View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Discrete Applied Mathematics 141 (2004) 209–223

DISCRETE APPLIED MATHEMATICS

www.elsevier.com/locate/dam

Phorma: perfectly hashable order restricted multidimensional arrays

Lauro Lins[a], Sóstenes Lins[b], Sílvio Melo[b]

[a] Centro de Informática, Universidade Federal de Pernambuco, UFPE, Recife, Brazil

[b] Departamento de Matemática, UFPE, Brazil

Received 26 July 2002; received in revised form 28 February 2003; accepted 22 March 2003

Abstract

In this paper we propose a simple and efficient data structure yielding a perfect hashing of quite general arrays. The data structure is named *phorma*, which is an acronym for *p*erfectly *h*ashable *o*rder *r*estricted *m*ultidimensional *a*rray.

© 2003 Elsevier B.V. All rights reserved.

MSC: 05A05; 05C90; 06F99

Keywords: Perfect hash function; Digraph; Implicit enumeration; Nijenhuis–Wilf combinatorial family

1. Motivation

Let $a = a_1a_2\dots a_n$ and $\alpha = \alpha_1\alpha_2\dots\alpha_n$ be n-sequences of positive integers, $\alpha \leqslant a$, meaning $\alpha_i \leqslant a_i$, $i = 1, 2, \dots, n$. Suppose that $f(\alpha)$ is a *symmetric function* on the variables α_i, that is, the value of $f(\alpha)$ does not change if the coordinates of α are permuted in an arbitrary way. To store the function f, it is enough to allocate space for the values of $f(\alpha)$, where $\alpha_i \geqslant \alpha_{i+1}$, $1 \leqslant i \leqslant n-1$. Thus, we need to enumerate the α's satisfying $\alpha \leqslant a$ and the boolean function

$$B^{n\geqslant}_{\text{sym}} = (\alpha_1 \geqslant \alpha_2) \wedge (\alpha_2 \geqslant \alpha_3) \wedge \cdots \wedge \cdots (\alpha_{n-1} \geqslant \alpha_n).$$

The motivation for this work is to enumerate and give a perfect hash function [2,4] for multidimensional arrays which have order restrictions on their entries. The simplest example of this situation is when the restrictions are given by $B^{n\geqslant}_{\text{sym}}$. We show

E-mail addresses: ldl@cin.ufpe.br (L. Lins), sostenes@dmat.ufpe.br (S. Lins), silvio@dmat.ufpe.br (S. Melo).

0166-218X/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0166-218X(03)00374-3

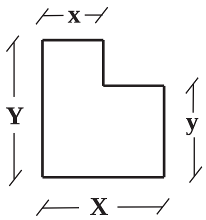

Fig. 1. The L-piece.

that quite general boolean functions can take the place of $B^{n\geqslant}_{\mathrm{sym}}$ and that the large class of enumerative/perfect hash associated problems can be put under a common framework.

To exemplify the appearance of a more complex boolean function, consider the problem of efficiently enumerate all the L-shaped pieces with vertices which fit in a $(p \times q)$ integer grid. This is a typical situation treated in [7]. An *L-shaped piece* is a rectangle R from which we have removed a smaller rectangle $r \subseteq R$. Moreover R and r have a corner in common. By effecting rotations, translations and reflections we may suppose that our L-shaped piece has a corner in the origin and the common vertex to r and R is the vertex opposite to the origin in rectangle R. Positioned in this way, the L-piece is represented by a quadruple of positive integers $(X, Y, x, y) = \alpha_1\alpha_2\alpha_3\alpha_4 \leqslant a_1a_2a_3a_4 = (p, q, p, q)$, as in Fig. 1.

The geometry imposes the restrictions: (1) $X \geqslant x$; (2) $Y \geqslant y$. Symmetry considerations enable us to partition the set of a-bounded L-pieces into equivalent classes and to distinguish a set A of representatives for these classes. For the occupancy purposes in [7], the L-pieces (X, Y, x, y) and (Y, X, y, x) must be considered equivalent. This implies the restrictions: (3) $X \geqslant Y$ and (4) $X = Y \Rightarrow x \geqslant y$. In terms of occupancy, (X, Y, X, y) with $y < X$, which is a degenerated L, can (and must) be replaced by the rectangle (X, Y, X, Y). Analogously, (X, Y, x, Y) with $x < X$ must be replaced by (X, Y, X, Y). In this way, the equivalence $(X = x) \Leftrightarrow (Y = y)$ holds. The equivalence is rewritten as two opposite implications in the disguised form: (5) $(X \neq x) \vee (Y = y)$ and (6) $(Y \neq y) \vee (X = x)$. Restrictions (1)–(6) are gathered in a boolean expression B_L in terms of the α_i's:

$$B_L = (\alpha_1 \geqslant \alpha_3) \wedge (\alpha_2 \geqslant \alpha_4) \wedge (\alpha_1 \geqslant \alpha_2) \wedge ((\alpha_1 \neq \alpha_2) \vee (\alpha_3 \geqslant \alpha_4))$$
$$\wedge((\alpha_1 \neq \alpha_3) \vee (\alpha_2 = \alpha_4)) \wedge ((\alpha_2 \neq \alpha_4) \vee (\alpha_1 = \alpha_3)).$$

So, we want to enumerate the 4-sequences $\alpha = \alpha_1\alpha_2\alpha_3\alpha_4$ of positive integers $\alpha \leqslant a$ and satisfying B_L. If, as it is typically needed in packing problems, a is of order $(120, 100, 120, 100) = (120, 100)^2$ then we have $23,094,225$ α's that satisfies B_L in a total of $144,000,000$ possibilities. If $a = (7, 5)^2$, then there is a total of 190 α's in 1225 possibilities. The valid 190 α's are in 1–1 correspondence with the st-paths in the digraph of Fig. 5.

2. The definition of phorma and the objective of the work

Let $\mathbb{N}$ be the set of natural numbers, $\mathbb{N}^\star = \mathbb{N} \setminus \{0\}$ and $N = \{1, 2, \ldots, n\}$. For $1 \leqslant m \leqslant n$, define $M = \{1, \ldots, m\}$. Let Y^X be the set of all functions from X into Y. Throughout this work, $\alpha = \alpha_1 \ldots \alpha_n$ is an n-sequence of positive integers, that is, $\alpha \in (\mathbb{N}^\star)^N$. The relation $\rho' \leqslant \rho$ for sequences ρ' and ρ of equal length means that $\rho'_i \leqslant \rho_i$, for each i-term of the sequences.

An *n-composition* $\delta = \delta_1 \ldots \delta_m$ is an element of $(\mathbb{N}^\star)^M$ such that $\sum_{1 \leqslant m \leqslant n} \delta_m = n$. The set of n-compositions is denoted by C^n. Given α, let m_α be the number of distinct entries in α and m_δ be the length of δ. Let $\bar{\alpha} = \bar{\alpha}_1 \ldots \bar{\alpha}_{m_\alpha} \in C^n$ denote the n-composition, where $\bar{\alpha}_i$ is the number of occurrences of the ith smallest entry of α.

