Optimal Algorithms for Structural Assembly

Danny Dolev, Kevin Karplus, Alan Siegel, Alex Strong, Jeffrey D. Ullman, Stanford University

In many VLSI design systems (Johannsen 1979), a common problem is to connect terminals (considered as points on the facing sides of two adjacent rectangles) in a designated order. We shall generally assume that there are no crossovers—*i.e.*, that the order is the same for both rectangles. Wire placement can be optimized by putting the rectangles as close to each other as possible. One seemingly promising way to minimize the separation is to use several layers for the wiring. Two- or three-layer wirings clearly are more general, and are necessary when the terminals of the two rectangles need not be connected to order. However, single-layer wiring is often used, because of the area lost when wires are run between two or more layers.

Figure 1 shows the wiring problem: to draw a wire from each P_i (a terminal on the lower row) to the corresponding Q_i (a terminal on the upper row), such that no two wires ever come within one unit of each other. (Of course, the actual size of a "unit" in terms of λ is determined by the specific design rules for the minimum width and separation of the particular layer used in the connection.) To meet the wire spacing requirement, the P_is must be placed at least one unit apart from each other, as are the Q_is. The wire width is set to zero whenever this simplification causes no material change in the algorithms. More detailed modeling is included where necessary. We also assume that wires must be vertical within one unit of the upper and lower rows, as shown in Figure 1. This assumption reflects the fact that the rectangles may contain unknown layers, which must be a unit distance from the routing wires. We consider principally the following two wiring disciplines:

1. *Wires can travel in any direction.* We call this unrestricted routing scheme the *general case.* It was considered by Tompa (1980), and an optimal wiring was obtained for the situation in which the relative position of the rectangles is fixed both horizontally and vertically, or is fixed only horizontally.
2. *Wires can travel only horizontally or vertically.* We call this the *rectilinear case.* It has been considered by Valiant (1979), Fischer and Paterson (1980), and Storer (1980), and is based on common design restrictions. To avoid an unduly complex description of the results, we assume in this case that the plane is a unit grid, as in Valiant (1979), and that wires can travel only along grid lines. Furthermore, the terminals are placed only on grid lines, although not necessarily on adjacent lines. If there are two or more layers, we assume that terminals can be connected between any layers, and that any grid line can "carry" a wire on each layer.

A version of this article appeared in the copyrighted Proceedings of the 13th ACM Symposium on the Theory of Computing, May 1, 1981.

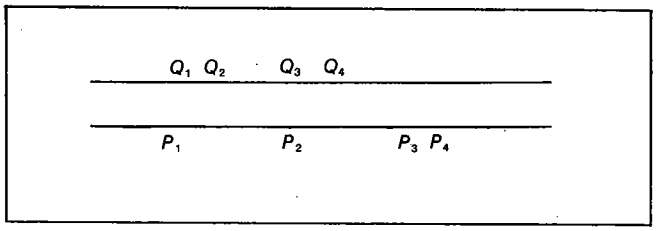

FIGURE 1. Basic wiring problem.

Much of what we say holds for a larger class of constraints on how wires may travel. For example, we can generalize the rectilinear case so that wires also travel along the diagonals of the squares of the grid. This model also reflects the actual constraints of some mask-making facilities.

We call the horizontal displacement of Q_1 and P_1 in Figure 1 the *offset,* and we refer to the number of tracks between the two rows of terminals as the *separation.* Thus, the separation is simply the physical distance, minus one. When non-rectilinear wiring schemes are used, the separation need not be an integer; in such cases, it has meaning only in terms of distance, not tracks. Consider the following optimization problem: given a fixed offset and a wiring discipline (*e.g.*, rectilinear), what is the minimum *separation problem.* Consider also the *offset problem*: for which offset is separation as small as possible?

The Rectilinear Model Separation Problem

In actual VLSI designs, it is often necessary to connect a set of terminals on a row (*e.g.*, inputs from pads) with another set of terminals on a different row (*e.g.*, input latches). The spacings between the terminals may be irregular. We would like to determine how close we can bring these rows of terminals to each other, while still being able to connect them legally. In addition to determining quickly this minimum separation, we would also like to have an algorithm for actually laying out the wires.

