Discrete Mathematics 28 (1979) 153–159

THE GREEDY ALGORITHM FOR PARTIALLY ORDERED SETS

Ulrich FAIGLE

FB Mathematik, AG Allgemeine Algebra, TH Darmstadt, D-6100 Darmstadt

Received 31 October 1978
Revised 26 April 1979

Those independence systems on finite partially ordered sets are characterized for which the greedy algorithm always works.

1. Introduction

After Kruskal [8] presented his algorithm to find a shortest spanning subgraph of a graph, it was soon realized by Rado [10] that this algorithm may be extended to matroids in order to determine optimal bases (see also Edmonds [2] and Welsh [12]). In fact, given a system $C(S)$ of subsets of a finite set S such that $C(S)$ contains all the subsets of any of its members and any assignment of non-negative weights to the elements of S, this "greedy algorithm" (the name is due to Edmonds) always finds an optimal member of $C(S)$ precisely when $C(S)$ is the system of independent sets of some matroid on S (see Gale [6]).

If $C(S)$ is the system of independent sets of some matroid on S and $w: S \to \mathbf{R}$ a non-negative weight-assignment, the greedy algorithm selects $x_1 \in S$ such that $\{x_1\} \in C(S)$ and $w(x_1)$ maximal. Then it selects $x_2 \in S - x_1$ such that $\{x_1, x_2\} \in C(S)$ and $w(x_2)$ maximal etc., and stops exactly when a basis of $C(S)$ is obtained. Suppose that $(x_1, \ldots, x_k)$ has been selected after k steps. Then $w(x_1) \geqslant w(x_2) \geqslant \cdots \geqslant w(x_k) \geqslant 0$. Moreover,

$$\sum_{i \leqslant k} w(x_i) = \max \left\{ \sum_{x \in I} w(x) : |I| = k,\ I \in C(S) \right\}.$$

So the greedy algorithm is optimal at any stage.

Analogs of the greedy algorithm for matroids have been investigated in more general contexts (see Edmonds [1], Dunstan and Welsh [4], Hammer et al. [7], Euler [3]).

In this paper, we consider the following situation. Let P be a (finite) set partially ordered by a priority relation, where we write $x \leqslant y$ if x is before or equals y, and $w: P \to \mathbf{R}$ be a non-negative weight assignment so that $w(x) \geqslant w(y)$ if x dominates y.

We are interested in an algorithm which selects one element of P at a time and respects the priority relation in the sense that for $x<y$, x cannot be selected *after* y has been selected. That is, we consider systems $C(P)$ of tuples $(x_1,\dots,x_k)$, $x_i \in P$, $x_i \leq x_j$ implies $i \leq j$, such that with each tuple $I \in C(P)$ all initial segments of I are also members of $C(P)$. Such a $C(P)$ is an "independence system" on P.

With respect to the weight-assignment $w: P \to \mathbf{R}$, the greedy algorithm may be formulated for the independence system $C(P)$. A natural question to ask is therefore: "Does the greedy algorithm always select an optimal member of $C(P)$?" It is easy to see that the general answer is "no".

The purpose of this paper is to provide a characterization of those independence systems for which the greedy algorithm always works. We call those systems "generating systems". Our main results are given in Sections 3 and 4. In Section 5 we briefly outline the connection between generating systems and "geometries on partially ordered sets", i.e., matroid analogs on partially ordered sets.

2. Independence systems

For a given (finite) partially ordered set P, we consider a collection $C(P)$ of tuples $(x_1,\dots,x_n)$ with $x_i \in P$ and $x_i \neq x_j$ for $i \neq j$, $1 \leq i, j \leq n$. By abuse of language, we will occasionally refer to those tuples as "sets" – still keeping in mind, however, that the order in which the elements are listed is important.

$C = C(P)$ is an *independence system* on P if

(IS_1) for any $I=(x_1,\dots,x_k) \in C$, $x_i \leq x_j$ implies $i \leq j$, $1 \leq i, j \leq k$.

(IS_2) for any $I=(x_1,\dots,x_k) \in C$, $I_m \in C$, where I_m is the initial segment $(x_1,\dots,x_m)$ of I of length m, $0 \leq m \leq k$.

Identifying $I_0 = \emptyset$, we see that the empty set is the smallest element of any independence system.

We define the *height* of C as

$$h(C) = \max\{|I| : I \in C\}$$

and call the element $B \in C$ a *basis* of C if $|B| = h(C)$. An element $p \in P$ which lies in every basis of C is an *isthmus* of C.

A subset $A \subset P$ is an (order) *ideal* of P if for all $y \in A$, $x \in P$, $x \leq y$ implies $x \in A$.