An *n-phorma* is a triple $P = (a, B, C)$ satisfying: (i) $a = a_1 a_2 \ldots a_n \in (\mathbb{N}^\star)^N$; (ii) B is a boolean function whose literals of B are of type $(\alpha_i \star \alpha_j)$, where $\alpha \in (\mathbb{N}^\star)^N$ and $\star \in \{\leqslant, \geqslant, <, >, =, \neq\}$; (iii) $C \subseteq C^n$ is a given set of n-compositions. The term n-phorma is an acronym for an n-dimensional *p*erfectly *h*ashable *o*rder *r*estricted *m*ultidimensional *a*rray.

The objective of this paper is to enumerate the set

$$A(P) = A(a, B, C) = \{\alpha \mid \alpha \leqslant a,\ \alpha \text{ satisfies } B,\ \bar{\alpha} \in C\}.$$

In the particular case when B is the empty boolean function, then there are no B-restrictions and $A(a, B, C)$ is the subset of $(\mathbb{N}^\star)^N$ consisting of all sequences $\alpha \leqslant a$, $\bar{\alpha} \in C$. We construct a bijection h: $A(P) \to \{0, 1, \ldots, |A(P)| - 1\}$, so that both h and h^{-1} are efficiently computable. Such functions are called *perfect hash functions* [2,4]. Their usefulness is well known.

As far as we know the problem of finding perfect hash functions for these quite general multidimensional arrays have not been considered before in the literature, whence the lack of more specific references and bibliography. Our solution is based on the theory of *combinatorial families* developed in [8]. Here we call these families *NW-families* and recall their definition in Section 4. The central idea is to associate a digraph to a collection of combinatorial objects in such a way that each object in the family is in 1–1 correspondence with a path in the digraph. A more detailed account of these combinatorial families appears in [9].

From a phorma (a, B, C) a digraph $G(a, B, C)$ with a single source s and a single sink t can be constructed so that the elements in $A(a, B, C)$ are in 1–1 correspondence with the st-paths. Indeed, $G(a, B, C)$ is *an NW-family* [8] encoding $A(a, B, C)$ with a simple perfect hash function h. We briefly review these families in Section 4. The digraph $G((7,5)^2, B_L, C^4)$ associated to the phorma $((7,5)^2, B_L, C^4)$ is shown in Fig. 5. In this example, the set C of 4-compositions is the whole set C^4.

3. More applications of phormas

The need to impose order restrictions on arrays appears frequently and in many cases it is not difficult to express these restrictions as a phorma. For a larger example, consider the 7-phormas arising from the generation of *T-shaped pieces*. In Fig. 2 we show the

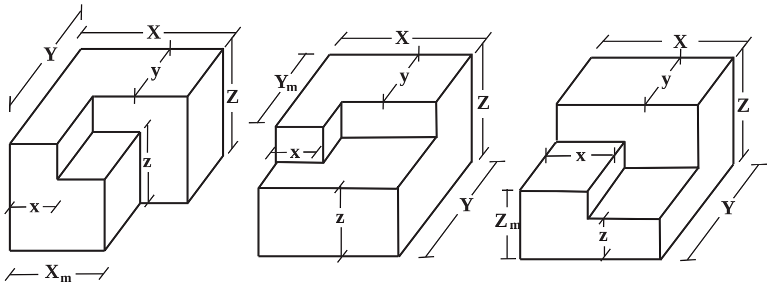

Fig. 2. The T-pieces T_x, T_y and T_z.

three kinds of such a piece. They are composed of a 3-block and a $3D$ L-piece. In the case of the T_z-piece, the L is truncated in one of its legs along the z-direction. These pieces are the $3D$ counterpart for the $2D$ L-shaped piece and they play an important role in $3D$ packing problems. They are described by seven parameters, which in the case of the T_z-piece are, (x, X, y, Y, z, Z_m, Z). To enumerate the T-pieces contained in a $(p \times q \times r)$-block was the motivating idea to formalize the notion of phorma. The need to effect this enumeration appears in [5,6].

As an example, for the T_z-piece, the restrictions coming from the geometry and the symmetry on the seven parameters $xXyYzZ_mZ = \alpha_1\alpha_2\alpha_3\alpha_4\alpha_5\alpha_6\alpha_7$ are of three types:

(1) $(X \geqslant x)$; $(Y \geqslant y)$; $(Z \geqslant Z_m \geqslant z)$;
(2) $(X \geqslant Y)$; $(X = Y) \Rightarrow (x \geqslant y)$;
(3) $(x = X) \Rightarrow (z = Z_m)$; $(y = Y) \Rightarrow (x = X) \wedge (z = Z)$.

The first type of restrictions is obvious. The second type expresses the fact that the T_z-piece can be rotated around a vertical axis without modifying its containment properties. The X- and Y-directions are equivalent. Other axis of rotations, implying similar restrictions, could be used if the boxes to be packed into the T_z-piece could change its vertical. The third type of restrictions deals with the degenerated cases, in which the T_z-piece becomes a simpler piece. In terms of a phorma type boolean function, the restrictions translate as a boolean function B_T^z with the following nine clauses:

$$B_T^z = (\alpha_2 \geqslant \alpha_1) \wedge (\alpha_4 \geqslant \alpha_3) \wedge (\alpha_7 \geqslant \alpha_6) \wedge (\alpha_6 \geqslant \alpha_5) \wedge (\alpha_2 \geqslant \alpha_4)$$
$$\wedge((\alpha_2 \neq \alpha_4) \vee (\alpha_1 \geqslant \alpha_3)) \wedge ((\alpha_1 \neq \alpha_2) \vee (\alpha_5 = \alpha_6))$$
$$\wedge((\alpha_3 \neq \alpha_4) \vee (\alpha_1 = \alpha_2)) \wedge ((\alpha_3 \neq \alpha_4) \vee (\alpha_5 = \alpha_7)).$$

In this case, once more, C is the whole set of 7-compositions C^7. If, just to be specific, $a = (15^2 17^2 19^3)$, then $|A(a, B_T^z, C^7)| = 7,510,130$, while $15^2 17^2 19^3 = 446,006,475$. The amount of memory required to store the digraph $G(a, B_T^z, C^7)$ is logarithmically smaller than $|A(a, B_T^z, C^7)|$ (see Fig. 6) and its construction takes only a few seconds of computer time. Along the same line, we can derive boolean functions B_T^x and B_T^y for the other T-pieces T_x and T_y shown in Fig. 2. The three T-pieces are inequivalent under reflections and rotations which maintain the vertical direction. They play a complementary role in $3D$ packing problems in which the vertical direction of the boxes to be packed must be preserved.