Given a fixed offset, and n pairs of terminals $(P_1, Q_1), \ldots, (P_n, Q_n)$, we take P_i to be both the name of a terminal on the bottom row and the horizontal position of that terminal; and we take Q_i to be a similar terminal on the upper row. Our goal is to find the minimum separation for which a legal wiring exists. Let us define a *right block* as a maximal sequence of pairs of terminals $(P_i, Q_i), \ldots, (P_j, Q_j)$ such that for $i \leq k \leq j$, $Q_k \geq P_k$. That is, all the connections in the block have a position in the upper row to the right of the corresponding position in the lower row. We define a *left block* in the obvious, symmetrical way. We call a left or right block a *block*.

A block can always be wired so that all wires in a right block

move monotonically to the right, and so that all wires in a left block move monotonically to the left; *i.e.*, a wire need never reverse its direction. Thus, wiring in a block never extends past the boundaries of the block; therefore, the blocks are wiring-independent. This fact implies that for a fixed offset, the separation is determined by the "worst" block.

Two important constraints force large separations. First, there is the *channel density* at any horizontal position h. This quantity is the number of values of i for which $P_i \leq h \leq Q_i$ or $P_i \geq h \geq Q_i$, but not $P_i = h = Q_i$. The second constraint is the *conflict number* for any two pairs of terminals, i and j, which we denote by $W(i,j)$, as defined in Table 1.

We can understand intuitively that the channel density indicates how many wires must cross the vertical line at position h. The *conflict number* for i and j is intended to measure the number of wires that must cross an imaginary line from P_j to Q_i. (This concept is an adaptation of the basic idea of Tompa (1980) to the rectilinear case.) The reason for the strange form of this number is that the wires may pass either horizontally or vertically between P_j and Q_i. If Q_i and P_j are sufficiently far apart horizontally, then the wires could pass vertically; the separation could be as little as 0. If Q_i and Q_j are close together horizontally, then the wires cannot all pass vertically; in this case, in the horizontal direction, one track per wire will be needed.

Example 1. Figure 2 shows an interesting case in which n pairs of terminals are offset by one unit. At any interior horizontal position, the channel density is 2. However, $P_n - Q_1 = n-2$, which is less than $n-1$. Thus, $W(1,n) = n$, from which we conclude that separation of n is required.

The channel density and the conflict number affect the number of tracks needed through a given line. These two notions are related; in fact, it can be shown that the largest channel density in a problem is never greater than the largest conflict number. We can also prove that the separation for any one-layer rectilinear connection is at least equal to the largest conflict number.

A very simple and useful algorithm in actual wiring is the following "greedy" algorithm. To wire a right block, we lay down wires sequentially, starting at the leftmost pair of terminals. When wiring a given pair, we move vertically, as long as we can. When we cannot proceed any farther in this direction, we move to the right, but revert to moving vertically as soon as we can do so legally. Left-block wiring follows the analogous procedure.

The relationships between the conflict numbers and channel densities, together with the "greedy" algorithm, let us prove the following results.

Theorem 1: In any rectilinear separation problem, tracks equal in number to the largest conflict number are necessary and sufficient for one-layer wiring.

Example 2: Figure 2 shows an application of the greedy algorithm.

The greedy algorithm was programmed by K. Karplus and A. Strong at Stanford University, for the layout of the "Digitar" chip. The algorithm is also part of the CHISEL design system being implemented for circuit layout. The program keeps a list of the x- and y-coordinates of the most recently routed wire, and uses this list to route the next wire, while simultaneously updating the list. The algorithm can also handle certain cases in which the terminals are not all on straight lines. An interesting characteristic of the greedy algorithm is that in the rectilinear case, the total wire length is minimal. This is a desirable feature, because the diffusion delay of a wire is proportional to the square of the length of the wire. This wiring scheme may take as much as $O(n^2)$ time to run, because the wires can have $O(n^2)$ "corners." The separation problem, however, can be solved more efficiently.

