PaToH: **Pa**rtitioning **To**ol for **H**ypergraphs*

Ümit V. Çatalyürek
Department of Biomedical Informatics
The Ohio State University
Columbus, OH 43210
catalyurek.1@osu.edu

Cevdet Aykanat
Computer Engineering Department
Bilkent University
Ankara, 06533 Turkey
aykanat@cs.bilkent.edu.tr

November, 1999

For additional information and documents on PaToH
http://bmi.osu.edu/∼umit/software.html

*This work is partially supported by the Commission of the European Communities, Directorate General for Industry under contract ITDC 204-82166, and Turkish Science and Research Council under grant EEEAG-160.

Contents

1 Introduction

Hypergraph partitioning has been an important problem widely encountered in VLSI layout design [19]. Recent works have introduced new application areas, including one-dimensional and two-dimensional partitioning of sparse matrices for parallel sparse-matrix vector multiplication [5, 4], sparse matrix reordering [4], permuting sparse rectangular matrices into singly-bordered block-diagonal form for parallel solution of LP problems [2]. The hypergraph partitioning problem can be defined as the task of dividing a hypergraph into two or more roughly equal sized parts such that a cost function on the hyperedges connecting vertices in different parts is minimized.

Kernighan-Lin (KL) based heuristics are widely used for graph/hypergraph partitioning because of their short run-times and good quality results. KL algorithm is an iterative improvement heuristic originally proposed for bipartitioning [17]. This algorithm became the basis for most of the subsequent partitioning algorithms. KL algorithm, starting from an initial bipartition, performs a number of passes until it finds a locally minimum partition. Each pass consists of a sequence of vertex swaps. The same swap strategy was applied to hypergraph partitioning problem by Schweikert-Kernighan [20]. Fiduccia-Mattheyses (FM) [8] introduced a faster implementation of KL algorithm for hypergraph partitioning. They proposed vertex move concept instead of vertex swap. This modification as well as proper data structures, e.g., bucket lists, reduced the time complexity of a single pass of KL algorithm to linear in the size of the graph and the hypergraph. Here, *size* refers to the number of edges and pins in a graph and hypergraph, respectively. Krishnamurthy [18] added to FM algorithm a look-ahead ability, which helps to break ties better in selecting a vertex to move. In FM-based algorithms, a vertex is locked as soon as it is moved in a pass, and it remains locked until the end of the pass. Hoffman [14], and Dasdan and Aykanat [7] introduced the dynamic locking approach to relax this restrictive locking mechanism.

The performance of KLFM algorithms deteriorates for large and too sparse graphs/hypergraphs. Here, sparsity of graphs and hypergraphs refer to their average vertex degrees. Furthermore, the solution quality of FM is not *stable* (*predictable*), i.e., average FM solution is significantly worse than the best FM solution, which is a common weakness of move-based iterative improvement approaches. Random multi-start approach is used in VLSI layout design to alleviate this problem by running FM algorithm many times starting from random initial partitions to return the best solution found [1]. However, this approach may not be viable in other applications because of high partitioning overhead. Most users will rely on one run of the partitioning heuristic, so that the quality of the partitioning tool depends equally on the worst and average partitionings than on just the best partitioning.

These considerations have motivated the *two–level* application of FM in hypergraph partitioning. In this approach, a clustering is performed on the original hypergraph $\mathcal{H}_0$ to induce a coarser hypergraph $\mathcal{H}_1$. Clustering corresponds to coalescing highly interacting vertices to supernodes as a preprocessing to FM. Then, FM is run on $\mathcal{H}_1$ to find a bipartition Π_1, and this bipartition is projected back to a bipartition Π_0 of $\mathcal{H}_0$. Finally, FM is re-run on $\mathcal{H}_0$ using Π_0 as an initial solution. Recently, the two–level approach has been extended to *multilevel* approaches [3, 11, 16] leading to fast and successful graph partitioning tools Chaco [12], MeTiS [15], WGPP [10] and reordering tools BEND [13], oMeTiS [15], and ordering code of WGPP [9]. We exploit the successful multilevel methodology to develop a new multilevel hypergraph partitioning tool, called PaToH (PaToH: **Pa**rtitioning **To**ols for **H**ypergraphs).

2 Preliminaries

2.1 Hypergraph Partitioning

A hypergraph $\mathcal{H}=(\mathcal{V},\mathcal{N})$ is defined as a set of vertices $\mathcal{V}$ and a set of nets (hyperedges) $\mathcal{N}$ among those vertices. Every net $n\in\mathcal{N}$ is a subset of vertices, i.e., $n\subseteq\mathcal{V}$. The vertices in a net n are called its *pins* and denoted as $pins[n]$. The size of a net is equal to the number of its pins, i.e., $s_n=|pins[n]|$. The set of nets connected to a vertex v is denoted as $nets[v]$. The degree of a vertex is equal to the number of nets it is connected to, i.e., $d_v=|nets[v]|$. Graph is a special instance of hypergraph such that each net has exactly two pins. Weights and costs can be respectively associated with vertices and nets of a hypergraphs. Let $w[v]$ and $c[v]$ denote the weight of vertex $v \in \mathcal{V}$ and the cost of net $n \in \mathcal{N}$.

$\Pi = \{\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_K\}$ is a *K-way partition* of $\mathcal{H}$ if the following conditions hold:

- each part $\mathcal{V}_k$ is a nonempty subset of $\mathcal{V}$, i.e., $\mathcal{V}_k \subseteq \mathcal{V}$ and $\mathcal{V}_k \neq \emptyset$ for $1 \leq k \leq K$,
- parts are pairwise disjoint, i.e., $\mathcal{V}_k \cap \mathcal{V}_\ell = \emptyset$ for all $1 \leq k < \ell \leq K$
- union of K parts is equal to $\mathcal{V}$, i.e., $\bigcup_{k=1}^{K} \mathcal{V}_k = \mathcal{V}$.

In a partition Π of $\mathcal{H}$, a net that has at least one pin (vertex) in a part is said to *connect* that part. *Connectivity set* Λ_n of a net n is defined as the set of parts connected by n. *Connectivity* $\lambda_n = |\Lambda_n|$ of a net n denotes the number of parts connected by n. A net n is said to be *cut* if it connects more than one part (i.e. $\lambda_n > 1$), and *uncut* otherwise (i.e. $\lambda_n = 1$). The cut and uncut nets are also referred to as *external* and *internal* nets, respectively. In a partition Π of $\mathcal{H}$, a vertex is said to be a *boundary* vertex if it is incident to a cut net. A K-way partition is also called a *multiway* partition if $K > 2$ and a *bipartition* if $K = 2$. A partition is said to be balanced if each part $\mathcal{V}_k$ satisfies the *balance criterion*

$$W_k \leq W_{avg}(1+\varepsilon), \quad for\ k = 1, 2, \ldots, K. \tag{1}$$

In (1), weight W_k of a part $\mathcal{V}_k$ is defined as the sum of the weights of the vertices in that part, i.e.,

$$W_k = \sum_{v \in \mathcal{V}_k} w[v], \tag{2}$$

W_{avg} denotes the weight of each part under the perfect load balance condition, i.e.,

$$W_{avg} = \left(\sum_{v \in \mathcal{V}} w[v]\right)/K, \tag{3}$$

and ε represents the predetermined maximum imbalance ratio allowed.

