on M_i, $0 \leq i \leq n$, and $x, y \in M_{n+1}$, $f(x) = (x_1, x_2)$, $f(y) = (y_1, y_2)$. Then

$$x < y :\iff \begin{cases} x_1 < y_1 \text{ or} \\ x_1 = y_1 \text{ and } x_2 < y_2 \end{cases}$$

extends $<$ to M_{n+1}. We complete the construction by $y < x$ for $x \in M_{n+1}$, $y \in M_i$ with $i \leq n$. The uniqueness is evident. $\square$

Definition 4 *The above total order relation on M is called Catalan relation.*

The following algorithm produces a list of $M_0, \ldots, M_n$ ordered by the Catalan relation:

```
begin
    produce(m_0)
    for i = 1 to n do
        for k = 0 to i - 1 do
            begin
                p = n - 1 - k
                for m_1 ∈ M_k do
                    for m_2 ∈ M_p do produce(f^{-1}(m_1, m_2))
            end
end
```

Note that two lists produced by the above algorithm for two different Catalan families represent the Catalan isomorphism between those two families: element n of the first list corresponds to element n of the second list. As an example we give the listings of the free bracket figures and the bracket figures:

M_0	$\varnothing$	x
M_1	()	$(x * x)$
M_2	()()	$(x * (x * x))$
	(())	$((x * x) * x)$
M_3	()()()	$(x * (x * (x * x)))$
	()(())	$(x * ((x * x) * x))$
	(())()	$((x * x) * (x * x))$
	(()())	$((x * (x * x)) * x)$
	((()))	$(((x * x) * x) * x)$
...	...	...

6 Catalan families as species

A *species* is a functor from the category $\mathcal{C}$ of finite sets and bijections into the category $\mathcal{D}$ of finite sets and functions (see [6], [9] or [8]). Two species A, B are called *isomorphic* (written $A \simeq B$) if there exists a natural isomorphism ψ between the functors A and B, i.e. for all objects U in $\mathcal{C}$ there exists a bijection $\psi_U : A[U] \to B[U]$ such that for every morphism $f : U \to V$ in $\mathcal{C}$ the following diagram commutes:

$$\begin{array}{ccc} A[U] & \xrightarrow{A[f]} & A[V] \\ {\scriptstyle \psi_U}\downarrow & & \downarrow{\scriptstyle \psi_V} \\ B[U] & \xrightarrow{B[f]} & B[V] \end{array}$$

Sometimes it is convenient to identify the finite sets and bijections of $\mathcal{C}$ as the sets $\{1, \ldots, n\}$ and the permutations.

Now, e.g. Catalan trees may be interpreted as the species of L-enriched rooted trees (L the species of linear orders) as described in [6] modulo the permutation group S_n on the vertices. But then the image of a permutation on the vertices under the action of the functor maps each equivalence class with respect to S_n onto itself and is hence just the identity. Since all equipotent species which map the morphisms of $\mathcal{C}$ to the identity are trivially isomorphic with respect to every bijection ψ, we certainly need a refined interpretation of Catalan families as species. The idea is to use the notion of *linear species* which we obtain by replacing in the above definition $\mathcal{C}$ by the category $\mathcal{L}$ of linearly ordered finite sets with bijections (see [6]) and sometimes it is also convenient to consider subcategories of $\mathcal{L}$.

Let us first define the (linear) species T of trivalent rooted trees: For the finite set $U_n = \{1, \ldots, 2n\}$ (with its natural linear order) we define $T[U_n]$ to be the set of trivalent rooted trees with $n+2$ leaves as described in Section 1.4. As morphisms we take the group of permutations of U_n generated by the transpositions τ_k exchanging element $2k-1$ and $2k$ ($k = 1, \ldots, n$). The image of τ_k under the action of the functor T is defined as follows: $T[\tau_k]$ horizontally flips every subtree starting in a vertex of level k (see Figure 15).