We briefly mention another application of phorma: finding all the solutions for *Cube It*. Let $x<y<z$ be real numbers. Consider the problem of finding all maximum packings of $(x\times y\times z)$-bricks into a cube of side $x+y+z$. If $y+z<3x$, then 27 is an upper bound on the number of bricks that can be packed, see [3]. There exists a phorma $(3^{81}, B_{It}^{\mathrm{Cube}}, c_{27}^3)$ of dimension 81 such that $A(3^{81}, B_{It}^{\mathrm{Cube}}, c_{27}^3)$ has 1008 elements coinciding with the 1008 distinct solutions for the problem of packing the maximum of 27 boxes. In this case, a is the sequence of 81 repetitions of 3, $a=3333\ldots3$ and $C=\{c_{27}^3\}$, where $c_{27}^3=(27,27,27)$. The expression for B_{It}^{Cube} and its justification are too long to be included in this paper. A higher dimensional analogue of B_{It}^{Cube} relates to an interesting open problem which is the subject of ongoing research: how to pack $5^5=3125$ $(a\times b\times c\times d\times e)$-boxes into a 5-cube of side $a+b+c+d+e$. Our implementation (not yet optimized) of the phorma $(3^{81}, B_{It}^{\mathrm{Cube}}, c_{27}^3)$ found the 1008 solutions in about a day of computer time. What is interesting to mention, is that there are no symmetries in these 1008 solutions. So, their set can be partitioned into 21 classes of 48 elements each, corresponding to the symmetry group of the cube. Representatives of these 21 classes are given in Fig. 3. The bricks orientations are

sol 1	sol 2	sol 3	sol 4	sol 5	sol 6	sol 7	sol 8	sol 9	sol 10	sol 11
acC	acC	acC	acC	acC	acC	acC	acC	acC	acB	acB
bCA	bBc	bBc	bBc	bBc	bBA	bAB	bAB	bAB	bCc	bCc
ABb	AaB	AaB	AaB	AaB	Bac	Bac	Bac	Bac	BAa	BAa
cbB	bBA	bBA	bAB	caB	cbB	cbB	cbB	cbB	caC	cbC
aAC	ACa	ACb	ACb	CAb	aAC	aCc	aCc	aBc	AbB	aBA
Bac	Cbc	Cca	Cca	Bca	ACb	ABa	ABa	ACa	aBc	ACb
BaA	caB	caB	cBa	bBA	BaA	BaA	BAa	BaA	BbA	BAa
cBb	CAb	CAa	CaA	ACa	Ccb	cBa	CbA	Cca	CAa	caB
Cca	Bca	Bbc	Bbc	Cbc	cBa	Ccb	cCb	cBb	cCb	Cbc

sol 12	sol 13	sol 14	sol 15	sol 16	sol 17	sol 18	sol 19	sol 20	sol 21
acB	acB	acB	acB	acB	acB	acB	acB	aBc	aBc
bCc	bCc	bCc	bCc	bCc	bCc	bCc	bAC	AaC	AaB
BAa	BAa	BAa	ABa	ABa	ABa	ABa	Bac	cbB	bcC
cbC	cbC	cbC	bAC	bAC	cbC	cbC	cbC	BAa	CAa
aAB	AaB	AaB	AaB	AbB	aBA	aAB	aBA	bcB	Bbc
ACb	aBc	aBc	Ccb	Cac	Bac	Bac	ACb	aCc	caB
BaA	BaA	BAa	cBa	cBa	BAa	BaA	BaA	Ccb	Bcb
cBa	cBa	CbA	CbA	CaA	caB	cBa	Ccb	cBa	cCa
Cbc	Ccb	cCb	Bac	Bcb	Ccb	Ccb	cBa	BaA	aBA

Fig. 3. All the solutions for cube it.

given by the conventions: $a \mapsto yzx$; $A \mapsto zyx$; $b \mapsto xzy$; $B \mapsto zxy$; $c \mapsto xyz$; $C \mapsto yxz$. The parameters of $G(3^{81}, B_{It}^{\text{Cube}}, c_{27}^3)$ are listed in Fig. 6. In particular, $H(3^{81}, B_{It}^{\text{Cube}}, c_{27}^3)$ has only 4 vertices and the whole difficulty is to find $\lfloor A(3^{81}, B_{It}^{\text{Cube}}, c_{27}^3)\rfloor$ which in this case coincides with $A(3^{81}, B_{It}^{\text{Cube}}, c_{27}^3)$.

4. NW-families

The following concept, introduced in [8], is the central tool for our hashing scheme. A *Nijenhuis–Wilf combinatorial family*, or simply an *NW-family*, is an acyclic digraph G whose vertex set is denoted by $V(G)$, having the properties below:

(1) $V(G)$ has a partial order (for $x, y \in V(G)$, $y \preccurlyeq x$ if there is a directed path from x to y) with a unique minimal element t. For each $v \in V(G)$ the set $\{x \in V(G) \mid x \preccurlyeq v\}$ is finite and includes t.
(2) Every vertex v, except t has a strictly positive outvalence $\rho(v)$. For each $v \in V(G)$, the set $E(v)$ of outgoing edges has a *v-local rank-label* ℓ_v, $0 \leqslant \ell_v(e) \leqslant \rho(v) - 1, e \in E(v)$.

A path starting at v and ending in t *is encoded* by the sequence of label-ranks of the sequence of its edges. Such a path is called an object of *order* v [8]. The beauty of this scheme is that we can perform various tasks on the family in an abstract way, without referring to the actual encoding/decoding of the objects as paths. An NW-family is especially suited to deal with the following 5 tasks. Tasks 1–4 are from [8]. Task 0 is emphasized here because of its applicability to the phorma: we need to calibrate the cardinality of $A(a, B, C)$ by choosing a in an adequate way.

Task 0. Counting: What is the family's cardinality? Algorithm: Given $v \in V(G)$, let $|v| = \sum\{|head(e)| \mid e \in E(G), tail(e) = v\}$. From this formula, $|v|$ is easily obtained by recursion. It is convenient to store it as an attribute of $v \in V(G)$ in a pre-processing phase, or *compilation time*.

Task 1. Sequencing: Given an object in the family, construct the "next" object. Algorithm: The *next path* of a given path π in coded form is, in coded form, the lexicographic successor of π.

Task 2. Ranking (perfect hashing): Given an object ω in the family, find the integer $h(\omega)$ such that ω is the $h(\omega)$th element in the order induced by Task 1. Algorithm: Let an element-path π of order v of an NW-family, $\pi = (e_1, e_2, \ldots, e_p)$ be given. The rank of π is defined as $h(\pi) = \sum_{i=1}^{p} \chi(e_i)$, where $\chi(e) = \sum\{|head(f)|$ with $\ell_v(f) < \ell_v(e)$, $f \in E(v)\}$.

Task 3. Unranking: Given an integer r, we need to construct the rth path from v to t. Define $pred_v(e)$ as the highest-rank edge of the set $\{f \in E(v) \mid \ell_v(f) < \ell_v(e)\}$, and let $|head(pred_v(e))| = 0$ if this set is empty. The required rth path π_r is generated as follows. Algorithm: $\pi_r \leftarrow \emptyset$; $r' \leftarrow 0$; $v' \leftarrow v$; **repeat** append to π_r the highest-rank edge e of $E(v')$ such that $r' + |head(pred_{v'}(e))| \leqslant r$; $r' \leftarrow r' + |head(pred_{v'}(e))|$; $v' \leftarrow head(e)$ **until** $v' = t$.

Task 4. Getting random object: Choose an object uniformly at random from the given family. Algorithm: Let $\xi \in [0,1]$ be uniformly chosen at random; return the $(|v| * \xi)$th object.