The set of external nets of a partition Π is denoted as $\mathcal{N}_E$. There are various [6, 21] *cutsize* definitions for representing the cost $\chi(\Pi)$ of a partition Π. Two relevant definitions are:

$$(a) \quad \chi(\Pi) = \sum_{n \in \mathcal{N}_E} c[n] \qquad and \qquad (b) \quad \chi(\Pi) = \sum_{n \in \mathcal{N}_E} c[n](\lambda_n - 1). \tag{4}$$

In (4.a), the cutsize is equal to the sum of the costs of the cut nets. In (4.b), each cut net n contributes $c[n](\lambda_n - 1)$ to the cutsize. The cutsize metrics given in (4.a) and (4.b) will be referred to here as *cut-net* and *connectivity* metrics, respectively. The hypergraph partitioning problem can be defined as the task of dividing a hypergraph into two or more parts such that the cutsize is minimized, while a given balance criterion (1) among part weights is maintained. The hypergraph partitioning problem is known to be NP-hard [19].

2.2 Recursive Bisection

The K-way graph/hypergraph partitioning problem is usually solved by recursive bisection. In this scheme, first a 2-way partition of $\mathcal{H}$ is obtained, and then this bipartition is further partitioned in a recursive manner. After $\lg_2 K$ phases, hypergraph $\mathcal{H}$ is partitioned into K parts. PaToH achieves K-way hypergraph partitioning by recursive bisection for any K value. That is, K is not restricted to be a power of 2.

The cutsize metrics given in (4) need special attention in K-way hypergraph partitioning by recursive bisection. Note that these two metrics become equivalent in hypergraph bisection. Consider a bipartition $\mathcal{V}_\mathcal{A}$ and $\mathcal{V}_\mathcal{B}$ of $\mathcal{V}$ obtained after a bisection step. It is clear that $\mathcal{V}_\mathcal{A}$ and $\mathcal{V}_\mathcal{B}$ and the internal nets of parts $\mathcal{A}$ and $\mathcal{B}$ will become the vertex and net sets of $\mathcal{H}_\mathcal{A}$ and $\mathcal{H}_\mathcal{B}$, respectively, for the following recursive bisection steps. Note that each cut net of this bipartition already contributes 1 (assuming

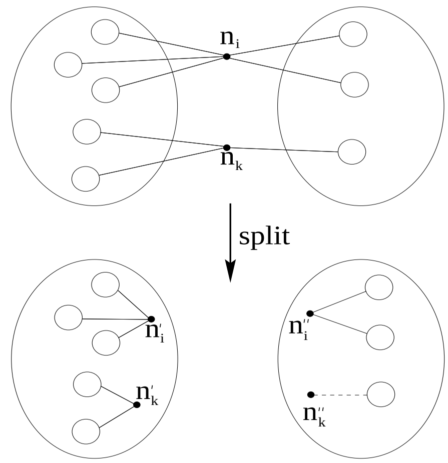

Figure 1: Cut-net splitting during recursive bisection.

unit cost nets) to the total cutsize of the final K-way partition to be obtained by further recursive bisections. Since each cut net will remain to be a cut net in the final K-way partition, all cut nets of this bipartition are discarded in the cut-net metric (4.a). However, in the connectivity metric (4.b), the further recursive bisections of $\mathcal{V}_\mathcal{A}$ and $\mathcal{V}_\mathcal{B}$ may increase the connectivity of these cut nets. Hence, after every hypergraph bisection step, each cut net n_i is split into two pin-wise disjoint nets $n' = pins[n] \bigcap \mathcal{V}_\mathcal{A}$ and $n'' = pins[n] \bigcap \mathcal{V}_\mathcal{B}$, and then these two nets are added to the net lists of $\mathcal{H}_\mathcal{A}$ and $\mathcal{H}_\mathcal{B}$ if $|n'| > 1$ and $|n''| > 1$, respectively. Note that the single-pin nets are discarded during the split operation since such nets cannot contribute to the cutsize in the following recursive bisection steps. Thus, the total cutsize according to (4.b) will become equal to the sum of the number of cut nets at every bisection step by using the above cut-net split method. Figure 1 illustrates two cut nets n_i and n_k in a bipartition, and their splits into nets n'_i, n''_i and n'_k, n''_k, respectively. Note that net n''_k becomes a single-pin net and it is discarded.

3 Multilevel Hypergraph Bisection

The multilevel hypergraph bisection algorithm used in PaToH consists of three phases: *coarsening*, *initial partitioning*, and *uncoarsening*. In the first phase, a bottom-up multilevel clustering is successively applied starting from the original graph by adopting various heuristics until number of vertices in the coarsened graph reduces below a predetermined threshold value. In the second phase, the coarsest graph is bipartitioned using various bottom-up heuristics. In the third phase, partition found in the second phase is successively projected back towards the original graph by refining the projected partitions on intermediate level uncoarser graphs using various top-down iterative improvement heuristics. The following sections briefly summarize these three phases. Although PaToH works on weighted nets, we will assume unit cost nets both for the sake of simplicity of presentation.

3.1 Coarsening Phase

In this phase, the given hypergraph $\mathcal{H}=\mathcal{H}_0=(\mathcal{V}_0,\mathcal{N}_0)$ is coarsened into a sequence of smaller hypergraphs $\mathcal{H}_1 = (\mathcal{V}_1,\mathcal{N}_1)$, $\mathcal{H}_2 = (\mathcal{V}_2,\mathcal{N}_2)$, $\ldots$, $\mathcal{H}_m = (\mathcal{V}_m,\mathcal{N}_m)$ satisfying $|\mathcal{V}_0| > |\mathcal{V}_1| > |\mathcal{V}_2| > \ldots > |\mathcal{V}_m|$. This coarsening is achieved by coalescing disjoint subsets of vertices of hypergraph $\mathcal{H}_i$ into *multinodes* such that each multinode in $\mathcal{H}_i$ forms a single vertex of $\mathcal{H}_{i+1}$. The weight of each vertex of $\mathcal{H}_{i+1}$ becomes equal to the sum of its constituent vertices of the respective multinode in $\mathcal{H}_i$. The net set of each vertex of $\mathcal{H}_{i+1}$ becomes equal to the union of the net sets of the constituent vertices of the respective multinode in $\mathcal{H}_i$. Here, multiple pins of a net $n \in \mathcal{N}_i$ in a multinode cluster of $\mathcal{H}_i$ are contracted to a single pin of the respective net $n' \in \mathcal{N}_{i+1}$ of $\mathcal{H}_{i+1}$. Furthermore, the single-pin nets obtained during this contraction are discarded. Note that such single-pin nets correspond to the internal nets of the clustering performed on $\mathcal{H}_i$. The coarsening phase terminates when the number of vertices in the coarsened hypergraph reduces below pre-determined number.

Clustering approaches can be classified as *agglomerative* and *hierarchical*. In the agglomerative clustering, new clusters are formed one at a time, whereas in the hierarchical clustering several new clusters may be formed simultaneously. In PaToH, we have implemented both randomized matching–based hierarchical clustering schemes and randomized hierarchic–agglomerative clustering schemes . The former and latter approaches will be abbreviated as matching–based clustering and agglomerative clustering, respectively.

The matching-based clustering works as follows. Vertices of $\mathcal{H}_i$ are visited in a random order. If a vertex $u \in \mathcal{V}_i$ has not been matched yet, one of its unmatched *adjacent* vertices is selected according to a criterion. If such a vertex v exists, we merge the matched pair u and v into a cluster. If there is no unmatched adjacent vertex of u, then vertex u remains unmatched, i.e., u remains as a singleton cluster. Here, two vertices u and v are said to be adjacent if they share at least one net, i.e., $nets[u] \cap nets[v] \neq \emptyset$.