	$i<j$	$i=j$	$i>j$
$P_j-Q_i>j-i$	0	1	$i-j+1$
$P_j-Q_i=j-i$	$j-i+1$	0	$i-j+1$
$P_j-Q_i<j-i$	$j-i+1$	1	0

TABLE 1. Definition of $W(i,j)$.

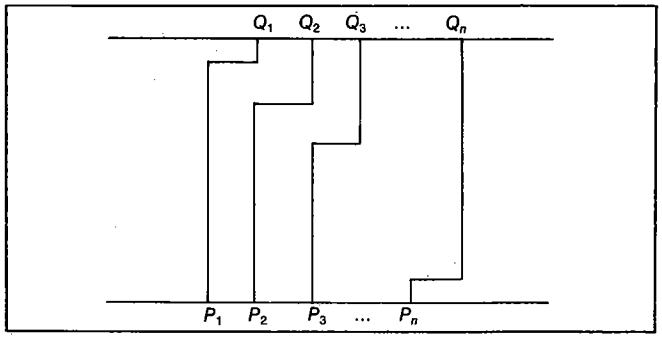

FIGURE 2. Large conflict number and small channel density.

Corollary 1: The number of tracks needed for the rectilinear separation problem can be determined in $O(n)$ time.

Proof: Normalize the problem so that all blocks are right blocks, by redefining $P_j = min(P_j, Q_j)$ and $Q_j = max(P_j, Q_j)$. This alteration doesn't change the number of tracks needed. Let c_j be the smallest $i \leq j$ such that $P_j - Q_i < j-i$ (or j if no such i). Because $W(c_j,j) = max_{i\in[1,j]}W(i,j)$, and $c_j \leq c_{j+1}$, we can find the maximum conflict number by incrementing j from 1 to n, searching for c_j (starting at c_{j-1}) and computing $W(c_j,j)$.

Multilayer Solutions

Two- and three-layer problems yield an entirely different sort of result. A solution to the three-layer rectilinear separation problem requres a number of tracks equal to one third the largest channel density. A two-layer solution requires a number of tracks equal to one-half the largest channel density. We cannot prove that this lower bound is always achievable. One can certainly achieve a wiring with a number of tracks equal to one-third the largest conflict number; but, as shown in Example 1, that number can be asymptotically much greater than one-third the largest channel density. However, by another sort of "greedy" algorithm, we can achieve a three-layer (or even two-layer) wiring proportional to the lower bound.

Theorem 2: A three-layer solution to the rectilinear separation problem exists, that uses no more tracks than the largest channel density. Furthermore, this solution can be implemented in two layers, if the grid size represents enough physical area for two wires to be run in the same layer.

Proof: First, as in Theorem 1, break the pairs of terminals into blocks. All wires in a block run first vertically in one layer, then horizontally in the second layer, and finally vertically again in the third layer. To wire a right block, select tracks from left to right. (To wire a left block, select tracks from right to left.) In either case, pick the lowest available track for the horizontal wire.

If $P_j = Q_i$, and the j^{th} pair uses a higher track than the i^{th}, then it looks as though two vertical wires occupy the same place. Avoid this problem by making the grid spacing large enough to handle two wires. (In practice, we can use a third layer to obtain a some-

what smaller grid spacing, without creating any transistors.)
Example 3: Figure 3 shows a wiring of the problem in Figure 2, using the greedy algorithm of Theorem 2.

Some new two-layer wiring algorithms do not require a larger grid size. For these schemes, $2h$-1 (*i.e.*, one less than twice the channel density) is sufficient (Rivest *et al.* 1981) and sometimes necessary (Leighton 1981) for routing.

One-Layer and Two-Layer Solutions

From Figures 2 and 3, one may assume that two-layer wirings can be arbitrarily better than one-layer wirings. However, we believe that in practice, this is not the case. We need a preliminary result indicating that as long as terminals are not packed too tightly along the rows, then a one-layer solution always exists, whose separation is proportional to the channel density, not to the conflict number.

Theorem 3: Let x be the maximum channel density. Suppose a constant $\alpha > 0$ such that for every $r \geq x(1+1/\alpha)$, no r consecutive grid lines have more than $r/(1+\alpha)$ terminals. Then, there always exists a one-layer solution to the rectilinear separation problem using no more than $x(1+1/\alpha)$ tracks.

