

Códigos Universais Adaptativos para Representação de Números Inteiros

Márcio J. de C. Lima, Valdemar C. da Rocha Jr. e Bruno T. Ávila.

Resumo— Para uma dada lista de números inteiros arbitrários, em geral não existe um único código universal que represente de forma ótima toda a lista, no sentido de representação com o menor comprimento médio das palavras-código. Essa é a principal motivação desse trabalho para introduzir os códigos universais adaptativos (UA), cuja construção é baseada em códigos com padrão (*pattern codes*). A construção de uma classe de códigos UA baseada nos códigos de Fibonacci é introduzida para uma representação de números inteiros, a qual possui um bom desempenho ao longo de um intervalo mais amplo de valores de números inteiros, em comparação com códigos universais conhecidos.

Palavras-Chave— códigos com padrão, códigos universais, sistemas de numeração, códigos de Fibonacci.

Abstract— For a given arbitrary list of integer numbers, in general there is no known single universal code which is an overall optimum in the sense of representing the whole list with the shortest average codeword length. This is the main motivation in this paper to introduce adaptive universal codes based on pattern codes. The construction of a class of adaptive universal codes based on Fibonacci codes is given for integer representation which perform well over a wider range of integer values in comparison to known universal code constructions.

Keywords— pattern codes, universal codes, numeration systems, Fibonacci codes.

I. INTRODUÇÃO

As várias maneiras de codificar um símbolo arbitrário, freqüentemente em uma representação compacta, representam um problema importante em áreas como teoria dos códigos de fonte [1], [2], [3], [4], compressão de dados [5], [6], [7], [8], teoria de transmissão de dados digitais [9], [10], sincronização de códigos [9], [11], [12], [13] e indexação para acesso rápido à memória [14].

Para uma dada lista de símbolos arbitrários, em que as probabilidades dos símbolos não são conhecidas e nem estimadas, em geral não existe um único código universal que represente de forma ótima toda a lista, no sentido de representação com o menor comprimento médio das palavras-código [6], [7], [8]. Essa é a principal motivação para introduzir os *Códigos Universais Adaptativos* (UA) baseados nos *códigos com padrão* (*pattern codes*), introduzidos por Gilbert [1]. Tais códigos possuem um padrão p que é uma seqüência binária particular finita, que ocorre apenas uma vez em uma palavra-código, quer seja como um prefixo ou como um sufixo. Gilbert

ainda mostra que existe uma classe de códigos com padrão $p = 01 \dots 111$, a qual possui um número máximo de possíveis palavras-código para um dado comprimento, estudados mais tarde por Guibas e Odlyzk [15].

Capocelli [12], em 1989, descreveu um classe de códigos universais capaz de atingir o número máximo de palavras-código descrito por Gilbert. Os códigos de Capocelli são relacionados à representação de Zeckendorf [16] dos números inteiros, em termos dos números generalizados de Fibonacci. Segundo Capocelli, para um dado número inteiro positivo r , esses códigos universais podem ser construídos considerando o conjunto de todas as seqüências binárias de comprimento maior que r , em que a seqüência formada pelo *bit* 0 seguido de r bits 1 consecutivos, denotado por 01_r , ocorre apenas uma vez como sufixo. Essas seqüências binárias formam um conjunto infinito contável de palavras-código de prefixo que é denotado por $S = S(r + 1, 01_r)$.

Em [12], também é demonstrado que os códigos S são códigos de prefixo [17], completos [17] e universais no sentido de Elias [18], [19]. Ainda afirma que os códigos S podem ser reconhecidos por um autômato finito (isto é, são códigos regulares) e, portanto, não podem ser assintoticamente ótimos [20]. Além disso, os códigos S são sincronizáveis e têm um atraso de sincronização igual a um, *i.e.*, caso ocorram erros, após uma única palavra-código o decodificador recupera o sincronismo, quando a sincronização tiver sido perdida [10]. Neste artigo é introduzida a construção de uma classe de códigos universais que possuem propriedades semelhantes aos códigos S . Contudo, apresentam um melhor desempenho, ao longo de um maior intervalo de números inteiros. Estes códigos são denotados por \mathcal{A} e resultam de uma combinação de, pelo menos, dois códigos universais.

Este artigo encontra-se organizado em cinco seções. Na *Seção II* encontra-se uma introdução aos códigos universais clássicos. Na *Seção III* apresenta-se os Códigos Universais Adaptativos (UA). Na *Seção IV* é descrita a construção de uma classe de códigos UA, os códigos Universais Adaptativos de Fibonacci (UAF) baseados nos códigos de Fibonacci, e são feitos comentários sobre os códigos propostos, comparando sua eficiência com a dos códigos universais. Por fim, na *Seção V* são apresentadas algumas conclusões.