5. Reducing, sorting, *a*-roofing: the digraph *G*(*a, B, C*)

If α has $m \leqslant n$ distinct entries, let $M_\alpha = \{1,\ldots,m\}$. The *reduction* of α, denoted by $\lfloor\alpha\rfloor$, is the unique surjection in $(M_\alpha)^N$ which is order compatible with α. That is, for $i \in N$, if α_i is the jth smallest entry in α, then $\lfloor\alpha\rfloor_i = j$. Let also $\alpha^{\searrow}$ denote the m-sequence of distinct entries of α in ascending order. We call $\alpha^{\searrow}$ the *sorting* of α. Given an ascending m-sequence γ, let $m_\gamma = m$.

Proposition 1. *The n-vector of positive integers* α *is recoverable from* $(\lfloor\alpha\rfloor, \alpha^{\searrow})$.

Proof. It is sufficient to observe that $\alpha_i = \alpha^{\searrow}_{\lfloor\alpha\rfloor_i}$. □

Since α induces the pair $(\lfloor\alpha\rfloor, \alpha^{\searrow})$ and, by Proposition 1, is recoverable from, it we can think of α as the pair $(\lfloor\alpha\rfloor, \alpha^{\searrow})$ and write $\alpha \equiv (\lfloor\alpha\rfloor, \alpha^{\searrow})$.

For $\alpha \in A(a,B,C)$ let the *a-roof* of α be $\lceil\alpha\rceil^a = \gamma^\star$ where $\gamma^\star$ is the lexicographically maximal increasing m-sequence with the property that $(\lfloor\alpha\rfloor, \gamma^\star) \in A(a,B,C)$. In particular, $\alpha^{\searrow}_i \leqslant \lceil\alpha\rceil^a_i = \gamma^\star_i$, $i \in N$.

Proposition 2. *The a-roof of* α, $\lceil\alpha\rceil^a$, *does not depend on* α *itself but only on* $\lfloor\alpha\rfloor$ *and* a, *in the sense that* $\lceil\alpha\rceil^a = \lceil\lfloor\alpha\rfloor\rceil^a$.

Proof. The a-roof $\lceil\alpha\rceil^a = \gamma_1^\star\gamma_2^\star\ldots\gamma_m^\star$ can be constructed as follows. Suppose that, for $1 \leqslant i \leqslant m$, i occurs at positions $p_{i1},\ldots,p_{ij_i}$ of $\lfloor\alpha\rfloor$. Then we must have $\gamma_m^\star = \min\{a_{p_{m1}}, a_{p_{m2}}, \ldots, a_{p_{mj_m}}\}$, due to a-dominance. For $i = m-1, m-2, \ldots, 1$, the definition implies that $\gamma_i^\star = \min\{a_{p_{i1}},\ldots,a_{p_{ij_i}}, \gamma_{i+1}^\star - 1\}$, by a-dominance and to insure the strict increase of $\gamma^\star$. Since the construction only depended on $\lfloor\alpha\rfloor$ and a, the Proposition is proved. □

Given a phorma (a,B,C) and the corresponding $A(a,B,C)$, three sets are defined

(i) $\lfloor A(a,B,C)\rfloor = \{\lfloor\alpha\rfloor \mid \alpha \in A(a,B,C)\}$,
(ii) $A^{\searrow}(a,B,C) = \{\alpha^{\searrow} \mid \alpha \in A(a,B,C)\}$,
(iii) $\lceil A(a,B,C)\rceil^a = \{\lceil\alpha\rceil^a \mid \alpha \in A(a,B,C)\}$.

Usually, but not necessarily (see the phorma $(3^{81}, B_{It}^{\mathrm{Cube}}, c_{27}^3)$), $|\lfloor A(a,B,C)\rfloor|$ is much smaller than $|A(a,B,C)|$. By Proposition 2, $|\lceil A(a,B,C)\rceil^a| \leqslant |\lfloor A(a,B,C)\rfloor|$. In general this inequality is also not tight. See examples in Fig. 6. The perfect hash function that is constructed for $A(a,B,C)$ depends on the explicit enumeration of the set $\lfloor A(a,B,C)\rfloor$. This set, in the case of our ongoing example, has nine elements,

$$\lfloor A((7,5)^2,B_L,C^4)\rfloor = \{1111,2121,2211,3211,3221,3321,4231,4312,4321\}.$$

The a-roof set has only seven elements because of two duplicates

$$\lceil A((7,5)^2,B_L,C^4)\rceil^a = \{5,57,45,457,457,345,4567,3457,3457\}.$$

Given a phorma (a,B,C) the *digraph* $\Lambda(a,B,C)$ is defined as follows. Its vertex set is $V(\Lambda(a,B,C)) = \{s\} \cup \lfloor A(a,B,C)\rfloor \cup \lceil A(a,B,C)\rceil^a$, where s is a single source. It is a simple graph, and so, each of its directed edges can be represented by an ordered pair of vertices. For each $\lfloor\alpha\rfloor \in \lfloor A(a,B,C)\rfloor$ there are edges $(s,\lfloor\alpha\rfloor)$ and $(\lfloor\alpha\rfloor,\lceil\alpha\rceil^a)$. These are all the edges of $\Lambda(a,B,C)$, concluding its definition. The digraph $\Lambda(a,B,C)$ is a subgraph of $G(a,B,C)$. In Fig. 5, the edges of $\Lambda((7,5)^2,B_L,C^4)$ are depicted in dashed gray. The edges of its complement $H((7,5)^2,B_L,C^4)$ in $G((7,5)^2,B_L,C^4)$ (which we define next) are depicted in solid lines. The number near a vertex v (the first number, when there are two) is the number of vt-paths in $G((7,5)^2,B_L,C^4)$.

Let H^∞ be the set of all finite strictly increasing sequences of positive integers. The empty sequence is in H^∞ and is denoted by t. Suppose $\gamma = \gamma_1\gamma_2\ldots\gamma_m \in H^\infty$. We define an NW-family H_γ as follows. If $\gamma_m > m$ let $\overleftarrow{\gamma}$ denote the increasing sequence of length m satisfying $\overleftarrow{\gamma}_m = \gamma_m - 1$ and $\overleftarrow{\gamma}^i = \min\{\overleftarrow{\gamma}_{i+1} - 1, \gamma_i\}$, for $i = m, m-1, \ldots, 1$. If $\gamma_m = m$, then $\overleftarrow{\gamma}$ does not exist. If $\gamma \neq t$, let $\swarrow\!\gamma$ be the sequence of length $m-1$ obtained from γ by removing its last entry: $\swarrow\!\gamma = \gamma_1\ldots\gamma_{m-1}$. If $\gamma = t$, then $\swarrow\!\gamma$ does not exist. Given $\bar{\gamma},\tilde{\gamma} \in H^\infty$, we say that $\bar{\gamma} \preccurlyeq \tilde{\gamma}$, if there is a sequence $(\tilde{\gamma} = \gamma^1, \gamma^2, \ldots, \gamma^p = \bar{\gamma})$, with $\gamma^i \in H^\infty$, such that, for each $i = 1,2,\ldots,p-1$, either $\gamma^{i+1} = \overleftarrow{\gamma}^i$ or else $\gamma^{i+1} = \swarrow\!\gamma^i$. The relation $\preccurlyeq$ makes H^∞ a partial ordered set, or *poset*. For $\gamma \in H^\infty$, let H_γ be the acyclic digraph whose vertex set is $V(H_\gamma) = \{\gamma' \mid \gamma' \preccurlyeq \gamma\}$. From each vertex $\gamma' \in V(H_\gamma)$ there are at most two outgoing edges: $(\gamma', \overleftarrow{\gamma}')$, of γ'-local rank-label 0, if $\overleftarrow{\gamma}'$ exists and $(\gamma', \swarrow\!\gamma')$, if $\swarrow\!\gamma'$ exists. The γ'-local rank-label of this last edge is either 1, if $\overleftarrow{\gamma}'$ exists or 0 otherwise. This concludes the definition of H_γ.

