Efficient VLSI Decompositions for de Bruijn Graphs *

Hiroyuki KAWAKITA[1], Toshinori YAMADA[2], and Shuichi UENO[1] †
1 Department of Communications and Integrated Systems, Tokyo Institute of Technology ‡
2 Department of Information and Computer Sciences, Saitama University §

1 Introduction

This paper shows efficient VLSI decompositions for de Bruijn graphs, which have been found using simulated annealing based on some known sufficient conditions.

A *VLSI decomposition* of a digraph (directed graph) G is a collection of isomorphic vertex-disjoint subdigraphs of G which together span G. A digraph H isomorphic to the subdigraphs comprising the decomposition is called a *building block* for G. The *efficiency* of H, denoted by $\mathit{eff}(H:G)$, is the fraction of the arcs (directed edges) of G which are present in the copies of H. If H is a building block for any digraph in a family of digraphs $\{G_n\}$, H is called a *universal building block* for $\{G_n\}$. Finding an efficient building block for G is motivated by the design of an efficient single VLSI chip with the property that many identical copies of this chip could be wired together to form a circuit represented by G.

We denote the vertex set and arc set of a digraph G by $V(G)$ and $A(G)$, respectively. An arc from vertex u to v is denoted by (u,v). The *nth order d-ary de Bruijn graph* B_n^d is a digraph with $V(B_n^d) = \{0,1,\ldots,d-1\}^n$, and arcs from each vertex $x_1x_2\cdots x_n$ to the vertices $0x_1x_2\cdots x_{n-1}$,$1x_1x_2\cdots x_{n-1}$, ..., and $(d-1)x_1x_2\cdots x_{n-1}$. A *universal d-ary de Bruijn building block of order n* is a universal building block for $\{B_m^d \mid m \geq n\}$. It is known that if H is a universal d-ary de Bruijn building block of order n then $\mathit{eff}(H:B_m^d) = |A(H)|/|A(B_n^d)|$ for every $m \geq n$, which means that the efficiency is independent of m[1, 5]. A binary de Bruijn graph represents circuit diagram for a parallel Viterbi decoder. A universal binary de Bruijn building block of order 7 with efficiency of 0.754 is used by JPL to construct a 64-chip VLSI decomposition for B_{13}^2, which is used to build a single-board Viterbi decoder[1, 2].

Let $X = (x_0, x_1, \ldots, x_k)$ be a sequence of $k+1$ vertices, and $Y = (y_1, y_2, \ldots, y_k)$ be a sequence of k arcs of a digraph. (X,Y) is called a path of length k if the following two conditions are satisfied: 1)$y_i = (x_i, x_{i-1})$ or $y_i = (x_{i-1}, x_i)$ for any i $(1 \leq i \leq k)$; 2)$x_i \neq x_j$ for any i and j with $0 \leq i < j \leq k$. (X,Y) is also called a (x_0,x_k)-path. A path (X,Y) is called a cycle if $x_0 = x_k$. y_i is called a forward arc if $y_i = (x_{i-1}, x_i)$, and a backward arc otherwise. A dipath is a path with no backward arcs. A dipath (X,Y) is also denoted by X for short. If a path has f forward arcs and b backward arcs, we define the *net length* of the path to be $|f-b|$. A cycle is said to be balanced if the net length of the cycle is equal to 0. A cycle is said to be unbalanced if it is not balanced. For a cycle (X,Y), the maximum subnet length of (X,Y) is defined to be the maximum net length of a subpath of (X,Y). A digraph G is connected if there exists a (u,v)-path for any vertices $u, v \in V(G)$.

For a connected digraph G, a mapping $\rho: V(G) \to \mathbb{Z}$ is called a rank function for G if $\rho(y) = \rho(x)+1$ for every $(x,y) \in A(G)$. A balanced cycle C is said to have *property $\mathcal{D}$* if both $|\{x \in V(C)|\ \rho(x) = \min_{y\in V(C)} \rho(y)\}|$ and $|\{x \in V(C)|\ \rho(x) = \max_{y\in V(C)} \rho(y)\}|$ are even for any rank function ρ for C. A digraph G is graded of rank r if there is a rank function $\rho: V(G) \to \{0,1,\ldots,r\}$. G is *graded* if G is graded of rank r for some r.