II. CÓDIGOS UNIVERSAIS CLÁSSICOS

Depois de examinar vários códigos universais [3], [4], [5], [9], [10], [12], [18], [19], [21], os melhores resultados, *i.e.*, códigos que têm comprimento médio de palavras-código mais curto, são obtidos com os códigos conhecidos como *códigos de Fibonacci* (códigos com padrão), como pode ser visto na

Márcio J. de C. Lima e Valdemar C. da Rocha Jr., Grupo de Pesquisa em Comunicações - CODEC, Departamento de Eletrônica e Sistemas, Universidade Federal de Pernambuco, Caixa Postal 7800, 50711-970, Recife, PE, Brasil, Emails: marcio.lima@ee.ufpe.br e vcr@ufpe.br. B. T. Ávila, Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro, 22453-900, Rio de Janeiro, RJ, Brasil. Email: bavila@inf.puc-rio.br. Este trabalho foi parcialmente financiado pela CAPES e pelo CNPq, projeto 306612/2007-0.

TABELA I

COMPARAÇÃO DO COMPRIMENTO DAS PALAVRAS-CÓDIGO PARA OS CÓDIGOS DE ELIAS [18] γ , δ , ω , OS CÓDIGOS DE WANG [3] $W(2)$, $W(3)$ E $W(4)$, OS CÓDIGO DE YAMAMOTO [4] $Y(2)$, $Y(3)$ E $Y(4)$, OS CÓDIGOS DE APOSTOLICO E FRAENKEL [9], $AF_1^{(2)}$ E $AF_1^{(3)}$, OS CÓDIGOS DE CAPOCELLI [12], $C(3)$ E $C(4)$, E OS CÓDIGOS ADAPTATIVOS $\mathcal{A}(3, 4)$, $\mathcal{A}(3, 5)$ E $\mathcal{A}(4, 5)$.

$\lfloor \log_2(N) \rfloor$	L_γ	L_δ	L_ω	$L_{W(2)}$	$L_{W(3)}$	$L_{W(4)}$	$L_{Y(2)}$	$L_{Y(3)}$	$L_{Y(4)}$	$L_{AF_1^{(2)}}$	$L_{AF_1^{(3)}}$	$L_{C(3)}$	$L_{C(4)}$	$L_{\mathcal{A}(3,4)}$	$L_{\mathcal{A}(3,5)}$	$L_{\mathcal{A}(4,5)}$
0	1	1	1	3,0	4,0	5,0	2,0	3,0	4,0	2	3	3	4	3	3	4
1	3	4	3	4,5	5,0	6,0	3,5	4,0	5,0	3	4	4	5	4	4	5
2	5	5	6	6,0	6,2	7,0	5,0	5,2	6,0	4	5	5	6	5	5	6
3	7	8	7	7,5	7,3	8,1	6,5	6,5	7,1	6	6	6	7	6	6	7
4	9	9	11	9,0	8,5	9,1	8,0	7,7	8,2	7	8	7	8	7	8	8
5	11	10	12	10,5	9,7	10,2	9,5	9,0	9,3	8	9	8	9	8	9	9
6	13	11	13	12,0	10,8	11,3	11,0	10,2	10,5	10	10	9	10	10	10	10
7	15	14	14	13,5	12,0	12,3	12,5	11,5	11,6	11	11	10	11	11	11	11
8	17	15	16	15,0	13,2	13,4	14,0	12,7	12,7	13	12	12	12	12	12	12
9	19	16	17	16,5	14,3	14,5	15,5	14,0	13,8	14	13	13	13	13	14	13
10	21	17	18	18,0	15,5	15,6	17,5	15,2	15,0	16	14	15	14	14	15	14
12	25	19	20	21,0	17,8	17,7	20,0	17,7	17,2	18	17	17	16	17	17	17
14	29	21	22	24,0	20,2	19,8	23,0	20,2	19,5	21	19	20	19	19	19	19
16	33	25	28	27,0	22,5	22,0	26,0	22,7	21,7	24	21	23	21	22	22	21
18	37	27	30	30,0	24,8	24,1	29,0	25,2	24,0	27	23	26	23	24	24	23
20	41	29	32	33,0	27,2	26,3	32,0	27,7	26,2	30	25	29	25	26	26	25
22	45	31	34	36,0	29,5	28,4	35,0	30,2	28,5	33	28	32	28	29	28	28
24	49	33	36	39,0	31,8	30,6	38,0	32,7	30,7	36	30	35	30	31	30	30
26	53	35	38	42,0	34,2	32,7	41,0	35,2	33,0	39	33	38	32	33	33	32
28	57	37	40	45,0	36,5	34,8	44,0	37,7	35,2			40	35			34
30	61	39	42	48,0	38,8	37,0	47,0	40,2	37,5			43	37			

Tabela I, em que L_Ω representa o comprimento do código Ω e N um número inteiro positivo qualquer, e têm uma representação compacta para grandes intervalos de números inteiros N , $1 \leq N \leq 2^M$, ($0 \leq M = \lfloor \log_2(N) \rfloor \leq 30$).

Um código com padrão (P -código) é um conjunto T de seqüências binárias, cada seqüência com comprimento $k \geq r$, em que $r + 1$ é o comprimento do padrão p . Para qualquer seqüência binária $x_1x_2 \dots x_{m+r} \in T$ ($m \geq 0$), de comprimento $k = m + r$, o padrão p ocorre precisamente uma vez, como um prefixo ou sufixo, e.g., o padrão $p = 0111$ de comprimento quatro é um sufixo na palavra-código 1100001000 0111. Note que o P -código é um código de prefixo (*prefix code*) e é assim unicamente decodificável, e é livre de vírgula (*comma free*) ou sincronizável [22].