The matching-based clustering allows the clustering of only pairs of vertices in a level. In order to enable the clustering of more than two vertices at each level, we have implemented a randomized agglomerative clustering approach. In this scheme, each vertex u is assumed to constitute a singleton cluster $C_u = \{u\}$ at the beginning of each coarsening level. Then, vertices are visited in a random order. If a vertex u has already been clustered (i.e. $|C_u| > 1$) it is not considered for being the source of a new clustering. However, an unclustered vertex u can choose to join a multinode cluster as well as a singleton cluster. That is, all adjacent vertices of an unclustered vertex u are considered for selection according to a criterion. The selection of a vertex v adjacent to u corresponds to including vertex u to cluster C_v to grow a new multinode cluster $C_u = C_v = C_v \cup \{u\}$. Note that no singleton cluster remains at the end of this process as far as there exists no isolated vertex.

3.2 Initial Partitioning Phase

The goal in this phase is to find a bipartition on the coarsest hypergraph $\mathcal{H}_m$. In PaToH, we use *Greedy Hypergraph Growing (GHG)* algorithm for bisecting $\mathcal{H}_m$. This algorithm can be considered as an extension of the GGGP algorithm used in MeTiS to hypergraphs. In GHG, we grow a cluster around a randomly selected vertex. During the coarse of the algorithm, the selected and unselected vertices induce a bipartition on $\mathcal{H}_m$. The unselected vertices connected to the growing cluster are inserted into a priority queue according to their FM gains. Here, the gain of an unselected vertex corresponds to the decrease in the cutsize of the current bipartition if the vertex moves to the growing cluster. Then, a vertex with the highest gain is selected from the priority queue. After a vertex moves to the growing cluster, the gains of its unselected adjacent vertices which are currently in the priority queue are updated and those not in the priority queue are inserted. This cluster growing operation continues until a predetermined bipartition balance criterion is reached. As also mentioned in MeTiS, the quality of this algorithm is sensitive to the choice of the initial random vertex. Since the coarsest hypergraph $\mathcal{H}_m$ is small, we run GHG multiple times starting from different random vertices and select the best bipartition for refinement during the uncoarsening phase.

3.3 Uncoarsening Phase

At each level i (for $i = m, m\text{–}1, \ldots, 1$), bipartition Π_i found on $\mathcal{H}_i$ is projected back to a bipartition Π_{i-1} on $\mathcal{H}_{i-1}$. The constituent vertices of each multinode in $\mathcal{H}_{i-1}$ is assigned to the part of the respective vertex in $\mathcal{H}_i$. Obviously, Π_{i-1} of $\mathcal{H}_{i-1}$ has the same cutsize with Π_i of $\mathcal{H}_i$. Then, we refine this bipartition by running a KLFM-based iterative improvement heuristics on $\mathcal{H}_{i-1}$ starting from initial bipartition Π_{i-1}. PaToH involves a wide range of KLFM-based refinement implementations as listed in Section 4.3.4. Here, we will only discuss the details of our Boundary FM (BFM) implementation. BFM is an FM algorithm that moves only the boundary vertices from the overloaded part to the under-loaded part, where a vertex is said to be a boundary vertex if it is connected to an at least one cut net.

BFM requires maintaining the *pin-connectivity* of each net for both initial gain computations and gain updates. The pin-connectivity $\sigma_k[n] = |n \cap \mathcal{P}_k|$ of a net n to a part $\mathcal{P}_k$ denotes the number of pins of net n that lie in part $\mathcal{P}_k$, for $k = 1,2$. In order to avoid the scan of the pin lists of all nets, we adopt an efficient scheme to initialize the σ values for the first BFM pass in a level. It is clear that initial bipartition Π_{i-1} of $\mathcal{H}_{i-1}$ has the same cut-net set with Π_i of $\mathcal{H}_i$. Hence, we scan only the pin lists of the cut nets of Π_{i-1} to initialize their σ values. For each other net n, $\sigma_1[n]$ and $\sigma_2[n]$ values are easily initialized as $\sigma_1[n]=s_n$ and $\sigma_2[n]=0$ if net n is internal to part $\mathcal{P}_1$, and $\sigma_1[n]=0$ and $\sigma_2[n]=s_n$ otherwise. After initializing the gain value of each vertex v as $g[v]=-d_v$, we exploit σ values as follows. We re-scan the pin list of each external net n and update the gain value of each vertex $v \in pins[n]$ as $g[v] = g[v] + 2$ or $g[v] = g[v] + 1$ depending on whether net n is *critical* to the part containing v or not, respectively. An external net n is said to be critical to a part k if $\sigma_k[n] = 1$ so that moving the single vertex of net n that lies in that part to the other part removes net n from the cut. Note that two-pin cut nets are critical to both parts. The vertices visited while scanning the pin-lists of the external nets are identified as boundary vertices and only these vertices are inserted into the priority queue according to their computed gains.

In each pass of the BFM algorithm, a sequence of unmoved vertices with the highest gains are selected to move to the other part. As in the original FM algorithm, a vertex move necessitates gain updates of its adjacent vertices. However, in the BFM algorithm, some of the adjacent vertices of the moved vertex may not be in the priority queue, because they may not be boundary vertices before the move. Hence, such vertices which become boundary vertices after the move are inserted into the priority queue according to their updated gain values. The refinement process within a pass terminates either no *feasible* move remains or the sequence of last $max\{50,\ 0.001|\mathcal{V}_i|\}$ moves does not yield a decrease in the total cutsize. A move is said to be feasible if it does not disturb the load balance criterion (1) with $K=2$. At the end of a BFM pass, we have a sequence of tentative vertex moves and their respective gains. We then construct from this sequence the maximum prefix subsequence of moves with the maximum prefix sum which incurs the maximum decrease in the cutsize. The permanent realization of the moves in this maximum prefix subsequence is efficiently achieved by rolling back the remaining moves at the end of the overall sequence. The initial gain computations for the following pass in a level is achieved through this rollback. The overall refinement process in a level terminates if the maximum prefix sum of a pass is not positive.

4 Library Interface

PaToH v3.0 library interface consists of two files; a header file `patoh.h` which contains constants, structure definitions and functions proto-types, and a library file `libpatoh.a`. The hypergraph representation used by the library interface is described in Section 4.1, then detail description of the functions are presented in Section 4.2. The parameter structure that is used by the PaToH's recursive multilevel hypergraph partitioner is discussed in the Section 4.3.

Before starting to discuss the details, lets look at a simple C program that partitions an input hypergraph using PaToH functions. The program is displayed in Figure 4. First statement is a function call to read the input hypergraph file which is given by the first command line argument. PaToH partition functions is customizable through a set of arguments, Although user (programmer) can set each of these arguments one by one, it is a good habit to call PaToH function `PaToH_Initialize_Parameters` to set all parameters to default values. After this call, user may prefer to modify the parameters according to his/her need before calling `PaToH_Alloc`. All memory that will be used by PaToH partitioning functions is allocated by `PaToH_Alloc` function, that is there will be no more dynamic memory allocation inside the partitioning functions. Now, we are ready to partition the hypergraph using PaToH's multilevel hypergraph partitioning functions. Call to `PaToH_Partition` will partition the hypergraph and resulting partition vector, part weights and cutsize will be returned in the parameters. Here, variable `cut` will hold the cutsize of the computed partition according to cutsize definition 4(b) since we requested to use this metric by initializing the parameters with constant `PATOH_CONCUT`. User may call partitioning functions as many times as he/she wants before calling function `PaToH_Free`. There is no need to re-allocate the memory before each partitioning call, unless hypergraph is changed. However, changing the coarsening algorithm and number of parts may also require a re-allocation.