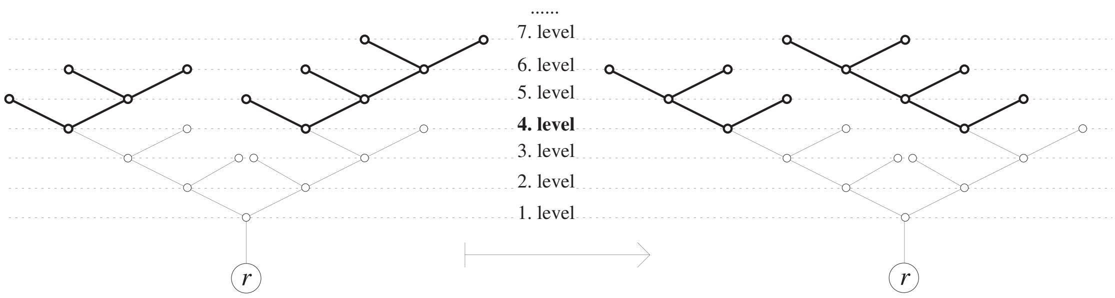

Figure 15: The action of $T[\tau_4]$ on a tree in $T[U_9]$

Now we define the (linear) species B of bracket figures: For the finite set $U_n = \{1, \ldots, 2n\}$ (with its natural linear order) we define $B[U_n]$ to be the set of bracket figures with $n+1$ variables as described in Section 1.2. Again, we take as morphisms the group of permutations of U_n generated by the transpositions τ_k. The image of τ_k under the action of the functor B is defined as follows: $T[\tau_k]$ flips every subfigure at bracket-depth k as indicated in the example below.

Example: The action of $B[\tau_4]$ on a bracket figure in $B[U_9]$

$$(((\underrightarrow{\boldsymbol{(x(xx))}}x)x)(x(\underrightarrow{\boldsymbol{(x(x(xx)))}}x))) \mapsto (((\underleftarrow{\boldsymbol{((xx)x)}}x)x)(x(\underleftarrow{\boldsymbol{(((xx)x)x)}}x)))$$

It is now easy to verify that the so defined species of trivalent rooted trees and bracket figures are isomorphic (as species) and the isomorphism is just the one given in Section 4. Moreover one can check that $T \simeq X \cdot T \cdot T + 1$ in the sense of species (and analogously for B) where X is the species of singletons and 1 is the species of the empty set (for this and the definition of the operations "$\cdot$" and "+" on linear species see [6]). This leads again to the formulas (1)–(3).

References

[1] W.G. Brown: Historical note on a recurrent combinatorial problem. American Mathematical Monthly **72/6** (1965), 973–977

[2] P. Flajolet, B. Salvy, W. Zimmermann: Automatic average-case analysis of algorithms. Theoretical Computer Science **79** (1991), 37–109

[3] I.P. Goulden, D.M. Jackson: Combinatorial enumeration. Wiley (1983)

[4] P. Hilton, J. Pedersen: Catalan numbers, their generalizations and their uses. The Mathematical Intelligencer **13** (1991), 64–75

[5] M. Jeger: Einführung in die Kombinatorik. Klett, Stuttgart (1979)

[6] A. Joyal: Une théorie combinatoire des séries formelles. Advances in Math. **42** (1981), 1–82

[7] J. Knott: A numbering system for binary trees. Communications of the A.C.M. **20** (1977), 113–115

[8] G. Labelle: Some new computational methods in the theory of species. Springer Lecture notes **1234** (1985), 192–209

[9] J. Labelle: Applications diverses de la théorie combinatoire des espèces de structures. Ann. sc. math. Québec **7** no. 1 (1983), 59–94

[10] C. Sprugnoli: The generation of binary trees as a numerical problem. Journal of the A.C.M. **39** (1992), 317–327

[11] R.P. Stanley: Enumerative combinatorics. Wadsworth & Brooks/Cole, Monterey (1986)

[12] J.H. van Lint: Combinatorial theory seminar Eindhoven university of technology. Springer Lectures Notes in Math. **382** (1974)

Author's address: Norbert Hungerbühler, Mathematik Departement, ETH Zürich, CH-8092 Zürich, Switzerland

e-mail: `buhler@math.ethz.ch`