TABELA II

PRIMEIRAS 16 PALAVRAS-CÓDIGO DOS CÓDIGOS $C(3)$, $C(4)$, $C(5)$, $\mathcal{A}(3, 4)$, $\mathcal{A}(3, 5)$ E $\mathcal{A}(4, 5)$.

N	$B_{C(3)}$	$B_{C(4)}$	$B_{C(5)}$	$B_{\mathcal{A}(3,4)}$	$B_{\mathcal{A}(3,5)}$	$B_{\mathcal{A}(4,5)}$
1	011	0111	01111	011	011	0111
2	0011	00111	001111	0011	0011	00111
3	1011	10111	101111	00011	00011	000111
4	00011	000111	0001111	01011	01011	010111
5	01011	010111	0101111	10111	00011	101111
6	10011	100111	1001111	00011	00111	000111
7	11011	110111	1101111	00111	01011	001111
8	000111	0001111	00011111	01011	10111	010111
9	001011	0010111	00101111	10011	000011	0110111
10	010011	0100111	01001111	11011	000111	1001111
11	100011	1000111	10001111	000011	001011	1101111
12	101011	1010111	10101111	000111	010011	0000111
13	110011	1100111	11001111	001011	010111	0001111
14	111011	1110111	11101111	010011	100111	0010111
15	000011	0000111	00001111	010111	110111	0011111
16	0001011	00001111	000011111	100111	0000011	0100111

Os códigos de Fibonacci apresentam outras propriedades, entre elas pode-se citar a estrutura fácil de ser compreendida,

com uma simples indicação do comprimento de cada palavra-código, ver Tabela II, em que B_Ω é chamado *dicionário* do esquema de codificação Ω . Sobre grandes intervalos de números inteiros, possuem representação compacta, ver Tabela I.

A seqüência generalizada dos números de Fibonacci de ordem m , $m \geq 2$, é definida pelo recursão linear $F_n^{(m)} = F_{n-1}^{(m)} + F_{n-2}^{(m)} + \dots + F_{n-m}^{(m)}$, $n \geq 1$, em que $F_{-m+1}^{(m)} = F_{-m+2}^{(m)} = \dots = F_{-2}^{(m)} = F_{-1}^{(m)} = 0$, $F_0^{(m)} = 1$.

Exemplo 2.1: O sistema de numeração de Fibonacci de ordem 2, cujos elementos são 1, 2, 3, 5, 8, 13, 21, 34, ..., pode ser usado como uma base na representação numérica denotada por $\mathcal{F}^{(2)}$. Logo, os números inteiros 11 e 17 têm suas representações em $\mathcal{F}^{(2)}$ da forma 00101 e 101001, respectivamente, como segue

$$\mathcal{F}^{(2)}(11) = \mathcal{F}^{(2)}(3 + 8) = \begin{matrix} & & & 1 & 2 & 3 & 5 & 8 \\ & & & 0 & 0 & 1 & 0 & 1 \end{matrix}$$

$$\mathcal{F}^{(2)}(17) = \mathcal{F}^{(2)}(1 + 3 + 13) = \begin{matrix} & & & & & & 1 & 2 & 3 & 5 & 8 & 13 \\ & & & & & & 1 & 0 & 1 & 0 & 0 & 1 \end{matrix}$$

◇

Os códigos de Fibonacci são baseados na propriedade de que qualquer número inteiro tem uma única representação binária no sistema de numeração de Fibonacci de ordem m , denotada por $\mathcal{F}^{(m)}$, que não contém m bits 1 consecutivos [11], [23]. A representação $\mathcal{F}^{(2)}$ é conhecida como representação de Zeckendorf [16].

Considere a classe generalizada dos códigos de Fibonacci descrita por Capocelli [12], denotando por $C(r)$ um código de Fibonacci que emprega um padrão de comprimento r . Note que $C(r) = S(r + 1, 01_r)$. Por exemplo, na Tabela II $C(4)$ denota o código de Fibonacci com padrão $p = 0111$ de comprimento quatro. Os códigos $C(3)$, $C(4)$ e $C(5)$, pertencentes à família dos códigos de Fibonacci, tal como

apresentado por Capocelli et al. [12], [20], [21], aparecem ilustrados na Tabela II.

III. CÓDIGOS UNIVERSAIS ADAPTATIVOS

A idéia de introduzir códigos universais adaptativos (UA) consiste em combinar dois ou mais *códigos com padrão*, de modo que o código adaptativo resultante se beneficie das propriedades dos códigos que o compõem pelo preço de uma pequena perda na eficiência. O código adaptativo \mathcal{A} combina vários códigos com padrão, com o objetivo de representar todos os intervalos de forma tão boa quanto o melhor código com padrão para um dado intervalo. Assim pode-se definir o Código Adaptativo \mathcal{A} da seguinte forma:

Definição 3.1: $\mathcal{A}(C_1, \dots, C_s)$ é um código composto pela combinação de s códigos com padrão C_1, C_2, \dots, C_s . $X \in \mathcal{A}(C_1, \dots, C_s)$ se X for uma seqüência binária da forma $x_1x_2 \dots x_kx_{k+1}x_{k+2} \dots x_{k+l}p_1p_2 \dots p_{m_i}$, em que X é escolhida do código C_i , $1 \leq i \leq s$, observando os k dígitos iniciais da seqüência X , i.e., $x_1x_2 \dots x_k$.