4.1 Hypergraph Representation

A hypergraph and its representation can be seen in Figure3.

4.2 Functions

Current PaToH interface contains three function categories; initialization and memory functions, partitioning functions, and utility functions. Following subsections present the detailed descriptions of the functions of each category.

4.2.1 Initialization and memory functions

```
int PaToH_Initialize_Parameters(PPaToH_Parameters pargs, int cuttype,
                                int SuggestByProblemType);
```

Description:
Initializes the parameters that will be used in the partitioning to some default values according to `SuggestByProblemType` parameter.

Parameters:

`pargs`	output	pointer to parameters structure described in Section 4.3. The structure that is pointed by this argument will be filled by this function
`cuttype`	input	must be either `PATOH_CUTPART` for cutnet metric (Equation 4(a)) or `PATOH_CONPART` for "Connectivy-1" metric (Equation 4(b)).
`SuggestByProblemType`	input	Must be set to one of • `PATOH_SUGPARAM_DEFAULT` • `PATOH_SUGPARAM_SPEED` • `PATOH_SUGPARAM_QUALITY`

```
int PaToH_Alloc(PPaToH_Parameters pargs, int _c, int _n, int _nconst,
                int *cwghts, int *nwghts, int *xpins, int *pins);
```

Description:
Allocates the memory that will be used by partitioning algorithms.

Figure 2: A simple C program that partitions an input hypergraph using PaToH functions

```
#include <stdio.h>
#include "patoh.h"

int main(int argc, char *argv[])
{
PaToH_Parameters args;
int             _c, _n, _nconst, *cwghts, *nwghts,
                *xpins, *pins, *partvec, cut, *partweights;


PaToH_Read_Hypergraph(argv[1], &_c, &_n, &_nconst, &cwghts,
                      &nwghts, &xpins, &pins);

printf("Hypergraph %10s -- #Cells=%6d  #Nets=%6d  #Pins=%8d #Const=%2d\n",
       argv[1], _c, _n, xpins[_n], _nconst);

PaToH_Initialize_Parameters(&args, PATOH_CONPART,
                            PATOH_SUGPARAM_DEFAULT);

args._k = atoi(argv[2]);
partvec =  (int *) malloc(_c*sizeof(int));
partweights  = (int *) malloc(args._k*sizeof(int));

PaToH_Alloc(&args, _c, _n, _nconst, cwghts, nwghts,
            xpins, pins);

PaToH_Partition(&args, _c, _n, cwghts, nwghts,
                xpins, pins, partvec, partweights, &cut);

printf("%d-way cutsize is: %d\n", args._k, cut);

free(partweights);
free(partvec);

PaToH_Free();
return 0;
}
```

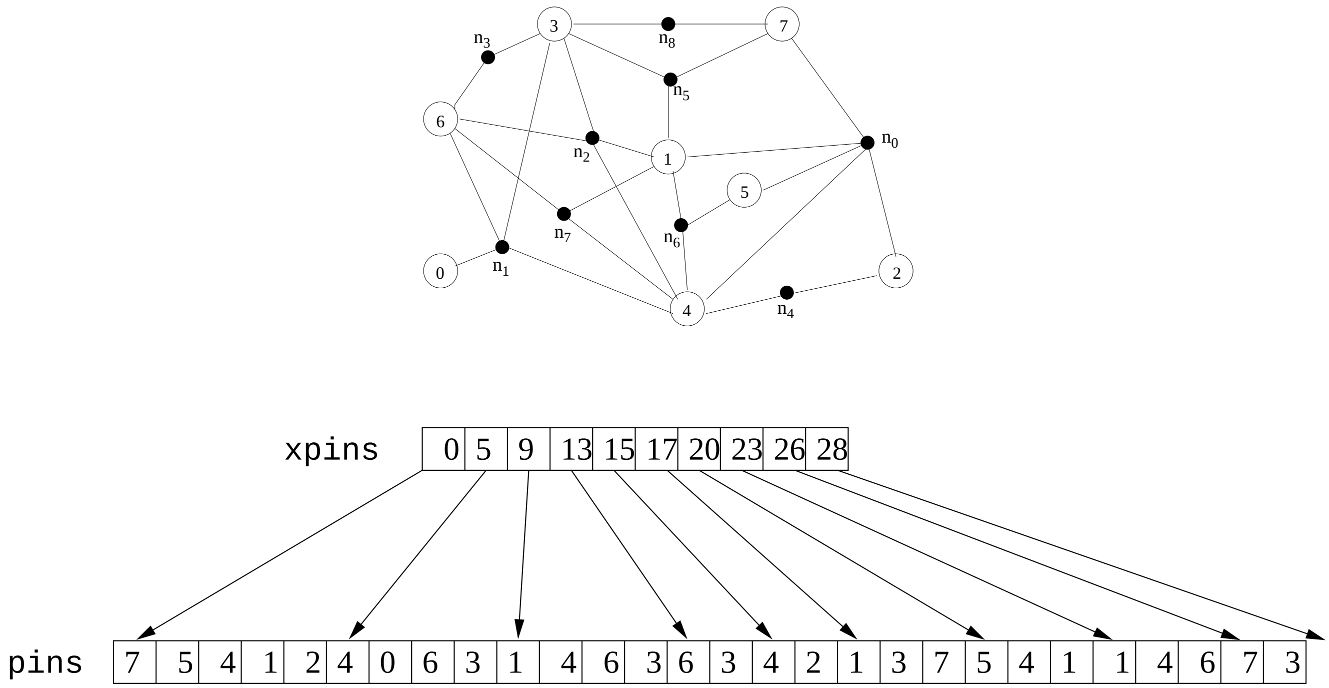

Figure 3: A hypergraph and its representation.

Parameters:

`pargs`	input	pointer to parameters structure described in Section 4.3. Allocation will be done using some of the parameters of this structure.
`_c`	input	number of cells of the hypergraph.
`_n`	input	number of nets of the hypergraph.
`_nconst`	input	number of constraints.
`cwghts`	input	array of size `_c`×`_nconst` that stores the weights of each cell. In multi-constraint partitioning, each cell v_i has `_nconst` weights, and they are stored in `cwghts[i*_ncost]` through `cwghts[(i+1)*_ncost-1]`
`nwghts`	input	array of size `_n` that stores the cost of each net. If hypergraph has unweighted nets, this parameter can be `NULL`.
`xpins`	input	array of size `_n+1` that stores the beginning index of cells connected to nets.
`pins`	input	array that stores the cell-lists of nets. Cells connected to net n_j are stored in `pins[xpins[j]]` through `pins[xpins[j+1]-1]`.

```
int PaToH_Free(void);
```

Description:
Frees the memory allocated by `PaToH_Alloc`.

4.2.2 Partitioning functions

```
int PaToH_Partition(PPaToH_Parameters pargs, int _c, int _n,
        int *cwghts, int *nwghts, int *xpins, int *pins,
        int *partvec, int *partweights, int *cut);
```

Description:
Partitions the hypergraph into `pargs->_k` parts using recursive multilevel hypergraph bisection algorithm.

