Semigroup Forum Vol. ?? (1991) ??–??

Ordering D-Classes and Computing Schein Rank is Hard

George Markowsky

Communicated by B. M. Schein

1. The Main Result

This paper demonstrates that deciding whether a principal ideal in the semigroup of binary relations contains another principal ideal is *NP*-complete. This implies that no efficient algorithm for solving this problem is likely to be found. By contrast the problem of deciding whether one left (right) ideal contains another is easy to solve efficiently. The proof also shows that determining the Schein rank of a binary matrix is not likely to be solved by an efficient algorithm.

Let $\mathcal{B}_n$ be the set of all $n \times n$ Boolean matrices with the usual Boolean matrix operations. If $A, B \epsilon \mathcal{B}_n$ it is easy to determine whether there exist $X \epsilon \mathcal{B}_n$ such that $B = XA$ or if there exists Y such that $B = AY$. In the first case, for each row of B simply take all the rows of A that are componentwise $\leq$ that row and take their union. Either you get the row of B or that row cannot be generated using A. If you succeed for each row of B, then $B = XA$ has a solution, while if you fail there is no solution. The matrix X has $X_{ij} = 1$ iff row j of A (A_{j*}) is componentwise $\leq$ row i of B (B_{i*}). For the equation $B = AY$ use the columns of A and B. Alternatively, take the transpose of both sides and use rows.

The previous result suggests that determining whether there exist X and Y such that $B = XAY$ might be simple. Unfortunately, as we will show, this is not the case since if you could solve this problem easily, you could solve the set basis problem which Larry Stockmeyer [St] has shown is *NP*-complete. *NP*-complete problems are the hardest problems for which one can verify that solutions are correct in polynomial time. An efficient solution of an *NP*-complete problem would provide efficient solutions to many problems that have been of interest for a long time. This includes problems such as finding Hamiltonian paths and graph coloring. A complete discussion of *NP*-complete problems and related problems is in [GJ]. While no one has been able to prove that there is no efficient algorithm for solving an *NP*-complete problem, most researchers in the field of algorithms believe that if efficient algorithms existed they would have been found in the more than 100 years that people have worked on this class of problems. The result in this paper shows that it is likely that deciding whether a solution of $B = XAY$ exists is "intrinsically" hard. We will show that it is hard even when A has a very simple form.

The set basis problem is the following. Given a set U of cardinality q and nonempty subsets $S_1, ..., S_m$ of U, decide whether for a given integer k we can find c nonempty subsets $T_1, ..., T_c$ of U with $c \leq k$ such that each S_i is a union of some of the T_i's. We may assume that $k < m$ since the S_i are a solution in the case $k \geq m$.

Given a set basis of the form described in the preceding paragraph, let $n = \max(m, q)$. Let B be the $n \times n$ matrix given by

$$B_{ij} = \begin{cases} S_i[j] & \text{for all } i \leq m \text{ and } j \leq q \\ 0 & \text{otherwise} \end{cases}$$

where $S_i[j]$ is the j-th element of S_i. Let A be the $k \times k$ identity matrix, I_k, in $\mathcal{B}_n$, i.e.,

$$A_{ij} = \begin{cases} 1 & \text{for all } 1 \leq i = j \leq k \\ 0 & \text{otherwise} \end{cases}$$

We now show that the original set basis problem has a solution if and only if there exist $n \times n$ matrices X and Y such that $B = XAY$.

First, suppose that there are X and Y such that $XAY = B$. Note that at most the first k rows of the product AY are non-zero since $A_{ij} = 0$ if $i > k$. Call this product W so that the first k rows will be $W_{1*}, W_{2*}, ..., W_{k*}$. Since $XW = B$ it is clear each row of B is a union of rows of W. Thus, the W_{i*}'s are a set basis for the S_j's.

Now suppose that the set basis problem has a solution $T_1, ..., T_c$ where $c \leq k$. Let the matrix Y be given by $Y_{ij} = T_i[j]$ for all $i \leq c$ and $j \leq q$ and $Y_{ij} = 0$ otherwise. Let X be a matrix given by

$$X_{ij} = \begin{cases} 1 & \text{if } T_j \text{ is componentwise} \leq S_i \\ 0 & \text{otherwise} \end{cases}$$

It is easy to see that $XAY = X(AY) = XY = B$.

The preceding shows that deciding whether one principal ideal contains another is at least as hard as an *NP*-complete problem. It is easy to see that this problem is in the class *NP* since a solution can be specified by giving matrices X and Y such that $XAY = B$. You can verify that a purported solution is correct by carrying out the multiplication. This shows that the problem of ordering *D*-classes is *NP*-complete.

Not only does the above show that deciding when one principal ideal contains another is *NP*-complete, but that this problem is *NP*-complete even when one of the matrices is something as simple as A which has 1's only along the diagonal. The fact that a simple version of the problem is *NP*-complete means that determining ordering relations among principal ideals does not become "significantly" more difficult for more complicated matrices in the sense that either both or neither can be done in polynomial time.

The difficulty of determining the ordering among principal ideals suggests that problems dealing with this ordering will be difficult to solve. Some of these problems are discussed in [Br].

The referee of the first version of this paper observed that a matrix is in the ideal of I_k if and only if the matrix has a Schein rank $\leq k$ (see [Ki; p.37]). Thus, it follows that determining whether a matrix has Schein rank $\leq k$ is an *NP*-complete problem. This last result was first established by Dana S. Nau [Na] in the context of specificity covers of reaction matrices.

Determining the Schein rank is at least as hard as determining if it is bounded by a particular number. Problems that are at least as hard as *NP*-complete problems are called *NP*-hard. Determining the Schein rank is *NP*-hard. For more details on *NP*-hard problems see [GJ].

References

[Br] Breen, Michael, "A Maximal Chain of Principal Ideals in the Semigroup of Binary Relations on a Finite Set", *Semigroup Forum*, 43(1991), no.1, 63-76..

[GJ] Garey, Michael R. and Johnson, David D., *Computers and Intractability*, W. H. Freeman, San Francisco, 1979..

[Ki] Kim, Ki Hang, *Boolean Matrix Theory and Applications*, Marcel Dekker, New York and Basel, 1982..

[Na] Nau, Dana S., "Specificity covering: immunological and other applications, computational complexity and other mathematical properties, and a computer program," A.M. Thesis, Technical Report CS-1976-7, Computer Science Department, Duke University, Durham, NC, 1976..

[St] Stockmeyer, Larry J., "The Set Basis Problem is *NP*-Complete", Research Report No. RC-5431, IBM Thomas J. Watson Research Center, Yorktown Heights NY, 1975..

Computer Science Dept.
University of Maine
Orono, ME 04469

Received February 6, 1991
and in final form September 16, 1991