Given a path π from $\gamma^\star$ to t in $H_{\gamma^\star}$, a *fall* of π is a vertex γ such that the edge $(\gamma, \swarrow\!\gamma)$ is used by π. Path π has exactly $m_{\gamma^\star}$ falls. In Fig. 4 the 4 falls of the path shown in thick edges are: 5678, 567, 34 and 3. The encoding/decoding of the increasing sequences $\gamma \preccurlyeq \gamma^\star$ as paths in the NW-family $H_{\gamma^\star}$ is particularly simple:

Proposition 3. *To a path π in $H_{\gamma^\star}$ from $\gamma^\star$ to t corresponds $\gamma_\pi \preccurlyeq \gamma^\star$ consisting of the last coordinates of the π falls (in reverse order). Reciprocally, to $\gamma \preccurlyeq \gamma^\star$, corresponds the unique path π_γ from $\gamma^\star$ to t such that the last entry of its ith fall coincides with the ith entry of γ. Moreover, $\pi_{\gamma_\pi} = \pi$, $\gamma_{\pi_\gamma} = \gamma$.*

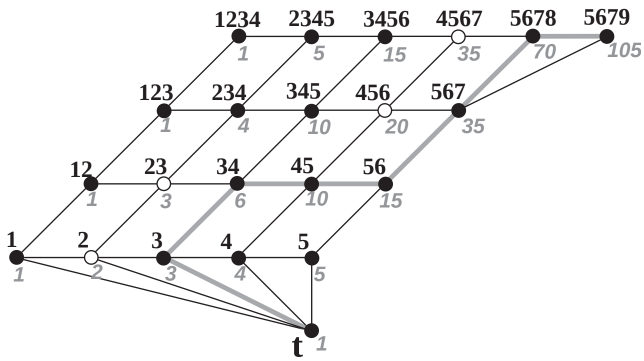

Fig. 4. Path π in H_{5679} with falls 5678, 567, 34 and 3 encoding $\gamma = 3478$, $h(\gamma) = 60$.

Proof. Straightforward from the definitions. □

Given a path π from $\gamma^\star$ to t in $H_{\gamma^\star}$, a *post-fall* of π is a vertex $\gamma' = \overleftarrow{\gamma}$ such that the edge $(\gamma, \swarrow\!\gamma)$ is used by π. Path π has at most $m_{\gamma^\star}$ falls. The set of post-falls of π is denoted $PostFall(\pi)$. In Fig. 4, calling π the path shown in thick edges, we have $PostFall(\pi) = \{4567, 456, 23, 2\}$ and their members are depicted as white vertices. The hash function $h_{\gamma^\star}$ in the NW-family $H_{\gamma^\star}$ takes a simple form:

Proposition 4. *The perfect hash function associated with the NW-family $H_{\gamma^\star}$ is $h_{\gamma^\star}(\gamma) = \sum\{|\gamma'| \,|\, \gamma' \in PostFall(\pi_\gamma)\}$.*

Proof. The result is an specialization of the rank function of a generic NW-family to $H_{\gamma^\star}$. It follows directly from the definitions. □

From this proposition it follows in Fig. 4 that $h_{5679}(3478) = 35 + 20 + 3 + 2 = 60$. The terms of the sum correspond to the orders of the white vertices, forming the set $PostFall(\pi_{3478})$.

Define $H(a, B, C) = \bigcup\{H_{\gamma^\star} \,|\, \gamma^\star \in \lceil A(a, B, C)\rceil^a\}$. Actually, in this union we need only to take maximal $\gamma^\star$'s. If $\gamma' \preccurlyeq \gamma^\star$, then $H_{\gamma'}$ is a subgraph of $H_{\gamma^\star}$ and it is irrelevant for the union. The digraph $H(7575, B_L, C^4)$ shown in Fig. 5, is formed by the union of 4 maximal $\gamma^\star$'s: $H_{3457} \cup H_{4567} \cup H_{457} \cup H_{57}$. In general, the *digraph of a phorma* $P = (a, B, C)$ is defined as $G(a, B, C) = \Lambda(a, B, C) \cup H(a, B, C)$.

In order to make $G(a, B, C)$ an NW-family, we need to define the v-local rank labels of the v-outgoing edges for each vertex v of $G(a, B, C)$. This can be accomplished by ordering lexicographically the elements of $\lfloor A(a, B, C)\rfloor$ and ranking them in the ascending order: $0, 1, \ldots, |\lfloor A(a, B, C)\rfloor| - 1$. The edge $(s, \lfloor\alpha\rfloor)$ gets as s-local rank the same rank as $\lfloor\alpha\rfloor$. An edge of type $(\lfloor\alpha\rfloor, \lceil\alpha\rceil^a)$ gets $\lfloor\alpha\rfloor$-local rank 0, because it is the unique $\lfloor\alpha\rfloor$-outgoing edge. For $\gamma \in V(H(a, B, C))$ we have already defined the γ-local label-ranks. With these local ranks the two conditions of NW-family are

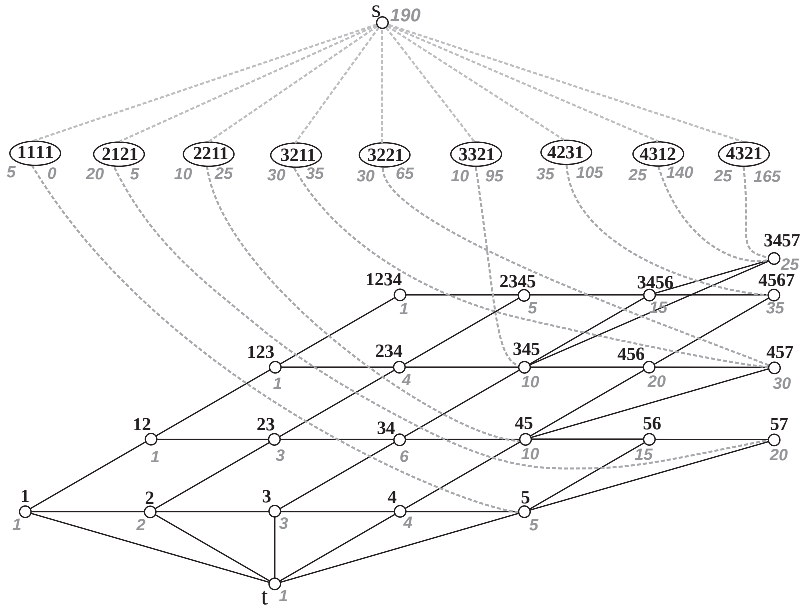

Fig. 5. Digraph $G(7575, B_L, C^4)$ encoding $A(7575, B_L, C^4)$.