Parameters:

`pargs`	input	pointer to parameters structure described in Section 4.3. Allocation will be done using some of the parameters of this structure.
`_c`	input	number of cells of the hypergraph.
`_n`	input	number of nets of the hypergraph.
`cwghts`	input	array of size `_c` that stores the weight of each cell
`nwghts`	input	array of size `_n` that stores the cost of each net.
`xpins`	input	array of size `_n+1` that stores the beginning index of cells connected to nets.
`pins`	input	array that stores the cell-lists of nets. Cells connected to net n_j are stored in `pins[xpins[j]]` through `pins[xpins[j+1]-1]`.
`partvec`	output	array of size `_c` that returns the part number of each cell.
`partweights`	output	array of size `pargs->_k` that returns the total part weight of each part.
`cut`	output	cutsize of the solution, according to the requested cutsize metric by `pargs->cuttype`.

```
int PaToH_Partition_with_FixCells(PPaToH_Parameters pargs, int _c,
        int _n, int *cwghts, int *nwghts, int *xpins, int *pins,
        int *partvec, int *partweights, int *cut);
```

Description:
Partitions the hypergraph into `pargs->_k` parts using recursive multilevel hypergraph bisection algorithm. Some of the cells of the hypergraph may have been pre-assigned.

Parameters:

`pargs`	input	pointer to parameters structure described in Section 4.3. Allocation will be done using some of the parameters of this structure.
`_c`	input	number of cells of the hypergraph.
`_n`	input	number of nets of the hypergraph.
`cwghts`	input	array of size `_c` that stores the weight of each cell.
`nwghts`	input	array of size `_n` that stores the cost of each net.
`xpins`	input	array of size `_n+1` that stores the beginning index of cells connected to nets.
`pins`	input	array that stores the cell-lists of nets. Cells connected to net n_j are stored in `pins[xpins[j]]` through `pins[xpins[j+1]-1]`.
`partvec`	in/out	array of size `_c` that stores the part number of each cell belong to. -1 indicates cell is free to assign any part, 0 to `pargs->_k` indicates that cell is pre-assigned to that part.
`partweights`	output	array of size `pargs->_k` that returns the total part weight of each part.
`cut`	output	cutsize of the solution, according to the requested cutsize metric by `pargs->cuttype`.

```
int PaToH_MultiConst_Partition(PPaToH_Parameters pargs, int _c,
        int _n, int _nconst, int *cwghts, int *xpins,
        int *pins, int *partvec, int *partweights, int *cut);
```

Description:
Partitions the hypergraph into `pargs->_k` parts using multi-constraint recursive multilevel hypergraph bisection algorithm

Parameters:

`pargs`	input	pointer to parameters structure described in Section 4.3. Allocation will be done using some of the parameters of this structure.
`_c`	input	number of cells of the hypergraph.
`_n`	input	number of nets of the hypergraph.
`_nconst`	input	number of constraints.
`cwghts`	input	array of size `_c`×`_nconst` that stores the weights of each cell. In multi-constraint partitioning, each cell v_i has `_nconst` weights, and they are store in `cwghts[i*_ncost]` through `cwghts[(i+1)*_ncost-1]`
`xpins`	input	array of size `_n+1` that stores the beginning index of cells connected to nets.
`pins`	input	array that stores the cell-lists of nets. Cells connected to net n_j are stored in `pins[xpins[j]]` through `pins[xpins[j+1]-1]`.
`partvec`	output	array of size `_c` that stores the part number of each cell belong to
`partweights`	output	array of size `pargs->_k`×`_nconst` that returns the total part weight of each part for each constraint.
`cut`	output	cutsize of the solution, according to the requested cutsize metric by `pargs->cuttype`.

4.2.3 Utility functions

```
int PaToH_Check_User_Parameters(PPaToH_Parameters pargs, int verbose);
```

Verifies the user parameters.

Parameters:

`pargs`	input	pointer to parameters structure described in Section 4.3.

```
int PaToH_Read_Hypergraph(char *filename, int *_c, int *_n,
        int *_nconst, int **cwghts, int **nwghts, int **xpins,
        int **pins);
```

Reads a hypergraph from the given file.

Parameters:

`pargs`	input	pointer to parameters structure described in Section 4.3.

```
int PaToH_Compute_Cut(PPaToH_Parameters pargs, int _c, int _n, int *nwghts,
        int *xpins, int *pins, int *partvec);
```

Computes the cut.

Parameters:

`pargs`	input	pointer to parameters structure described in Section 4.3.

```
int PaToH_Compute_Part_Weights(PPaToH_Parameters pargs, int _c, int _nconst,
                               int *cwghts, int *partvec, int *partweights);
```

Computes the part weights.

Parameters:

`pargs`	input	pointer to parameters structure described in Section 4.3.

4.3 Data Structures

User controls the execution of the multilevel bisection algorithm by setting appropriate parameters. First argument of PaToH partitioning functions is a pointer to a structure of type `PaToH_Arguments`. This structure is defined in file `patoh.h`. We have categorized the parameters in four groups. Following subsections briefly describes the each parameter.

4.3.1 Miscellaneous Parameters

- `cuttype`: determines the cost function for partitioning.
- `_k`: number of parts.
- `outputdetail`: detail of verbose output.
- `seed`: seed of the random generator.
- `doinitperm`: if set to a non-zero value, PaToHshuffles the pins and nets lists of the hypergraph prior to partitioning.
- `bisec_fixednetsizetrsh`: During the each bisection nets with size larger than this value will be discarded. Please note that, if such a larger net is splitted during the recursive bisection it may be considered in the further partitionings.
- `bisec_netsizetrsh`: Nets with size larger than `bisec_netsizetrsh`$\times s_{avg}$ are discarded during the each bisections step, where s_{avg} is average net size.
- `bisec_partmultnetsizetrsh`: Nets with size larger than `bisec_partmultnetsizetrsh`$\times K$ are discarded during the each bisections step..

- `bigVcycle`: the maximum number of big V-cycles.
- `smallVcycle`: the maximum number of big V-cycles.
- `usesamematchinginVcycles`: if set to a non-zero value PaToHwill use the same coarsening algorithm during the V-cycles. If it is zero, PaToHwill automatically selects a coarsening algorithm for each V-cycle.
- `usebucket`: PaToHcan use both heap and bucket as priority queue. If this parameter always bucket is used, if it is 0 always heas is used, if it is -1 PaToH determines when to use heap using following parameters.
- `maxcellinheap`: Heap will not be used if the current hypergraph has more cell than this number.
- `heapchk_mul`: Heap will be used if $bs \times$ `heapchk_mul/heapchk_div` $< |\mathcal{V}_i|$ at the level i, where bs is required bucket size.
- `heapchk_div`: Heap will be used if $bs \times$ `heapchk_mul/heapchk_div` $< |\mathcal{V}_i|$ at the level i, where bs is required bucket size.
- `MemMul_CellNet`: PaToHallocates three large continues arrays to be able to run Multilevel Partitioning. The first array holds the internal cell and net structures. This parameter tells to PaToH to allocate `MemMul_CellNet` times much memory that is required to hold the cell and net structures of the original hypergraph.
- `MemMul_Pins`: The second large array is used to store net-lists of cells (nets array) and cell-lists of nets (pins array). This parameter tells to PaToH to allocate `MemMul_Pins` times much memory that is required to hold pins and nets arrays of the original hypergraph.
- `MemMul_General`: The last large array is used to store temporary working arrays required during the multilevel partitioning. This parameter tells to PaToH to allocate `MemMul_Pins` times much memory that is required to hold pins array of the original hypergraph.