Pela definição 3.1 pode-se ressaltar as seguintes observações:

1. A seqüência X pertence tanto ao código $\mathcal{A}(C_1, \dots, C_s)$ quanto ao código C_i .
2. Como o segmento inicial de comprimento k da seqüência X determina qual código C_i com padrão é usado, k deve obedecer à desigualdade

$$k \geq \lceil \log_2 s \rceil, \quad (1)$$

garantindo a utilização dos r códigos com padrão.

3. O código $\mathcal{A}(C_1, \dots, C_s)$ não é necessariamente um código com padrão, segundo a definição de Gilbert [1], contudo respeita o fato do padrão $p_i = p_1p_2 \dots p_{m_i}$ ocorrer somente uma vez, no final da seqüência X .

Exemplo 3.1: Sejam C_1 e C_2 códigos com padrão, $p_1 = 011$ e $p_2 = 100$, respectivamente, em que o padrão só ocorre uma vez como sufixo. As 16 primeiras palavras-código dos códigos C_1 e C_2 são mostradas nas duas primeiras colunas da Tabela III. Dessa forma, tem-se $s = 2 \Rightarrow k \geq \lceil \log_2 2 \rceil = 1$. Assumindo $k = 1$, o primeiro *bit* das palavras de \mathcal{A} identifica qual código está sendo usado. Pode-se designar que as palavras-código de \mathcal{A} iniciadas por 0 pertencem ao código C_1 e as palavras-código de \mathcal{A} iniciadas por 1 pertencem ao código C_2 . As 16 primeiras palavras-código do código adaptativo $\mathcal{A}(C_1, C_2)$ são mostradas na terceira coluna da Tabela III, as palavras em negrito na coluna do código $\mathcal{A}(C(3), C(4))$ provêm do código C_2 e as não destacadas são provenientes do código C_1 . \diamond

Exemplo 3.2: Considere os códigos $C(3)$ e $C(4)$ de Fibonacci, definidos por Capocelli, que usam respectivamente os padrões 011 e 0111. Dessa forma, tem-se $s = 2 \Rightarrow k \geq \lceil \log_2 2 \rceil = 1$. Assumindo $k = 1$, o primeiro *bit* das palavras-código de \mathcal{A} identifica qual código está sendo usado. Pode-se designar que as palavras-código de \mathcal{A} iniciadas por 0 pertencem ao código $C(3)$ e as palavras-código de \mathcal{A} iniciadas por 1 pertencem ao código $C(4)$. As 16 primeiras palavras-código dos códigos $C(3)$, $C(4)$ e $\mathcal{A}(C(3), C(4))$ são mostradas na Tabela III, as palavras em negrito na coluna

do código $\mathcal{A}(C(3), C(4))$ provêm do código $C(4)$ e as não destacadas são provenientes do código $C(3)$. \diamond

TABELA III

PRIMEIRAS 16 PALAVRAS-CÓDIGO DOS CÓDIGOS COM PADRÃO C_1, C_2 $C(3), C(4)$ E DOS CÓDIGO ADAPTATIVOS $\mathcal{A}(C_1, C_2)$ E $\mathcal{A}(C(3), C(4))$.

N	B_{C_1}	B_{C_2}	$B_{\mathcal{A}(C_1, C_2)}$	$B_{C(3)}$	$B_{C(4)}$	$B_{\mathcal{A}(C(3), C(4))}$
1	011	100	011	011	0111	0011
2	0011	1100	100	0011	00111	00011
3	1011	0100	0011	1011	10111	01011
4	00011	11100	1100	00011	000111	10111
5	01011	10100	00011	01011	010111	000011
6	10011	01100	01011	10011	100111	001011
7	11011	00100	11100	11011	110111	010011
8	000011	111100	10100	000011	0000111	100111
9	001011	110100	000011	001011	0010111	110111
10	010011	101100	001011	010011	0100111	0000011
11	100011	011100	010011	100011	1000111	0001011
12	101011	010100	111100	101011	1000111	0010011
13	110011	001100	110100	110011	1010111	0100011
14	111011	000100	101100	111011	1100111	0101011
15	0000011	1111100	0000011	0000011	1110111	1000111
16	0001011	1110100	0001011	0001011	00000111	1010111

IV. CÓDIGOS UNIVERSAIS ADAPTATIVOS DE FIBONACCI

Como ilustrado na Tabela I, um único código de Fibonacci não é uniformemente bom para representar números inteiros contidos em um intervalo relativamente grande. Essa observação também se aplica a outras famílias de códigos universais, quando usadas na representação de números inteiros. A idéia de introduzir códigos universais adaptativos de Fibonacci (UAF) consiste em combinar dois ou mais códigos de Fibonacci, de modo que o código adaptativo resultante se beneficie das propriedades dos códigos de Fibonacci que o compõem.