satisfied by $G(a,B,C)$. It remains to verify that its *st*-paths encode the elements of $A(a,B,C)$:

Theorem 1 (Main Theorem). *For every phorma $P=(a,B,C)$ the st-paths of $G(a,B,C)$ are in* 1–1 *correspondence with the elements of $A(a,B,C)$.*

Proof. Given an $\alpha \in A(a,B,C)$, let $\lceil\alpha\rceil^a = \gamma^\star$. Define $\pi_\alpha = (s,\lfloor\alpha\rfloor)\circ(\lfloor\alpha\rfloor,\lceil\alpha\rceil^a)\circ\pi_{\alpha^{\searrow}}$. Reciprocally, given an *st*-path π in $G(a,B,C)$, let β be the second vertex of π, $\gamma^\star$ be its third vertex and γ be such that $\pi = (s,\beta)\circ(\beta,\gamma^\star)\circ\pi_\gamma$. Define $\alpha_\pi \equiv (\beta,\gamma)$. These definitions imply $\pi_{\alpha_\pi} = \pi$ and $\alpha_{\pi_\alpha} = \alpha$. □

Given $\lfloor A(a,B,C)\rfloor$ ordered lexicographically and $\beta \in \lfloor A(a,B,C)\rfloor$ define $\|\beta\| = \sum\{|\beta'|$ such that $\beta' < \beta\}$. In Fig. 5 the values of $\|\beta\|$ appear as the second number near each vertex β. The hash function h for a phorma assumes a particularly simple expression:

Proposition 5. *Given $\alpha \equiv (\lfloor\alpha\rfloor, \alpha^{\searrow}) \in A(a,B,C)$, the perfect hash function h of $G(a,B,C)$ is*

$$h(\alpha) = \|\lfloor\alpha\rfloor\| + h_{\lceil\alpha\rceil^a}(\alpha^{\searrow}).$$

Proof. This value of $h(\alpha)$ follows from the general algorithm for ranking in an abstract NW-family, when specialized to phormas. □

6. Implementation aspects

The need of the boolean function B in a phorma (a,B,C) is just to enable the enumeration of $\lfloor A(a,B,C)\rfloor$. If the size of this set is small, then an explicit list of its elements, $\{\beta^1,\beta^2,\ldots,\beta^u\}$, can be given in place of B. If this is not the case, then a convenient way to input a generic phorma type boolean function is by means of a tree $T(B)$ with three types of internal nodes: $\vee$-nodes, $\wedge$-nodes, $\neg$-nodes. The leaves of the tree correspond to the basic constituent boolean functions of type $\alpha_i \star \alpha_j$, where $\star \in \{\leqslant, \geqslant, <, >, =, \neq\}$. The $\neg$-nodes (negation operator) must have at most one child. Note that each subtree rooted at an internal $\Diamond$-node v ($\Diamond \in \{\vee,\wedge,\neg\}$) is a boolean tree obtained by taking the $\Diamond$-operation of the boolean tree(s) corresponding to the children of v. Given an α, it is possible to decide its B-satisfiability, by evaluating from the leaves up and arriving to the root of $T(B)$. See [1] for more details.

We also admit two ways of inputting C: by means of an explicit list of its elements, $\{\delta^1,\delta^2,\ldots,\delta^z\}$, if $z=|C|$ is small, or by a phorma type of boolean restrictions on the coordinates of the δ's. In this case, C is itself a boolean expression with clauses of type $(\delta_i \star \delta_j)$. In the case $C=C^n$, this boolean expression is empty. We define an NW-family encoding $\bigcup_{n\in\mathbb{N}^\star}\{C^n\}$: consider the digraph L^∞, whose vertex set is the set of points in the plane which have positive integer coordinates. There are at most two edges from a point $(p,q)\in V(L^\infty)$, namely a *west edge* $((p,q),(p-1,q))$, if $p\geqslant 2$, and a *southwest edge* $((p,q),(p-1,q-1))$, if $p,q\geqslant 2$. The (p,q)-local rank-label of the first edge is 0, if it exists, and the (p,q)-local rank-label of the second edge is 1, if both edges exist. In the case that only the second edge exists, then its (p,q)-local rank-label is 0. Let C^n_m be the subset of C^n of n-compositions which have length m.

Proposition 6. *The paths from (n,m) to $(1,1)$ in L^∞ are in 1–1 correspondence with the elements of C^n_m. Thus L^∞ is an NW-family encoding the n-compositions for all $n\in\mathbb{N}$.*

Proof. Let $\delta\in C^n_m$ be given. Construct a path π_δ from (n,m) to $(0,0)$ in L^∞ as follows. Let $\delta'\leftarrow\delta$ and $\pi'\leftarrow$ the empty path. Repeat n times: if $\delta'_1>1$, then $\delta'_1\leftarrow\delta'_1-1$, extend π' with a west edge; if $\delta'_1=1$, then δ' becomes δ' without its first part; extend π' with a southwest edge. After the n iterations of this loop, δ' is the composition 1 of 1 in 1 part and define $\pi_\delta=\pi'$. Reciprocally, given a path π from (n,m) to $(1,1)$ in L^∞, construct a $\delta_\pi\in C^n_m$ as follows. Let $\delta'\leftarrow 1$ and $\pi'\leftarrow\pi$. For $i=1,2,\ldots,n$ do: if the ith edge of π is a southwest edge, let $\delta'\leftarrow(1,\delta')$; if the ith edge of π is a west edge, let $\delta'_1\leftarrow\delta'_1+1$. Define $\delta_\pi=\delta'$. These definitions imply $\delta_{\pi_\delta}=\delta$ and that $\pi_{\delta_\pi}=\pi$, establishing a 1–1 correspondence between C^n_m and the paths from (n,m) to $(1,1)$ in L^∞. $\square$

By using Proposition 6 it is possible to generate in an efficient way the δ's satisfying the boolean expression C via a C-restricted implicit enumeration based on L^∞.

The crucial task to construct (at compiler time) the digraph $G(a,B,C)$ is to explicitly generate $\lfloor A(a,B,C)\rfloor$. Since α and $\lfloor\alpha\rfloor$ are order isomorphic, one possibility to produce $\lfloor A(a,B,C)\rfloor$ is to generate all the n^n members of N^N and to test each such sequence for reducibility and B-satisfiability [2]. This simple minded approach is suitable for small dimension n. In our four-dimensional phorma $((7,5)^2, B_L, C^4)$ there are only 256 tests to be made. When n increases this simple minded method becomes inapplicable. For example, for the 7-phorma (a, B_T^z, C^7) there are $7^7 = 823,543$ tests to be made and a better approach is needed to generate the 1134 elements of $\lfloor(A(a, B_T^z, C^7))\rfloor$ as well as the 20 elements of $\lceil A(a, B, C^7)\rceil^a$ (for $a = (15^2 17^2 19^3)$). The basic idea is to implement a *B-restricted* implicit enumerating scheme *which takes into account only reduced sequences* in generating the set $\lfloor(A(a, B, C))\rfloor$. This methodology extends substantially the realm of the phorma applicability.