Proof: (sketch) Figure 4 shows the crucial case. Note that k is the channel density at Q_i (thus $k \leq x$) and $k + m$ is the conflict number $W(i,j)$ if the latter is not 0. We must prove that $W(i,j)$ is either proportional to x or is 0, thus showing that the maximum conflict number and the channel density are proportional. If $m \leq x/\alpha$, then $W(i,j)$ is no larger than $x(1 + 1/\alpha)$. If $m > x/\alpha$, then reason as follows:

For $W(i,j)$ not to be 0, we must have $k + m \geq P_j - Q_i$. But by our assumption about the sparseness of terminals, we can show that $m \leq (P_j - Q_i)/(1 + \alpha)$. It follows from these two inequalities that $k + m \geq m(1 + \alpha)$, or $m \leq k/\alpha$. Because we assumed $m > x/\alpha \geq k/\alpha$, we see that $W(i,j)$ must be 0.

Corollary 2: If the minimal distance between two neighboring terminals is $1 + \alpha$, for $\alpha \geq 0$, then a one-layer solution to the rectilinear separation problem always exists using no more than $\min(n, x(1 + 1/\alpha))$ tracks, where n is the number of pairs to be connected.

Even when the one-layer solution is "worse," it still has some advantages in actual routing. We can implement the one-layer solution with either polysilicon or diffusion, and then do other independent routings simultaneously with the third (metal) layer.

There is a qualitative difference between the situation in which terminals are not as densely packed as the grid lines (however close to 1 the ratio may be), and the case in which terminals are packed one to a grid line. In the former case, the separation depends only on the channel density; in the latter case, it depends on the conflict number.

Now we can apply Theorem 3 to particular values of α that reflect real design rules. (We use here the design rules of Mead and Conway (1980).) First, in terms of λ (the fundamental unit for design rules), we can run a single-layer wiring in polysilicon with grid units equal to 4λ. (In fact, we can use a 3λ grid if we alternate polysilicon and diffusion; but the capacitance inherent in diffusion wires provides good reason not to do so.) However, consider the grid size for a two-layer wiring. If, according to Theorem 2, the wires run vertically in polysilicon, horizontally in metal, then vertically in polysilicon, we have several choices, none very good. (These wires must be diffusion if the horizontal grid is less than 10λ; certain other combinations of horizontal and vertical grid sizes also require that this wire be diffusion.)

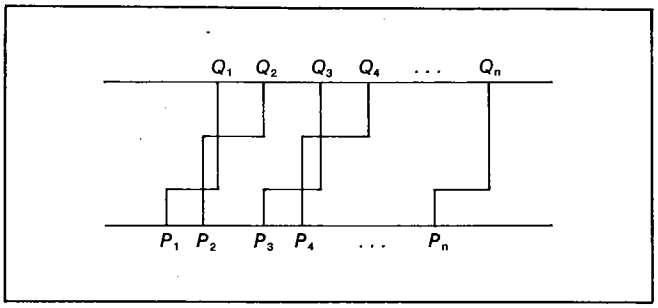

FIGURE 3. Two-layer "greedy" wiring.

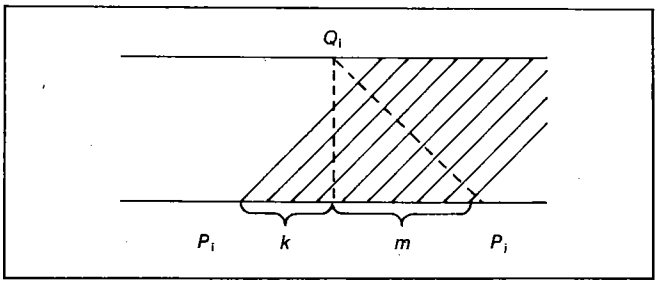

FIGURE 4. Diagram for Theorem 5.