A. Construção Geral

Para os números inteiros positivos dados u e v , $u < v$, $C(u)$ e $C(v)$ denotam códigos de Fibonacci (códigos de Capocelli). A seguir será demonstrado como combinar os códigos de Fibonacci $C(u)$ e $C(v)$ para produzir o código Adaptativo de Fibonacci (UAF) $\mathcal{A}(u, v)$.

Seja $\mathcal{B}_{\mathcal{A}(u, v)}$ o dicionário do código $\mathcal{A}(u, v)$, e denote por $\mathcal{B}_{\mathcal{A}(u, v)}(N)$, $\mathcal{B}_{\mathcal{A}(u, v)}(N) \in \mathcal{B}_{\mathcal{A}(u, v)}$ a palavra-código que representa o inteiro positivo N em $\mathcal{A}(u, v)$. Faça $G_k^{(u)} = \sum_{i=0}^k F_i^{(u)}$ e $S_k = G_{k-(v-u)-1}^{(u)} + F_k^{(u-1)}$, em que $G_k^{(u)} = 0$, $k < 0$. Faça $M_k = M_{k-1} + S_k$, para $k \geq 0$, em que $M_k = 0$ para $k < 0$. Finalmente, denote por I_k o intervalo $I_k = [1 + M_{k-1}, M_k]$. A associação de um dado número inteiro positivo N com uma palavra-código de $\mathcal{A}(u, v)$ é construída da seguinte forma:

1. Se $N = 1$, emitir a palavra-código 01_{u-1} .
2. Se $N > 1$, localizar k , tal que $N \in I_k$.
3. Calcular $Q = N - M_{k-1} - 1$.
4. Se $Q \leq F_k^{(u-1)} - 1$, emitir o símbolo 0 seguido pela representação binária $\mathcal{F}^{(u-1)}(Q)$ de comprimento $k - (u - 2)$. Se necessário, completar $\mathcal{F}^{(u-1)}(Q)$ com zeros a esquerda para obter o comprimento $k - (u - 2)$. Adicionar o sufixo do código $C(u)$.

- Se $Q \geq F_k^{(u-1)}$, emitir o símbolo 1 seguido pela representação binária $\mathcal{F}^{(v-1)}(Q - F_k^{(u-1)})$ de comprimento $k - (v - 2)$. Se necessário, completar $\mathcal{F}^{(v-1)}(Q - F_k^{(u-1)})$ com zeros à esquerda para obter o comprimento $k - (v - 2)$. Adicionar o sufixo do código $C(v)$.

B. Código Adaptativo $\mathcal{A}(3, 4)$

Descreve-se em seguida como combinar os códigos de Fibonacci $C(3)$ e $C(4)$ para produzir o UA $\mathcal{A}(3, 4)$. Denote por $\mathcal{B}_{\mathcal{A}(3,4)}$ o dicionário de $\mathcal{A}(3, 4)$, e seja $\mathcal{B}_{\mathcal{A}(3,4)}(N)$, $\mathcal{B}_{\mathcal{A}(3,4)}(N) \in \mathcal{B}_{\mathcal{A}(3,4)}$ a palavra-código que representa o número inteiro positivo N em $\mathcal{A}(3, 4)$. Faça $G_k^{(3)} = \sum_{i=0}^k F_i^{(3)}$ e faça $S_k = G_{k-2}^{(3)} + F_k^{(2)}$, em que $G_k^{(3)} = 0$, $k < 0$. Seja $M_k = M_{k-1} + S_k$, para $k \geq 0$, em que $M_k = 0$ para $k < 0$. Finalmente, denote por I_k o intervalo $I_k = [1 + M_{k-1}, M_k]$. Para um número inteiro positivo dado N , a palavra-código associada ao código adaptativo $\mathcal{A}(C(3), C(4))$ é construída como segue.

- Se $N = 1$, emitir a palavra-código 011.
- Se $N > 1$, localizar k , tal que $N \in I_k$.
- Calcular $Q = N - M_{k-1} - 1$.
- Se $Q \leq F_k^{(2)} - 1$, emitir o símbolo 0 seguido pela representação binária $\mathcal{F}^{(2)}(Q)$ de comprimento $k - 1$. Se necessário, completar $\mathcal{F}^{(2)}(Q)$ com zeros à esquerda para obter o comprimento $k - 1$. Adicionar o sufixo 011 do código $C(3)$.
- Se $Q \geq F_k^{(2)}$, emitir o símbolo 1 seguido pela representação binária $\mathcal{F}^{(3)}(Q - F_k^{(2)})$ de comprimento $k - 2$. Se necessário, completar $\mathcal{F}^{(3)}(Q - F_k^{(2)})$ com zeros à esquerda para obter o comprimento $k - 2$. Adicionar o sufixo 0111 do código $C(4)$.

Na Tabela IV são mostrados os números envolvidos na construção do código $\mathcal{A}(3, 4)$ para $k \leq 8$.