4.3.2 Coarsening Parameters

- `crs_VisitOrder`: cell visit order for coarsening algorithms
  - `0`: Sequential,
  - `1`: Random,
  - `2`: Cell degree sorted,
  - `3`: Maximum net size sorted,
  - `4`: Minimum net size sorted,
  - `5`: Minimum of net size sum sorted,
  - `6`: Sweep.
- `crs_alg`: coarsening algorithm choices:

  In matching-based clustering schemes listed below (1—8), vertex u denotes the unmatched vertex visited according the order determined by the `crs_VisitOrder` parameter. Vertex u is the source of the current matching process and an unmatched vertex v is selected according to a criterion among all unmatched vertices adjacent to u. Recall that two vertices u and v are said to be adjacent if they share at least one net, i.e., $nets[u] \cap nets[v] \neq \emptyset$. Here, $\mathcal{N}_{uv}$ denotes the set of nets shared by vertices u and v, and $N_{uv} = |\mathcal{N}_{uv}|$ denotes the number of nets shared between u and v.

- 1: Heavy connectivity matching (HCM). Vertex v has *maximum* connectivity value N_{uv}.
- 2: Probabilistic heavy connectivity matching(PHCM).
- 3: Manhathan distance (MANDIS). Vertex v has *minimum* Manhattan Distance $M_{uv} = d_u + d_v - 2N_{uv}$.
- 4: Average distance (AVEDIS). Vertex v has *minimum* average Manhattan Distance $M_{uv}/(d_u - N_{uv})$.
- 5: Canberra metric (CANBERRA). Vertex v has *minimum* $M_{uv}/(d_u + d_v)$ ratio.
- 6: Absorption Matching (ABS). Vertex v has *maximum* sum $\sum_{n \in N_{uv}} 1/(s_n - 1)$. This similarity metric favors matching vertex pairs connected via nets of small sizes.
- 7: Greedy Cut Matching (GCM). Vertex v has *minimum* $d_u + d_v - N_{uv}$ value.
- 8: Scaled Heavy Connectivity Matching (SHCM). Vertex v has *maximum* $N_{uv}/(d_u + d_v - N_{uv})$ ratio.

In agglomerative clustering schemes listed below (9—15), vertex u denotes the unclustered vertex visited according the order determined by the `crs_VisitOrder` parameter. Vertex u is the source of the current clustering process and all vertices adjacent to vertex u are considered for selection according to a criterion. The selection of a vertex v adjacent to u corresponds to including vertex u to singleton or multinode cluster C_v that contains vertex v to grow a new multinode cluster $C_{uv} = \{u\} \cup C_v$.

- 9: Heavy Connectivity Clustering (HCC). This metric is the agglomerative version of HCM. That is, v has maximum N_{u,C_v}, which denotes the number of nets shared between vertex u abd cluster C_v.
- 10: Heavy Pin Clustering (HPC). Cluster v has maximum $\sum_{n \in \mathcal{N}_{u,C_v}} |pins[n] \cap C_v|$.
- 11: Absorption Clustering using Nets (ABSHCC), This metric is the agglomerative version of ABS. That is, v has maximum $\sum_{n \in N_{u,C_v}} 1/(s_n - 1)$.
- 12: Absorption Clustering using Pins (ABSHPC) ,
- 13: Connectivity Clustering (CONC),
- 14: Greedy Cut Clustering (GCC). This is agglomerative version of GCM. With this metric v_j is chosen to form a cluster with v_i that has minimum $d_i + d_j - N_{i,j}$.
- 15: Scaled Heavy Connectivity Clustering (SHCC),

In the net-based clustering algorithms listed below (16–17), nets are visited in random order and their pins are considered for clustering.

- 16: Net Clustering (NC). All pins of a net are clustered if none of them has been clustered yet.
- 17: Modified Net Clustering (MNC). All pins of a net that are not currently clustered gathered to form a cluster.

• `crs_coarsento`: limits the number of cells in the coarsest hypergraph.

• `crs_coarsentokmult`: Number of cells in the coarsest hypergraph is set to maximum of $K \times$ `crs_coarsentokmult`.

• `crs_coarsenper`: Stops coarsening when number of cells is not reduced by `crs_coarsenper`

• `crs_maxallowedcellwmult`: limits the construction of large cells. Maximum weight of a cell can be at most `crs_maxallowedcellwmult` $\times W_{avg}$.

- `crs_idenafter`: starting level of identical net detection in coarsening. Supplying negative values results in automatic computation of the parameter.

- `crs_iden_netsizetrh`: Threshold net size for identical net detection. Nets whose sizes are equal or less than this values will be checked.

- `crs_useafter`: Changes the coarsening algorithm after that level to `crs_useafteralg`.

- `crs_useafteralg`: Coarsening algorithm that will be used after level `crs_useafter`.

4.3.3 Initial Partitioning Parameter

- `nofinstances`: PaToH can refine multiple partitions during the uncoarsening phase. This parameter sets the number of partitioning instance to be constructed in initial partitioning phase. Each of these instances will be refined during the uncoarsening phase.

- `initp_alg`: Determines the initial partitioning algorithm, here is the list of the implemented algorithms:

  - `1`: Greedy Hypergraph Growing Partition (`GHGP`). In this algorithm we grow a cluster around a randomly selected vertex. During the coarse of the algorithm, the selected and unselected vertices induce a bipartitioning on the coarsest hypergraph. The unselected vertices connected to the growing cluster are inserted into a priority queue according to their FM gains. Here, the gain of an unselected vertex corresponds to the decrease in the cutsize of the current bipartition if the vertex moves to the growing cluster. Then, a vertex with the highest gain is selected from the priority queue. After a vertex moves to the growing cluster, the gains of its unselected adjacent vertices which are currently in the priority queue are updated and those not in the priority queue are inserted into the queue. This cluster growing operation continues until a predetermined bipartition balance criterion is reached. Since the coarsest graph is small, GHGP algorithm is run four times starting from different random vertices and select the best bipartition for refinement during the uncoarsening phase.
  - `2`: Agglomerative Match and Bin Packing (`AGG_MATCH`).
  - `3`: Bin Packing (`BINPACK`).
  - `4`: Breadth-First (`BF`).
  - `5`: A random initial partitioning (`RANDOM1`).
  - `6`: Another random initial partitiong (`RANDOM2`).
  - `7`: Yet another random initial partitioning (`RANDOM3`).
  - `8`: Greedy hypergraph Growing with Max Pin (`GHG_MAXPIN`).
  - `9`: GreedyHypergraph Growing with Max Net (`GHG_MAXNET`).
  - `10`: GreedyHypergraph Growing with Max only-Pos FM Gain (`GHG_MAXPOSGAIN`).
  - `11`: Component bin-pack and Greedy Hypergraph Growing Partition (`COMP_GHGP`).
  - `12`: Greedy Component bin-pack and Greedy Hypergraph Growing Partition (`GREEDY_COMP_GHGP`).
  - `13`: use one of the above at each instance (`ALL`).

- `initp_runno`: the number of runs for each instance's initial partition.

- `initp_ghg_trybalance`: if it is set to a non-zero value, PaToHtries to find better balanced partitions during greedy hypergraph partitioning.

- `initp_refalg`: refinement algorithm that will be used after each initial partitioning, please refer to `ref_alg` for a list of available algorithms.