Example 4: We could use a horizontal grid of 10λ and a vertical grid of 7λ. That is, terminals that must be separated by 10λ. But if this is the case, we could use a 4λ grid and a one-layer wiring, and claim that $\alpha = 3/2$. Then, by Theorem 3, a wiring exists with no more tracks than 5/3 the maximum channel density. If x is that maximum, then, in terms of λ, the separation in the one-layer is $6.67\lambda x$; for the two-layer case, with its 7λ vertical grid, we need $7\lambda x$ separation, which is greater.

We could also use a three-layer wiring with an 8λ horizontal grid and 7λ vertical grid, which provides slightly better separation than does the one-layer solution. Other grid sizes for the two-layer case (such as a 7λ horizontal grid coupled with a 14λ vertical grid) yield the same conclusion: in practice, one-layer wirings are as good or better than two- or three-layer wirings. The only way to escape this conclusion would be to improve substantially the upper bound of Theorem 2, or to discover better plans for three-layer wirings. (Determining the set of pairs h and v for which two- and three-layer wirings exist for an $h\lambda$ by $v\lambda$ grid is an open problem.)

Other Results

For Theorem 4: When wires can run along the eight compass directions, the separation problem can be solved in $O(n)$ time.

Proof: (sketch) The conflict number can be broken into two parts: one resulting from the rectilinear (flat) portions of the wiring, and one resulting from a restriction to the 45-degree pieces. As in the rectilinear case, the maximum contribution from the flat portions can be found from a linear scan. The other contribution can also be evaluated by a linear scan. In this instance, a priority queue must be maintained to indicate which P_j connection point gives a maximum (restricted) conflict for a current point Q_i. When i is incremented a new restricted $W(i,j)$ value is computed, and the data structure is updated. As j is incremented, P_j is inserted into the data structure. The linearity results from the fact that during the insertion, old data (that contributes a conflict less than that from a new P_j entry) can be deleted. Furthermore, during the updating, P_h entries giving former (but not current) maximum conflict values can be deleted.

For Theorem 5: For the general wiring case, the separation problem can be solved in $O(n\log n)$ time (Siegel and Dolev 1981).

Proof: (sketch) We first find a P_{i_0} maximizing $W(i, \frac{n}{2})$, the separation induced by the pairs P_i and $Q_{\frac{n}{2}}$. It turns out (Siegel and Dolev 1981) that the maximum separation is among separations restricted to the intervals $[P_1, P_{i_0}]$ and $[Q_1, Q_{\frac{n}{2}}]$, or $[P_{i_0}, P_n]$ and $[Q_{\frac{n}{2}}, Q_n]$. Repeating the divide and-conquer step on the Q coordinates requires a total of $O(n\log n)$ comparisons.

Although we obtain the same separation as Tompa (1980) in these two cases, our total wire length is not normally minimized.

Another problem is that of "fluid" terminals. Assume we can decide where to put the terminals on the row so as to minimize the separation. The interesting case is when the terminals on one row (say, the upper) are fixed, and we have to determine the placement of the terminals on the lower row. Using our earlier results, we can achieve the minimal separation by packing consecutive terminals as closely as possible, on the left-hand side of the row (as long as they belong to a left block), and on the right hand side of the row (as long as they belong to a right block). The rest of the terminals (if any) can be placed just across from the corresponding terminals. The algorithm minimizes the largest block; therefore, it produces the minimal channel density, which lets us achieve a minimal separation. The last result can be extended so that we can simultaneously minimize the size of the occupied part of the row we construct.

In the offset problem, the relative positions of the terminals in each row are fixed; but we can slide the rows relative to one another. The problem is to find the offset that minimizes the separation.

Theorem 6: For the rectilinear, general, and 8-compass-point wire models, the offset problem can be solved in time $O(T(n)\log n)$, where $T(n)$ is the time to find the separation (Dolev *et al.*, 1981; Siegel and Dolev 1981).

Some more general algorithms for river-routing multiple modules across a channel can be found in Leiserson and Pinter (1981).

A variety of new routing results have appeared recently. Brown and Rivest (1981) examine a two-layer model that minimizes cross-coupling capacitances. This model also permits non monotone wires, to reduce the separation. LaPaugh (1980) analyzes two-layer routing of a rectangular module, and Pinter (1981) analyzes the problem of track assignment for wires that must change channels.