TABELA IV

NÚMEROS USADOS NA CONSTRUÇÃO DO CÓDIGO ADAPTATIVO $\mathcal{A}(3, 4)$ PARA $k \leq 8$.

k	0	1	2	3	4	5	6	7	8
$F_k^{(2)}$	1	1	2	3	5	8	13	21	34
$F_k^{(3)}$	1	1	2	4	7	13	24	44	81
$G_k^{(3)}$	1	2	4	8	15	28	52	96	177
$S_k = G_{k-2}^{(3)} + F_k^{(2)}$	1	1	3	5	9	16	28	49	86
$M_k = M_{k-1} + S_k$	1	2	5	10	19	35	63	112	198
I_k	1-1	2-2	3-5	6-10	11-19	20-35	36-63	64-112	113-198

Exemplo 4.1: Considere a geração da palavra-código $\mathcal{B}_{\mathcal{A}(3,4)}(15)$ representando o número inteiro $N = 15$ usando o código $\mathcal{A}(3, 4)$. Ver Tabelas IV e V, que descrevem o desenvolvimento do código $\mathcal{A}(3, 4)$.

- O valor de k para que $N \in I_k$, i.e., para que $10 < N \leq 19$, é $k = 5$;
- Assim $Q = N - M_{k-1} - 1 = 15 - 10 - 1 = 4$;
- Como $Q < F_4^{(2)} - 1$, emitir o símbolo 0 e representar Q com $\mathcal{F}^{(2)}(Q) = \mathcal{F}^{(2)}(4)$ de comprimento $k - 1 = 3$, i.e., $\mathcal{F}^{(2)}(4) = 101$;

- Finalmente, $\mathcal{B}_{\mathcal{A}(3,4)}(15) = 0101\ 011$ é produzido. ◇

TABELA V

DESENVOLVIMENTO DO CÓDIGO ADAPTATIVO $\mathcal{A}(3, 4)$.

N	$\mathcal{B}_{\mathcal{A}(3,4)}$	Q
1	011	-
2	0 011	0
3	00 011	0
4	01 011	1
5	1 0111	0
6	000 011	0
7	001 011	1
8	010 011	2
9	10 0111	0
10	11 0111	1
11	0000 011	0
12	0001 011	1
13	0010 011	2
14	0100 011	3
15	0101 011	4
16	100 0111	0
17	101 0111	1
18	110 0111	2
19	111 0111	3

O código adaptativo $\mathcal{A}(3, 4)$ combina os dois códigos com padrão $C(3)$ e $C(4)$ ($p_1 = 011$ e $p_2 = 0111$ de comprimento 3 e 4 respectivamente). As palavras-código iniciadas pelo bit 0 são as palavras-código com padrão p_1 e as que são iniciadas pelo bit 1 são as palavras-código com padrão p_2 .

C. Códigos Adaptativos $\mathcal{A}(3, 5)$ e $\mathcal{A}(4, 5)$

Seguindo as notações e definições já descritas, $\mathcal{A}(3, 5)$ e $\mathcal{A}(4, 5)$ denotam códigos universais adaptativos resultantes da combinação dos códigos de Capocelli $C(3)$ e $C(5)$, e $C(4)$ e $C(5)$ respectivamente, usando a construção geral descrita na Seção IV-A. Os valores requeridos para a construção dos códigos $\mathcal{A}(3, 5)$ e $\mathcal{A}(4, 5)$ são apresentados nas Tabelas VI e VII, respectivamente, para $k \leq 8$.

TABELA VI

NÚMEROS USADOS NA CONSTRUÇÃO DO CÓDIGO ADAPTATIVO $\mathcal{A}(3, 5)$ PARA $k \leq 8$.

k	0	1	2	3	4	5	6	7	8
$F_k^{(2)}$	1	1	2	3	5	8	13	21	34
$F_k^{(4)}$	1	1	2	4	8	15	29	56	108
$G_k^{(4)}$	1	2	4	8	16	31	60	116	224
$S_k = G_{k-3}^{(4)} + F_k^{(2)}$	1	1	2	4	7	12	21	37	65
$M_k = M_{k-1} + S_k$	1	2	4	8	15	27	48	85	150
I_k	1-1	2-2	3-4	5-8	9-15	16-27	28-48	49-85	86-150

D. Comentários

A Tabela II apresenta as palavras-código dos códigos adaptativos $\mathcal{A}(3, 4)$, $\mathcal{A}(3, 5)$ e $\mathcal{A}(4, 5)$ para os 16 primeiros números inteiros positivos. A Tabela I mostra os comprimentos das palavras-código para vários códigos universais e para os códigos adaptativos, cobrindo um vasto intervalo de números inteiros N .