The cycle space of a digraph G is a vector space generated by the cycles of G. A basis of the cycle space is called a *fundamental k-basis* if the basis is consisting of fundamental cycles with respect to a spanning tree such that the maximum subnet length of each fundamental cycle is at most k.

The following sufficient conditions can be found in the literature.

Theorem I [5] *Let H be a connected spanning subdigraph of B_n^2. If H is graded and has a fundamental $(n+1)$-basis of the cycle space, and every cycle of H has property $\mathcal{D}$ then H is a universal binary de Bruijn building block of order n.* ■

Theorem II [4] *Let H be a connected spanning subdigraph of B_n^d. If H is graded and has a fundamental n-basis of the cycle space then H is a universal d-ary de Bruijn building block of order n. In particular, a spanning tree of B_n^d is a universal d-ary de Bruijn building block of order n.* ■

Our universal building blocks have been found using simulated annealing based on sufficient conditions above.

2 Results

We used simulated annealing to find an efficient universal d-ary de Bruijn building block of order n. A configuration is a spanning tree of B_n^d. The neighborhood of a spanning tree T is the spanning trees T' such that $|A(T')-A(T)| = 1$. For a spanning tree T of B_n^d, let H_T be a unique maximal universal d-ary de Bruijn building block of order n obtained from T by adding arcs of B_n^d in such a way that H_T satisfies the condition of Theorem I if $d = 2$, and the condition of Theorem II if $d \geq 3$. The cost of T is defined as $1-\mathit{eff}(H_T : B_n^d)$.

*DeBruijn グラフの効率的な VLSI 分解
†川喜田 裕之, 山田 敏規, 上野 修一
‡東京工業大学 集積システム専攻
§埼玉大学 情報システム工学科

$n \backslash d$	2			3			4			5			6		
	a_n	e_n	[2]	a_n	e_n	[3]	a_n	e_n	[3]	a_n	e_n	[3]	a_n	e_n	[3]
1	1	0.250	0.250	2	0.222	0.000	4	0.250	0.000	6	0.240	0.000	9	0.250	0.000
2	3	0.375	0.375	11	0.407	0.222	27	0.422	0.188	53	0.424	0.160	92	0.426	0.139
3	8	0.500	0.500	44	0.543	0.296	140	0.547	0.234	343	0.549	0.192	703	0.542	0.162
4	19	0.594	0.594	153	0.630	0.395	647	0.632	0.340	1946	0.623	0.294	4840	0.622	0.258
5	43	0.672	0.672	499	0.684	0.519	2790	0.681	0.457	10476	0.670	0.403	–	–	0.359
6	92	0.719	0.719	1584	0.724	0.604	11707	0.715	0.546	–	–	0.490	–	–	0.442
7	193	0.754	0.754	4954	0.755	0.661	–	–	0.612	–	–	0.560	–	–	0.512
8	399	0.779	0.777	–	–	0.700	–	–	0.662	–	–	0.616	–	–	0.570
9	819	0.800	–	–	–	0.726	–	–	0.700	–	–	0.661	–	–	0.619
10	1673	0.817	–	–	–	0.743	–	–	0.728	–	–	0.697	–	–	0.659
11	3412	0.833	–	–	–	0.755	–	–	0.749	–	–	0.725	–	–	0.693

Table 1: The most efficient known universal d-ary de Bruijn building blocks.