4.3.4 Uncoarsening Parameters

- `ref_alg`: Determines the refinement algorithm that will be used during the uncoarsening. Current version of PaToH contains 18 KLFM-based refinement algorithms:
  * `0`: No refinement (`NONE`).
  * `1`: Fiduccia-Matheyses (`FM`)
  * `2`: FM with tight balance (`FM_TB`)
  * `3`: Boundary FM (`B_FM`)
  * `4`: BFM with tight balance (`BFM_TB`).
  * `5`: FM with dynamic locking (`FM_DL`).
  * `8`: FM with Krishnamurty's multilevel gain (`FM_MG`).
  * `9`: BFM with dynamic locking (`BFM_D`).
  * `10`: BFM with Krishnamurty's multilevel gain (`BFM_MG`).
  * `11`: Kernihgan-Lin (`KL`)
  * `12`: KL with dynamic locking (`KL_DL`).
  * `13`: Boundary Kernighan-Lin (`BKL`).
  * `14`: BKL with dynamic locking (`BKL_DL`).
  * `15`: FM and KL (`FMKL`).
  * `16`: BFM and BKL (`BFMKL`).
  * `17`: BFM and BKL with dynamic locking (`BFMKL_DL`).
  * `18`: FM and Kernighan-Lin with dynamic locking (`FMKL_DL`).
- `ref_useafter`: After that level of coarsening refinement algorithm `ref_useafter_alg` will be used.
- `ref_useafteralg`: Refinement algorithm that will be used after level `ref_useafter`.
- `ref_passcnt`: Limits the number of passes at each level of uncoarsening.
- `ref_maxnegmove`: Limits the number of consecutive negative-gain moves.
- `ref_maxnegmovemult`: Limits the number of consecutive negative-gain moves to `ref_maxnegmovemult` $\times|\mathcal{V}_i|$ at the i-th level of uncoarsening..
- `ref_dynamiclockcnt`: Limits the maximum number of moves of a cell in a pass. `ref_dynamiclockcnt` $=$ 1 results the clasic FM algorithm.
- `init_imbal`: imbalance ratio of the coarsest hypergraph, i.e., maximum part weight in the coarsest hypergraph can be at most $W_{avg} \times (1+$`init_imbal`$)$.
- `final_imbal`: imbalance ratio of the final partition.
- `fast_initbal_mult`: At the beginning of each uncoarsening level partition is forced to have maximum imbalance ratio of
  `fast_initbal_mult` $\times\varepsilon$, then selected refinement algorithm is executed.
- `init_sol_discard_mult`: At the coarsest level, instances which have `init_sol_discard_mult` times worse cutsize than the partition with minimum cutsize are discarded.
- `final_sol_discard_mult`: At the final partition, instances which have `init_sol_discard_mult` times worse cutsize than the partition with minimum cutsize are discarded. Note that, PaToHlinearly interpolates this discard multiplier at each level of uncoarsening.

5 Stand-Alone Program

Distribution includes a stand-alone program, called `patoh`, for single constraint partitioning [1]. `patoh` gets its paramaters from command line arguments. You can run the PaToH from command line as follows

```
> patoh <hypergraph-file> <number-of-parts>
  [[parameter1] [parameter2] ....]
```

You can tune the parameters using optional `[parameter]` arguments. The syntax of these optional parameters is as follows; two-letter abbreviation of a parameter is followed by an equal sign and a value. For example, parameter `ref_alg` is abbreviated as “RA” and to select “Boundary FM with dynamic locking” (3rd algorithm) use “RA=3”. For a complete example, lets say we have an hypergraph named “myhypergraph.txt”, and we want to partition this hypergraph into 4 parts using the Kernighan-Lin refinement algorithm with cutsize metric. Command

```
> patoh myhypergraph.txt 4 UM=U RA=6
> patoh myhypergraph.txt 4 PQ=Q RA=6 II=0.004 FI=0.003 OD=3
```

will partition the hypergraph using default parameter settings optimized for high quality partitions. Additionally initial and final imbalance ratios are set to 0.4% and 0.3%, respectively, and a detailed output of recursive bisection will be displayed.

```
+++++++++++++++++++++++++++++++++++++++++++++++++++++++
+++ PaToH v2.5 (c) Sep 1999, by  Umit V. Catalyurek +++
+++++++++++++++++++++++++++++++++++++++++++++++++++++++

********************************************************************************
Hypergraph : /home/work/ubora/lpComputational/NL.rn
     #Cells :  9718   #Nets :  7039   #Pins : 41428
********************************************************************************
 4-way partitioning results of PaToH:

 'Con - 1' Cost:   712
 Cell Weights: Min= 2420 (0.004) Max= 2442 (0.005)
 Net Weights : Min= 1579 (0.045) Max= 1803 (0.090)
---- BALANCE IS NOT TIGHT ENOUGH ----
--------------------------------------------------------------------------------
I/O          :       0.160
I.Perm/Cons.H:       0.070  ( 2.7%)
Coarsening   :       1.740  (66.9%)
Partitioning :       0.070  ( 2.7%)
Uncoarsening :       0.620  (23.8%)
Split        :       0.080  ( 3.1%)
Total        :       2.600
Total (w I/O):       2.760
--------------------------------------------------------------------------------
```

Tables 1 and 2 display the command-line abbreviation of each parameter, the value-types of the parameters and the valid ranges of the values.

[1] Please note that this executable will not work with multiple vertex weights. For multi-constraint partioning use the provided library interface.

Table 1: Stand-alone program parameters

Parameter	Abbreviation	Type	Range
Miscellaneous Parameters			
`outputdetail`	OD	i	0, 1, 2, 3
`seed`	SD	i	-1: random, otherwise sets seed
`doinitperm`	DP	i	0, 1
`bisec_fixednetsizetrsh`	FS	i	[1-
`bisec_netsizetrsh`	NT	f	[0.5-
`bisec_partmultnetsizetrsh`	NM	i	[1-
`bigVcycle`	BV	i	[1-
`smallVcycle`	SV	i	[1-
`usesamematchinginVcycles`	UM	i	0, 1
`usebucket`	UB	i	-1, 0, 1
`maxcellinheap`	HC	i	[0-
`heapchk_mul`	HM	i	[1-
`heapchk_div`	HD	i	[1-
`MemMul_CellNet`	A0	i	[1-
`MemMul_Pins`	A1	i	[1-
`MemMul_General`	A2	i	[1-
Coarsening Parameters			
`crs_VisitOrder`	VO	i	[0-6]
`crs_alg`	MT	i	[1-17]
`crs_coarsento`	CT	i	[10-
`crs_coarsentokmult`	CK	i	[1-
`crs_coarsenper`	CP	i	[1-100]
`crs_maxallowedcellwmult`	CM	f	[0.01-1.0]
`crs_idenafter`	ID	i	[-1 -
`crs_iden_netsizetrh`	IT	i	[2-
`crs_useafter`	FL	i	[0-
`crs_useafteralg`	FM	i	[1-17]
Initial Partitioning Parameter			
`nofinstances`	NI	i	[1-
`initp_alg`	PA	i	[1-13]
`initp_runno`	IR	i	[1-
`initp_ghg_trybalance`	TB	i	0, 1
`initp_refalg`	IA	i	[0-10]

Table 2: Stand-alone program parameters (continued)

Parameter	Abbreviation	Type	Range
Uncoarsening Parameters			
`ref_alg`	RA	i	[0-10]
`ref_useafter`	RL	i	[0-
`ref_useafteralg`	RF	i	[0-10]
`ref_passcnt`	RP	i	[1-
`ref_maxnegmove`	RN	i	[5-
`ref_maxnegmovemult`	RU	f	[0.0001-1.0]
`ref_dynamiclockcnt`	LC	i	-1, 1, 2, 3
`init_imbal`	II	f	[0.00-0.50]
`final_imbal`	FI	f	[0.00-0.50]
`fast_initbal_mult`	FB	f	[0.5-2.0]
`init_sol_discard_mult`	DI	f	[0.01-1.00]
`final_sol_discard_mult`	DF	f	[0.01-1.00]
Parameter for Stand-alone program			
total # of runs	NR	i	[1..