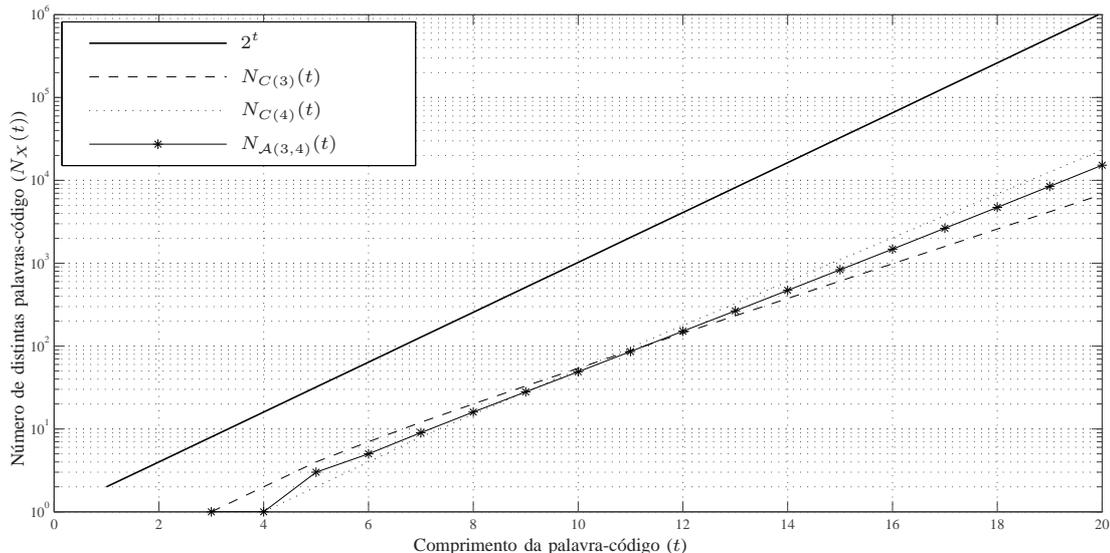


Fig. 1. Comparação do número de palavras-código de comprimento t para a representação natural na base 2, os códigos de Capocelli $C(3)$ e $C(4)$, e o código adaptativo $\mathcal{A}(3, 4)$.

TABELA VII

NÚMEROS USADOS NA CONSTRUÇÃO DO CÓDIGO ADAPTATIVO $\mathcal{A}(4, 5)$ PARA $k \leq 8$.

k	0	1	2	3	4	5	6	7	8
$F_k^{(3)}$	1	1	2	4	7	13	24	44	81
$F_k^{(4)}$	1	1	2	4	8	15	29	56	108
$G_k^{(4)}$	1	2	4	8	16	31	66	116	224
$S_k = G_{k-2}^{(4)} + F_k^{(3)}$	1	1	3	6	11	21	40	75	141
$M_k = M_{k-1} + S_k$	1	2	5	11	22	43	83	158	299
I_k	1-1	2-2	3-5	6-11	12-22	23-43	44-83	84-158	159-299

Observe na Figura 1, em que $N_{\Omega}(t)$ é o número de distintas palavras-código de comprimento t para a representação do esquema de codificação Ω , que o código Universal Adaptativo de Fibonacci $\mathcal{A}(3, 4)$ é mais eficiente que o código de Fibonacci $C(4)$ para $t, t < 11$, e que é mais eficiente do que o código de Fibonacci $C(3)$ para $t, t > 11$. Da mesma forma, na Figura 2, *i.e.*, que o código Universal Adaptativo de Fibonacci $\mathcal{A}(3, 5)$ é mais eficiente que o código $C(5)$ para $t, t < 15$, e que é mais eficiente do que o código $C(3)$ para $t, t \geq 15$.

V. CONCLUSÕES

Códigos Universais Adaptativos (UA) foram introduzidos neste artigo, para representação de números inteiros combinando dois ou mais *códigos com padrão*. Esta representa uma nova abordagem ao problema clássico de representação compacta de uma lista de números inteiros arbitrários. Foi apresentada uma construção de uma classe de códigos UA, baseada nos códigos de Fibonacci, de modo que o código adaptativo resultante se beneficiasse das propriedades dos códigos que o compõem, com o preço de uma pequena perda na eficiência. Esse resultado implica que os códigos adaptativos representam alternativas relevantes para a codificação de símbolos arbitrários, freqüentemente em uma representação

compacta, com aplicações nas diversas áreas que utilizam códigos universais.

AGRADECIMENTOS

Este trabalho recebeu apoio parcial da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), projetos 306612/2007-0 e 141346/2008-6, e do Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal de Pernambuco.

REFERÊNCIAS

- [1] E. N. Gilbert, "Synchronization of binary techniques," *IEEE Trans. Inform. Theory*, COM-28, pp. 1204-1213, Setembro 1960.
- [2] J. Rissanen e G. G. Langdon, "Universal modeling and coding", *IEEE Trans. Inform. Theory*, vol. IT-27, no. 2, pp. 12-23, 1981.
- [3] M. Wang e A.S. Fraenkel, "Almost asymptotically optimal flag encoding of the integers," *IEEE Trans. Inform. Theory* vol. IT-34, no. 2, pp. 238-245, 1988.
- [4] M. Wang e A.S. Fraenkel, "A new asymptotically optimal code for the positive integers," *IEEE Trans. Inform. Theory* vol. IT-37, no. 5, pp. 1420-1429, 1991.
- [5] P. M. Fenwick, "Variable length integer codes based on the Goldbach conjecture and other Additive Codes," *IEEE Trans. Inform. Theory* vol. IT-48, no. 8, pp. 1839-1848, 2002.
- [6] P. Fenwick, "Universal Codes," em *Lossless Compression Handbook*, pp. 55-78. Academic Press, Agosto 2003.
- [7] K. Sayood, *Introduction to Data Compression*, 3a ed. Morgan Kaufmann, Dezembro 2005.
- [8] D. Salomon, *Data Compression: The Complete Reference*, 3a ed. Springer, Dezembro 2006.
- [9] A. Apostolico e A.S. Fraenkel, "Robust transmission of unbounded strings using Fibonacci representations," *IEEE Trans. Inform. Theory* vol. IT-33, no. 2, pp. 238-245, 1987.
- [10] A.S. Fraenkel e S.T. Klein, "Robust universal complete codes for transmission and compression," *Discrete Applied Mathematics* vol. 64, pp. 31-55, Julho 1996.
- [11] W. H. Kautz, "Fibonacci codes for synchronization control," *IEEE Trans. Inform. Theory*, vol. IT-11, pp. 284 - 292, 1965.
- [12] R. M. Capocelli, "Comments and additions to 'Robust transmission of unbounded strings using Fibonacci representations'," *IEEE Trans. Inform. Theory*, vol. 35, no. 1, pp. 191-193, Janeiro 1989.