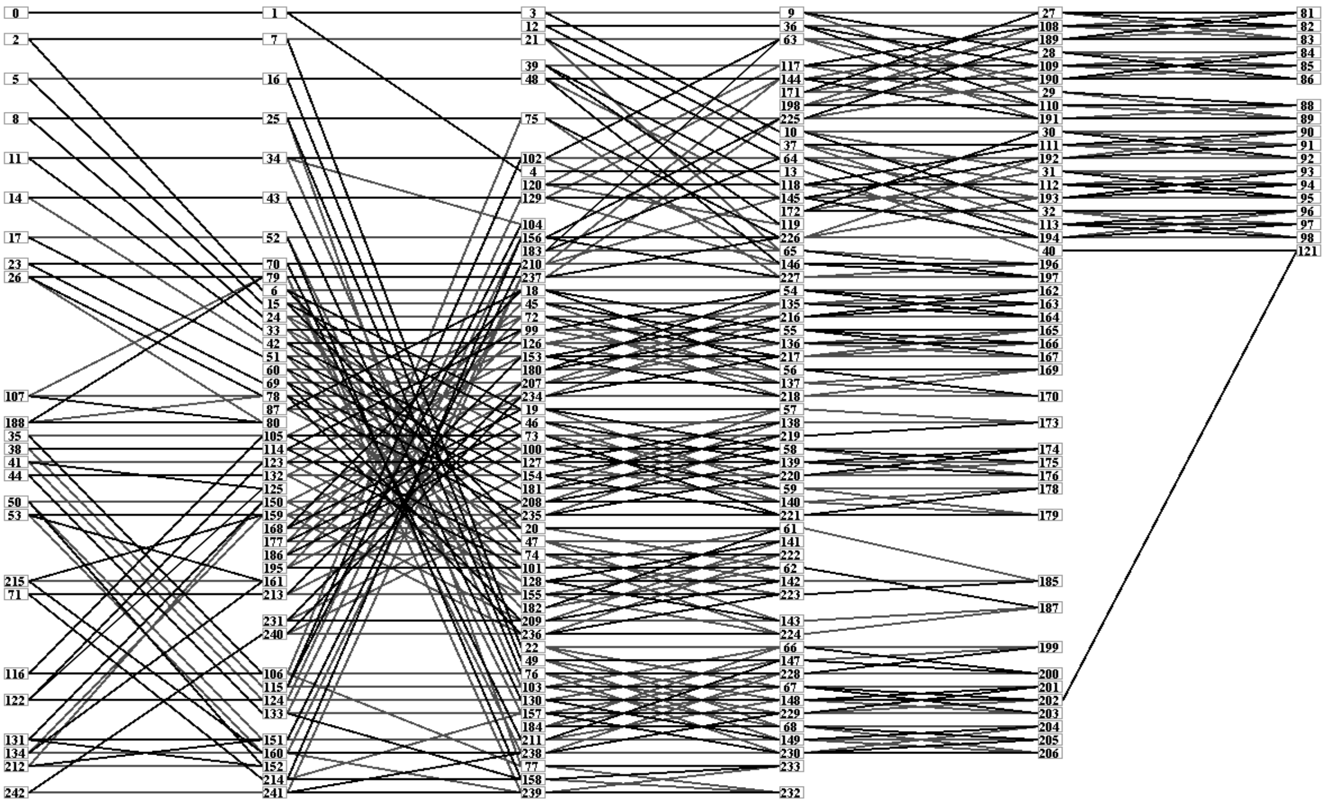

Figure 1: The most efficient known universal ternary de Bruijn building block of order 5.

Table 1 lists the number of arcs and efficiency of the most efficient H_T we have been able to find using simulated annealing, together with the efficiency obtained so far in the literature, [2] for $d = 2$ and [3] for $d \geq 3$. In the table a_n and e_n denote the number of arcs and the efficiency of H_T, respectively.

The best known universal ternary de Bruijn building block of order 5 is shown in Figure 1, in which the ternary strings are represented by their decimal equivalents, and all edges are directed from left to right.

References

[1] O. Colins, S. Dolinar, R. McEliece, and F. Pollara. A VLSI decomposition of the de Bruijn graph. *J. ACM,* Vol. 39, No. 4, pp. 931–948, Oct. 1992.

[2] S. Dolinar, T.-M. Ko, and R. McEliece. Some VLSI decompositions of the deBruijn graphs. *Discrete Math.,* Vol. 106/107, pp. 189–198, 1992.

[3] P. Kopřiva and P. Tvrdík. Decompositions of de Bruijn Networks. *Proc. 1998 International Conference on Parallel and Distributed Systems,* pp. 580–587, 1998.

[4] T. Nishiyama, T. Yamada, and S. Ueno: On VLSI Decompositions for d-ary de Bruijn Graphs; *Technical Report of the Institute of Electronics, Information and Communication Engineers,* Vol.100, No.568, pp.47-54, 2001.

[5] T. Yamada, S. Imai, and S. Ueno. On VLSI Decomposotions for deBruijn Graphs. *Proc. 1999 IEEE International Symposium on Circuits and Systems,* pp. VI 165–VI 169, 1999.