Let A be any ideal of P. Then we may define the *sub-system*

$$C(A) = \{I \in C : I \subset A\}$$

of $C(P)$. Clearly, $C(A)$ is an independence system if $C(P)$ is. We also use the notation $C(A) = C(P) - (P - A)$.

3. The greedy algorithm

A *natural weighting* w of the partially ordered set P is a function $w: P \to \mathbf{R}$ such that for all $x, y \in P$, $x \leqslant y$ implies $w(x) \geqslant w(y) \geqslant 0$. (In christening an order *reversing* function "natural" we follow Stanley [11].)

w extends to a non-negative function on the independence system $C(P)$ in the obvious manner: for all $I \in C(P)$,

$$w(I) = \begin{cases} \sum_{x \in I} w(x) & \text{if} \quad I \neq \emptyset, \\ 0 & \text{if} \quad I = \emptyset. \end{cases}$$

If $I = (x_1, \ldots, x_k) \in C(P)$ so that $w(I) = \max\{w(J) : J \in C(P)\}$, then I is *optimal*. I is *w-feasible* if $w(x_1) \geqslant w(x_2) \geqslant \cdots \geqslant w(x_k)$.

The *greedy algorithm* is a procedure which determines a w-feasible element of the independence system $C = C(P)$ as follows:

Step 1. Choose $x_1 \in P$ so that $w(x_1)$ is maximal and $(x_1) \in C$. If no such choice is possible, stop. Otherwise continue.

Step 2. Choose $x_2 \in P - x_1$ so that $(x_1, x_2) \in C$ and $w(x_2)$ is maximal among those $w(x)$ with $w(x) \leqslant w(x_1)$ and $(x_1, x) \in C$. If no such choice is possible, stop. Otherwise continue.

Step k. Choose $x_k \in P - \{x_1, \ldots, x_{k-1}\}$ so that $(x_1, \ldots, x_{k-1}, x_k) \in C$ and $w(x_k)$ is maximal among those $w(x)$ with $w(x) \leqslant w(x_{k-1})$ and $(x_1, \ldots, x_{k-1}, x) \in C$. If no such choice is possible, stop. Otherwise continue.

So the greedy algorithm will always exhibit a w-feasible element of C. However, this element will, in general, not be optimal. We illustrate this with the following

Example. Let $P = \{a, b, c\}$ be a partially ordered set with the only non-trivial relation $b < c$, and consider the independence system

$$C(P) = \{\emptyset, (a), (b), (b, c)\}.$$

If the weights are $W(a) = \frac{1}{2}$, $w(b) = w(c) = \frac{1}{3}$, the greedy algorithm will select (a), which is not optimal.

We say that the greedy algorithm *works* if it always selects an optimal set in C.

Also note that by property (IS_1) of an independence system the greedy algorithm, at no stage, selects an element which, in the priority relation, is before an element already selected.

Suppose now that $C = C(P)$ is an independence system on the partially ordered set P for which the greedy algorithm always works, no matter how the natural weighting w on P may be defined.

(GS_1) For all $I, I' \in C$ such that $|I|<|I'|$, there exists $x \in I'$ and $y \in P$ with $y \leq x$ and $(I, y) \in C$.

Proof. Define for $x \in P$,

$$w(x)=\begin{cases}1 & \text{if there is a } z \in I \cup I', \quad x \leq z,\\ 0 & \text{otherwise.}\end{cases}$$

Then w is a natural weighting on P, and I is in compliance with the greedy algorithm with respect to w. Because of $w(I)<w(I')$ and the hypothesis that the greedy algorithm works, (GS_1) follows. □

Note that (GS_1) says in particular that any element of C can be completed to a basis of C, and that all bases of C are equicardinal.

(GS_2) For all ideals $A, B \subset P$, with $A \subset B$, if the element $p \in A$ is an isthmus of $C(B)$, then p is an isthmus of $C(A)$.

Proof. Define for $x \in P$,

$$w(x)=\begin{cases}1 & \text{if } x \in A,\\ 0 & \text{otherwise.}\end{cases}$$

Suppose $I \in C$ is a basis of $C(A)$ with $p \notin I$. Since I is in compliance with the greedy algorithm, we may, by (GS_1), use the greedy algorithm again to complete I to a basis of $C(B)$. If x is the next element chosen, then $x \notin A$ because I already was a basis of $C(A)$. So $w(x)=0$. Recalling that $w(p)=1$ and that the greedy algorithm produces a w-feasible basis, we see that there exists a basis I' of $C(B)$ with $p \notin I'$, i.e., p is not an isthmus of $C(B)$. Here we have implicitly used the fact that if the greedy algorithm always works for C, it always works for any subsystem of C. □

4. Generating systems

We call an independence system $C(P)$ a *generating system* (g.s.) if $C(P)$ satisfies (GS_1) and (GS_2). A motivation for the name "generating system" will be indicated in the next section.