Given a phorma (a, B, C) and $\delta \in C$. Define

$$\lfloor A(a, B, C, \delta)\rfloor = \{\alpha \in \lfloor A(a, B, C)\rfloor \mid \bar{\alpha} = \overline{\lfloor \alpha \rfloor} = \delta\}.$$

As we know how to generate $\delta \in C$, the generation of $\lfloor A(a, B, C)\rfloor$ reduces to the generation of each $\lfloor A(a, B, C, \delta)\rfloor$, because ($\dot\cup$ means disjoint union)

$$\lfloor A(a, B, C)\rfloor = \dot{\bigcup_{\delta \in C}} \lfloor A(a, B, C, \delta)\rfloor.$$

The *m-dimensional grid digraph* J^m is the digraph whose vertices are the points of $\mathbb{R}^m$ with integer coordinates. There is an edge from $p = p_1 \ldots p_j \ldots p_m$ to $q = q_1 \ldots q_j \ldots q_m$ if $p_i = q_i$ except for $i = j$, where $p_j = q_j + 1$.

Proposition 7. *An element of* $\lfloor A(a, B, C, \delta)\rfloor$ *corresponds to a path from the point* δ *to the origin in* J^{m_δ}.

Proof. Given $\beta = \beta_1\beta_2 \ldots \beta_{m_\delta} \in \lfloor A(a, B, C, \delta)\rfloor$ we define a path named π_β in digraph J^{m_δ} from δ to the origin as follows. Path π_β starts at δ and its ith edge is the edge parallel to the β_ith axis. It follows from the definitions that π_β finishes at the origin. □

From Proposition 7 a B-restricted implicit enumeration scheme based on paths in J^m, only produces reduced words. The construction of $\lfloor A(a, B, C)\rfloor$, $\lceil A(a, B, C)\rceil^a$, and as a consequence, the construction of the digraph $\Lambda(a, B, C)$ are efficiently performed in this way.

Now we turn our attention to the construction and storage of the digraph $H(a, B, C)$. Let $\mathscr{L}(r, m) = \{\gamma \in V(H(a, B, C)) \mid \gamma \in (\mathbb{N}^\star)^M,\ \gamma_m = r\}$ and $\lceil A(a, B, C)\rceil^a_{\max} = \{\gamma^\star \in \lceil A(a, B, C)\rceil^a,\ \gamma^\star \text{ maximal}\}$.

Proposition 8. $|\mathscr{L}(r, m)| \leqslant |\lceil A(a, B, C)\rceil^a_{\max}|$.

Proof. For each element $\gamma \in \mathscr{L}(r, m)$ choose some $\gamma^\star \in \lceil A(a, B, C)\rceil^a_{\max}$ such that $\gamma \preccurlyeq \gamma^\star$. This defines a function f from $\mathscr{L}(r, m)$ to $\lceil A(a, B, C)\rceil^a_{\max}$, given by $f(\gamma) = \gamma^\star$. It is enough to prove that f is injective. Let γ and γ' be distinct elements of $\mathscr{L}(r, m)$. Note that ${}^{\swarrow}\gamma \neq {}^{\swarrow}\gamma'$. Suppose that $\bar{\gamma}$ and $\bar{\gamma}'$ are such that $\gamma \preccurlyeq \bar{\gamma}$ and $\gamma' \preccurlyeq \bar{\gamma}'$. Then it follows that $\bar{\gamma} \neq \bar{\gamma}'$ because the first $m-1$ entries of $\bar{\gamma}$ form ${}^{\swarrow}\gamma$ and the first $m-1$ entries of $\bar{\gamma}'$ form ${}^{\swarrow}\gamma'$. So, f is injective. □

Let $a_\star = \max\{a_i\}$, $n_\star = \max\{m \mid \exists\ \delta \in C \text{ with } m_\delta = m\}$ and v the number of non-empty $\mathscr{L}(r,m)$'s.

Proposition 9. $|V(H(a,B,C))| \leqslant 1 + |\lceil A(a,B,C)\rceil^a_{\max}|(a_\star - (n_\star - 1)/2)n_\star$.

Proof. Clearly, $v \leqslant (a_\star - (n_\star - 1)/2)n_\star$. The term 1 is for the sink t. The inequality follows from Proposition 8. □

The bound given in Proposition 9 is not tight. In general, the maximum value of $|\mathscr{L}(r,m)|$, λ, tends to be much smaller than $|\lceil A(a,B,C)\rceil^a_{\max}|$. A more informative parameter related to the size of $H(a,B,C)$ is μ defined as

$$\mu = |V(H(a,B,C))|/v.$$

For phormas arising in the realm of the applications that we have explored, μ is rather small. Given a vertex γ of this digraph, $\swarrow\!\gamma$ and $\overleftarrow{\gamma}$ are easily obtainable. So the edges of $H(a,B,C)$ do not need to be stored. Each one of the v $\mathscr{L}(r,m)$'s is kept as a lexicographically ordered list indexed by an $(a_\star \times n_\star)$-array. The (r,m)-entry of this array is a pointer to the list $\mathscr{L}(r,m)$. A binary search can then be used to locate a specific member of $\mathscr{L}(r,m)$, when computing h and h^{-1}.

The amount of work needed to compute $h(\gamma)$ is basically proportional to m_γ, the length of γ. Indeed, from Proposition 4 we need only to find the m elements of the set $PostFall(\pi_\gamma)$ and add their orders. These orders are stored at the construction of H_γ. This makes the time for computing $h(\alpha)$ independent of $a_\star$.

Fig. 6 displays basic parameters of various phormas. The following shortcuts are used: $v_G = |V(G(a,B,C))|$, $v_H = |V(H(a,B,C))|$, $\alpha^a = |\lceil A(a,B,C)\rceil^a|$, $\alpha^a_{\max} = |\lceil A(a,B,C)\rceil^a_{\max}|$. The last column of Fig. 6 is $10^4 \times d$, with $d = |A(a,B,C)|/\ (\prod_{i\in N} a_i)$ the *density* of (a,B,C). It is interesting to observe how fast the densities of the symmetric phormas (the ones with $B = B^{n\geqslant}_{\text{sym}}$) go to zero as n increases. We present parameters for the phormas $(9^n, B^{n>}_{\text{sym}}, C^n)$, $2 \leqslant n \leqslant 9$. The boolean functions B^n_{sym} for these phormas are obtained from $B^{n\geqslant}_{\text{sym}}$ by replacing the inequalities $\geqslant$ by the strict inequalities $>$. Thus, only strictly decreasing sequences are permitted. Note that $A(9^{10}, B^{n>}_{\text{sym}}, C^{10}) = \emptyset$.