Acknowledgments

The authors' work was supported by Hertz and Weizmann fellowships, by NSF grant DCS-80-12907, and by DARPA contract MDA-903-80-C-107.

References

Brown, D.J. and R.L. Rivest. October 1981. "New Lower Bounds for Channel Width." *Carnegie-Mellon Conference on VLSI Systems and Computations.*

Dolev, D., K. Karplus, A. Siegel, A. Strong and J.D. Ullman. 1981. "Optimal Wiring Between Rectangles." *Proceedings of the 13th Annual ACM Symposium on the Theory of Computing.*

Fisher, M.J. and M.S. Paterson. 1980. "Optimal tree layout." *Proc. of the 12th Annual ACM Symposium on the Theory of Computing.*

Johannsen, D. January 1979. "Bristle blocks: a silicon compiler." *Caltech Conference on VLSI.* See also *Sixteenth Design Automation Proceedings,* June 1979.

LaPaugh, A.S. 1980. "A polynominal time algorithm for optimal routing around a rectangle." *Proceedings of the Twenty-first Annual IEEE Symposium on Foundations of Computer Science.*

Leighton, F.T. 1981. "New Lower Bound Techniques for VLSI." *Proceedings of the Twenty-second Annual IEEE Symposium on Foundations of Computer Science.*

Leiserson, C.E. and R.Y. Pinter. October 1981. "Optimal Placement for River Routing." *Carnegie-Mellon Conference on VLSI Systems and Computations.*

Mead, C.A. and L.A. Conway. 1980. *Introduction to VLSI Systems.* Reading, MA: Addison-Wesley.

Pinter, R.Y. October 1981. "Optimal Routing in Rectilinear Channels." *Carnegie-Mellon Conference on VLSI Systems and Computations.*

Rivest, R.L., A.E. Baratz and G. Miller. October 1981. "Provably Good Channel Routing Algorithms." *Carnegie-Mellon Conference on VLSI Systems and Computations.*

Siegel, A. and D. Dolev. October 1981. "The Separation for General Single-layer Wiring Barriers." *Carnegie-Mellon Conference on VLSI Systems and Computations.*

Siegel, A. and D. Dolev. 1981. "Some Geometry for General River Routing." in *VLSI Systems and Computations,* editors H.T. Kung, B. Sproull, and G. Steele, Rockville, MD:Computer Science Press.

Storer, J.A. 1980. "The node cost measure for embedding graphs on the planar grid." *Proceedings of the Twelfth Annual ACM Symposium on the Theory of Computing.*

Tompa, M. 1980. "An optimal solution to a wire-routing problem." *Proceedings of the Twelfth Annual ACM Symposium on the Theory of Computing.*

Valiant, L. 1979. "Universality Considerations in VLSI Circuits." Memorandum, University of Warwick, Computer Science Department, Warwick, England.

About the Authors

Danny Dolev received the BS degree in physics from the Hebrew University, Jerusalem (1971) and the MS and PhD degrees in applied mathematics and computer science (1973 and 1979) from the Weizmann Institute of Science, Rehovot, Israel. He has six-years' experience as a systems analyst and applications programmer, and is now a post-doctoral scholar at Stanford.

Kevin Karplus received the BS degree in mathematics from Michigan State University (1976) and is currently completing a PhD in computer science at Stanford University. His research interests include VLSI design tools, chip design, analysis of alogorithms, and digital music synthesis.

Alan Siegel is a Ph.D. candidate in the computer science department at Stanford University. He holds a MS degree from New York University and a BS degree from Stanford. He has published papers on river routing and on classical mathematics.

Alex Strong received a BA in mathematics and physics from Wesleyan Univ. (1974). He is currently working toward a PhD in computer science at Stanford University. His research interests include VLSI design and design tools, digital audio synthesis, and microprocessor architecture.

Jeffrey D. Ullman graduated from Columbia in 1963 and received the PhD from Princeton in 1966. He spent three years at Bell Laboratories, and was on the faculty of Princeton University from 1969–1979, before joining the faculty at Stanford University.