5.1 Input File Format

The input hypergraph $\mathcal{H}=(\mathcal{V},\mathcal{N})$ is stored in a plain text file. The first line after the possible comment lines describes the size of the hypergraph, the indexing method used at the rest of the file and the weighting scheme. The rest of the file contains information for each net, and possibly for each vertex–depending on the weighting scheme. Any line beginning with '%' is a comment line and skipped.

The first line contains 4, optionally 6 integers. The first one–either 1, or 0– is the numbering used in indexing the elements of $\mathcal{V}$ and $\mathcal{N}$. Next the sizes of the sets $|\mathcal{V}|, |\mathcal{N}|$, and *pins* should be present. The fifth integer is optional and describes the weighting scheme of the hypergraph, if present. The hypergraph can have weights associated with cells, nets, or both, 1,2,3 respectively. The sixth integer also optional and denotes the number of constraints, in other words number of weights for each cells. If it is omitted it is assumed to be 1.

The next $|\mathcal{N}|$ lines contain the information about the nets. i^{th} line (excluding the comments) contains the cell list of net n_i. In the case of the weighted nets, each line begins with an integer representing the weight of n_i.

If cells are weighted, following the net lines, each cell's weight must be supplied. If there are more than one weight constraints then each cell should have number of constraint weights.

Input file corresponding to the sample hypergraph of Section 4.1 is displayed in Figure 4.

5.2 Output File Format

The output file corresponding to K-way partitioning of $\mathcal{H} = (\mathcal{V},\mathcal{N})$ contains $|\mathcal{V}|$ integers in range $[0,\ldots,K-1]$. i^{th} entry in this file represents the part number whom the cell i is assigned to.

References

[1] C. J. Alpert and A. B. Kahng. Recent directions in netlist partitioning: A survey. *VLSI Journal*, 19(1–2):1–81, 1995.

[2] C. Aykanat, A. Pınar, and Ü. V. Çatalyürek. Permuting sparse rectangular matrices into singly-bordered block-diagonal form for parallel solution of lp problems. *submitted for publication.*

```
1 8 9 28
8 6 3 5 2
4 5 1 7
4 2 5 7
4 7
3 5
8 2 4
6 5 2
5 7 2
8 4
```

(a)

```
1 8 9 28 2
10 8 6 3 5 2
15 4 5 1 7
13 4 2 5 7
18 4 7
25 3 5
20 8 2 4
14 6 5 2
27 5 7 2
29 8 4
```

(b)

```
1 8 9 28 1
8 6 3 5 2
4 5 1 7
4 2 5 7
4 7
3 5
8 2 4
6 5 2
5 7 2
8 4
80 85 30 55 42 39 90 102
```

(c)

```
1 8 9 28 3
10 8 6 3 5 2
15 4 5 1 7
13 4 2 5 7
18 4 7
25 3 5
20 8 2 4
14 6 5 2
27 5 7 2
29 8 4
80 85 30 55 42 39 90 102
```

(d)

Figure 4: (a) Hypergraph file without weights, (b) Hypergraph file with net weights, (c) Hypergraph file with cell weights, (d) Hypergraph file with weights on nets and cells

[3] T. N. Bui and C. Jones. A heuristic for reducing fill in sparse matrix factorization. In *Proc. 6th SIAM Conf. Parallel Processing for Scientific Computing*, pages 445–452, 1993.

[4] Ü. V. Çatalyürek. *Hypergraph Models for Sparse Matrix Partitioning and Reordering.* PhD thesis, Bilkent University, Computer Engineering and Information Science, Nov 1999.

[5] Ü. V. Çatalyürek and C. Aykanat. Hypergraph-partitioning based decomposition for parallel sparse-matrix vector multiplication. *IEEE Transactions on Parallel and Distributed Systems*, 10(7):673–693, 1999.

[6] C.-K. Cheng and Y.-C. Wei. An improved two-way partitioning algorithm with stable performance. *IEEE Transactions on Computer-Aided Design*, 10(12):1502–1511, December 1991.

[7] A. Dasdan and C. Aykanat. Two novel multiway circuit partitioning algorithms using relaxed locking. *IEEE Transactions on Computer-Aided Design*, 16(2):169–178, February 1997.

[8] C. M. Fiduccia and R. M. Mattheyses. A linear-time heuristic for improving network partitions. In *Proceedings of the 19th ACM/IEEE Design Automation Conference*, pages 175–181, 1982.

[9] A. Gupta. Fast and effective algorithms for graph partitioning and sparse matrix ordering. Technical Report RC 20453, IBM T. J. Watson Research Center, Yorktown Heights, NY, 1996.

[10] A. Gupta. Watson graph partitioning package. Technical Report RC 20453, IBM T. J. Watson Research Center, Yorktown Heights, NY, 1996.

[11] B. Hendrickson and R. Leland. A multilevel algorithm for partitioning graphs. Technical report, Sandia National Laboratories, 1993.

[12] B. Hendrickson and R. Leland. *The Chaco user's guide, version 2.0.* Sandia National Laboratories, Alburquerque, NM, 87185, 1995.

[13] B. Hendrickson and E. Rothberg. Effective sparse matrix ordering: just around the bend. In *Proc. Eighth SIAM Conf. Parallel Processing for Scientific Computing.*

[14] A.G Hoffmann. Dynamic locking heuristic— a new graph partitioning algorithm. In *Proceedings of IEEE International Symposium on Circuits and Systems*, pages 173–176, 1994.

[15] G. Karypis and V. Kumar. *MeTiS A Software Package for Partitioning Unstructured Graphs, Partitioning Meshes, and Computing Fill-Reducing Orderings of Sparse Matrices Version 3.0.* University of Minnesota, Department of Comp. Sci. and Eng., Army HPC Research Center, Minneapolis, 1998.

[16] G. Karypis and V. Kumar. A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM Journal on Scientific Computing*, to appear.

[17] B. W. Kernighan and S. Lin. An efficient heuristic procedure for partitioning graphs. *The Bell System Technical Journal*, 49(2):291–307, February 1970.

[18] B. Krishnamurthy. An improved min-cut algorithm for partitioning VLSI networks. *IEEE Transactions on Computers*, 33(5):438–446, May 1984.

[19] T. Lengauer. *Combinatorial Algorithms for Integrated Circuit Layout.* Willey–Teubner, Chichester, U.K., 199.

[20] D. G. Schweikert and B. W. Kernighan. A proper model for the partitioning of electrical circuits. In *Proceedings of the 9th ACM/IEEE Design Automation Conference*, pages 57–62, 1972.

[21] Y.-C. Wei and C.-K. Cheng. Ratio cut partitioning for hierarchical designs. *IEEE Transactions on Computer-Aided Design*, 10(7):911–921, July 1991.