7. Conclusion

We have defined a data structure generator which permits the perfect hash of order restricted multidimensional arrays $A(a,B,C)$. The restrictions accord a general type of boolean functions B formed by order restricting pairs of entries of the array in arbitrary ways. The boolean function B is used in forming a reduced set $\lfloor A(a,B,C)\rfloor$, inducing a partition of $A(a,B,C)$. An $\lfloor\alpha\rfloor \in \lfloor A(a,B,C)\rfloor$ corresponds to a member subset $\lfloor A(a,B,C,\overline{\lfloor\alpha\rfloor})\rfloor$ of this partition. The elements of $\lfloor A(a,B,C,\delta)\rfloor$, $\delta \in C$, are in 1–1 correspondence with paths from δ to the origin in the m_δ-dimensional integer grid digraph J^{m_δ}, and can be efficiently found in a B-restricted implicit enumeration scheme which produces only reduced sequences. To generate all $c \in C$, which might be itself a boolean function on the δ's, we use the NW-family R^∞ in a C-restricted implicit

Phorma	v_G	v_H	$\lvert\lfloor A\rfloor\rvert$	$\lvert A\rvert$	α^a	α^a_{max}	λ	μ	$10^4 d$
$9^2 B^{2\geq}_{sym}/B^{2>}_{sym}C^2$	20/19	17/17	2/1	45/36	2/1	1/1	1	1.0000	5556/4444
$9^3 B^{3\geq}_{sym}/B^{3>}_{sym}C^3$	29/24	24/22	4/1	165/84	3/1	2/1	1	1.0000	2263/1152
$9^4 B^{4\geq}_{sym}/B^{4>}_{sym}C^4$	39/27	30/25	8/1	495/126	4/1	3/1	1	1.0000	754/192
$9^5 B^{5\geq}_{sym}/B^{5>}_{sym}C^5$	52/28	35/26	16/1	1287/126	5/1	4/1	1	1.0000	218/21
$9^6 B^{6\geq}_{sym}/B^{6>}_{sym}C^6$	72/27	39/25	32/1	3003/84	6/1	5/1	1	1.0000	57/1.58
$9^7 B^{7\geq}_{sym}/B^{7>}_{sym}C^7$	107/24	42/22	64/1	6435/36	7/1	6/1	1	1.0000	13/0.0752
$9^8 B^{8\geq}_{sym}/B^{8>}_{sym}C^8$	173/19	44/17	128/1	12870/9	8/1	7/1	1	1.0000	3/0.0030
$9^9 B^{9\geq}_{sym}/B^{9>}_{sym}C^9$	302/12	45/10	256/1	24310/1	9/1	8/1	1	1.0000	$0.6/10^{-5}$
$9^{10} B^{10\geq}_{sym}C^{10}$	557	45	511	43758	9	8	1	1.0000	0.1
$(7,5)^2 B_L C^4$	32	22	9	190	7	4	2	1.0476	1551
$(40,30)^2 B_L C^4$	164	154	9	245670	7	4	2	1.0621	1706
$(50,40)^2 B_L C^4$	204	194	9	652910	7	4	2	1.0486	1632
$(60,50)^2 B_L C^4$	244	234	9	1420325	7	4	2	1.0400	1578
$(99,50)^2 B_L C^4$	400	390	9	5196500	7	4	2	1.1404	2121
$(100,50)^2 B_L C^4$	404	394	9	5317825	7	4	2	1.1420	2127
$10^7 B^z_T C^7$	1184	49	1134	237325	7	6	1	1.0000	237
$15^7 B^z_T C^7$	1219	84	1134	3853200	7	6	1	1.0000	226
$20^7 B^z_T C^7$	1254	119	1134	28226800	7	6	1	1.0000	221
$25^7 B^z_T C^7$	1289	154	1134	132916875	7	6	1	1.0000	218
$30^7 B^x_T C^7$	1324	189	1134	472460925	7	6	1	1.0000	216
$15^2 17^2 19^3 B^z_T C^7$	1262	127	1134	7510130	20	13	3	1.1651	168
$25^2 27^2 29^3 B^z_T C^7$	1332	197	1134	204089675	20	13	3	1.1006	184
$10^2 50^2 12^3 B^z_T C^7$	1201	66	1134	390270	17	10	2	1.0645	9
$3^{81} B^{Cube}_{It} c^3_{27}$	1013	4	1008	1008	1	1	1	1.0000	10^{-32}

Fig. 6. Parameters values for $G(a,B,C)$ of various phormas.

enumeration search. The whole scheme is summarized by two facts: (i) an $\alpha \in A(a,B,C)$ induces three pieces of information, $\lfloor\alpha\rfloor$, $\overset{\searrow}{\alpha}$ and $\lceil\alpha\rceil^a$ and is recoverable from the first two, $\alpha \equiv (\lfloor\alpha\rfloor, \overset{\searrow}{\alpha})$; (ii) this decomposition reflects in the rank formula for a perfect hashing of $A(a,B,C)$: $h(\alpha) = \|\lfloor\alpha\rfloor\| + h_{\lceil\alpha\rceil^a}(\overset{\searrow}{\alpha})$. This encoding scheme has the power of perfectly addressing huge and quite intricate arrays $A(a,B,C)$ by means of the logarithmically smaller NW-family $G(a,B,C)$. This general type of perfect hash scheme does not seem to have been treated before in the literature. In particular, its use in database systems is a possible source of relevant applications and remains to be investigated.

Acknowledgements

The authors thank A. Bondy for bringing reference [3] to their attention. They also thank three anonymous referees for helpful comments improving the legibility of the paper. The financial support of CNPq, (contract no. 30.1103/80) in the case of the second author, is acknowledged.

References

[1] G. Booch, Object Oriented Design with Applications, The Benjamin/Cummings Publishing Company, Inc., Mento Park, CA, 1991, ISBN 0-8053-0091-0.
[2] T. Cormen, C. Leiserson, R. Rivest, Introduction to Algorithms, The MIT Electrical Engineering and Computer Science Series, The MIT Press, Cambridge, MA, McGraw-Hill Book Company, New York, 1990, ISBN 0-262-03141-8.
[3] D.G. Hoffman, Packing Problems and Inequalities, The Mathematical Gardner, Waderworth, Boston, 1981, ISBN 0-534-98015-5.
[4] D. Knuth, The Art of Computer Programming, Vol. 3, 2nd Edition, Addison-Wesley, Reading, MA, 1975.
[5] L. Lins, Packing in g-blocks, Master Thesis, CIN-UFPE, 2003.
[6] L. Lins, S. Lins, R. Morabito, An n-tet graph approach to non-guillotine packings of n-dimensional boxes into an n-container, Eur. J. Oper. Res. 141 (2002) 421–439.
[7] L. Lins, S. Lins, R. Morabito, An L-approach for packing (ℓ, w)-rectangles into rectangular and L-shaped pieces, J. Oper. Res. Soc. 2003, to appear.
[8] A. Nijenhuis, H.S. Wilf, Combinatorial Algorithms for Computers and Calculators, 2nd Edition, Academic Press, New York, 1978.
[9] H.S. Wilf, East side, west side ... Available in PDF at the home page of the author: www.cis.upenn.edu/∼wilf, 1990.