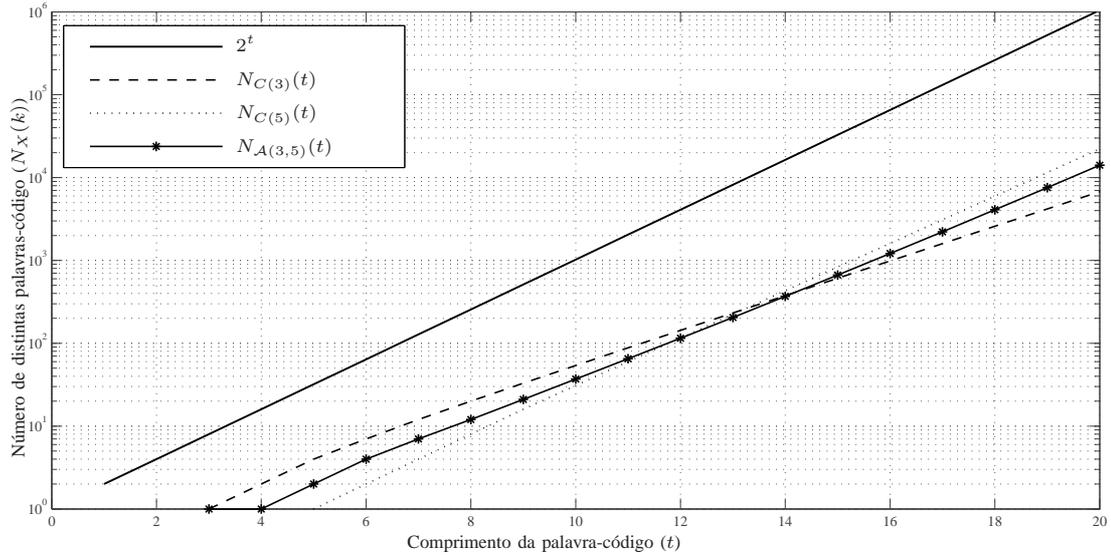


Fig. 2. Comparação do número de palavras-código de comprimento t para a representação natural na base 2, os códigos de Capocelli $C(3)$ e $C(5)$, e o código adaptativo $A(3,5)$.

- [13] A. J. Wijngaarden e H. Morita, "Partial-prefix synchronizable codes," *IEEE Transactions on Information Theory* vol. IT-47, no. 5, pp. 1839–1848, 2001.
- [14] H. E. Williams e J. Zobel, "Compressing integers for fast access," *The Computer Journal* vol. 42, no. 3, pp. 193–201, 1999.
- [15] L. J. Guibas e A. M. Odlyzko, "Maximal Prefix-Synchronized Codes," *SIAM Journal on Applied Mathematics*, vol. 35, no.2, pp. 401–408, Setembro, 1978.
- [16] E. Zeckendorf, "Représentation des nombres naturels par une somme des nombres de Fibonacci ou de nombres de Lucas," *Bull. Soc. Roy. Sci. Liège* 41, pp. 179–182, 1972.
- [17] J. M. Berstel e D. Perrin, *Theory of Codes*. Orlando, FL: Academic, 1985.
- [18] P. Elias, "Universal codeword sets and representations of the integers", *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 194–203, 1975.
- [19] K. B. Lakshmanan, "On universal codeword sets", *IEEE Trans. Inform. Theory*, vol. IT-27, no. 5, pp. 659–662, 1981.
- [20] R. M. Capocelli e A. De Santis, "Regular universal codeword sets," *IEEE Trans. Inform. Theory*, vol. IT-32, no. 1, pp. 129–133, 1986.
- [21] R. M. Capocelli, "Flag encoding related to the Zeckendorf representation of integers," in *Sequences, Combinatorics, Compression, Security, and Transmission*, R. M. Capocelli, Ed. New York: Springer-Verlag, 1990.
- [22] R. Scholtz, "Maximal and variable word-length comma-free codes," *IEEE Transactions on Information Theory* vol. IT-15, no. 2, pp. 300–306, 1969.
- [23] D. E. Knuth, *The Art of Computer Programming*, Volumes 1-3 Boxed Set. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1998.