The purpose of this section is to show that for any g.s. $C(P)$ and any natural weighting $w: P \to \mathbf{R}$, the greedy algorithm works.

The proofs in this section will make use of the observation that every subsystem of a g.s. is a g.s.. This enables us to proceed by induction on the cardinality $|P|$ of the underlying partially ordered set P. By P^+ we mean $\{p \in P: p < q \text{ for no } q \in P\}$, the set of maximal elements.

So let $C = C(P)$ be a fixed g.s. and $w: P \to \mathbf{R}$ a fixed natural weighting on P.

(i) There exists a w-feasible optimal basis of C.

Proof. Let $B = (b_1, \ldots, b_n)$ be an optimal basis of C. If there is $e \in P^+$ with $e \notin B$, then B is also an optimal basis of the g.s. $C - e$ with the induced natural weighting. Hence (i) holds by induction on $|P|$.

So assume $P^+ \subset B$. In particular, $b_n \in P^+$ because of (IS_1). If $h(C - b_n) = h(C)$, $B_{n-1} = (b_1, \ldots, b_{n-1})$ may be completed to a basis (B_{n-1}, b) of $C - b_n$. Now $P^+ \subset B$ implies $b \leq b_n$. Hence $w(b_n) \leq w(b)$ and $w(B_{n-1}, b) \geq w(B)$, i.e., the induction hypothesis on $C - b_n$ may be applied.

If $h(C - b_n) = h(C) - 1$, then b_n is an isthmus of C and therefore (B', b_n) must be a basis of C, for every basis B' of $C - b_n$. If there is an optimal basis B' of $C - b_n$ with $e \notin B'$ for some $e \in P^+$, $e \neq b_n$, then (B', b_n) is optimal in $C - e$ and in C so that (i) follows as before.

Therefore, by the induction hypothesis, we may assume that B_{n-1} is optimal and w-feasible in $C - b_n$. If $w(b_n) \leq w(b_{n-1})$, there is nothing left to prove. Otherwise, in particular $b_{n-1} \not\leq b_n$, and hence $b_{n-1} \in P^+$ and $w(b_{n-1}) = \min\{w(x) : x \in P\}$.

Consider now $B_{n-2} = (b_1, \ldots, b_{n-2}) \in C - b_{n-1}$.

If $h(C - b_{n-1}) = h(C)$, B_{n-2} may be completed to a basis (B_{n-2}, x, y) of $C - b_{n-1}$. Since, by (GS_2), b_n is an isthmus of $C - b_{n-1}$, $b_n \in (B_{n-2}, x, y)$ and the minimality of $w(b_{n-1})$ shows that (B_{n-2}, x, y) is optimal in C. So the induction hypothesis applies again to $C - b_{n-1}$.

In the final case, $h(C - b_{n-1}) = h(C) - 1$, b_{n-1} and b_n are isthmi of C. Thus (B_{n-2}, b_n) is a basis of $C - b_{n-1}$ and consequently (B_{n-2}, b_n, b_{n-1}) a basis of C. Now, keeping in mind that $w(b_{n-1}) \leq w(x)$, for all $x \in P$, we may repeat the argument with (B_{n-2}, b_n, b_{n-1}) instead of B. □

(ii) If $B_k = (b_1, \ldots, b_k) \in C$ is constructed according to the greedy algorithm, then $w(x) \leq w(b_k)$, for all $x \in P$ with $(B_k, x) \in C$. In particular, the greedy algorithm constructs a basis of C.

Proof. Suppose that B_k is a smallest (with respect to k) counterexample and $b \in P$ so that $w(b) > w(b_k)$, $(B_k, b) \in C$.

Then we may assume $P^+ \subset (B_k, b)$ and, in particular, $b \in P^+$ since otherwise (ii) is seen to be true by induction on $|P|$.

If $h(C - b) = h(C)$, then there is $b' \in P$ such that $(B_k, b') \in C - b$. But because of $P^+ - b \subset B_k$, $b' \leq b$ and $w(b') \geq w(b) > w(b_k)$, in contradiction to the induction hypothesis for $C - b$.

Therefore $h(C - b) = h(C) - 1$, i.e., b is an isthmus. Due to $w(b_k) < w(b)$ and $P^+ - b \subset B_k$, we have $b_k \in P^+$.

If $h(C - b_k) = h(C)$, $B_{k-1} = (b_1, \ldots, b_{k-1})$ may be augmented to a basis (B_{k-1}, x, y) of $C - b_k$ in accordance to the greedy algorithm since, by hypothesis,

(ii) holds for $C-b_k$. Since b is an isthmus of $C-b_k$, $x=b$ or $y=b$. In any case, $w(x)\geqslant w(b)>w(b_k)$, contradicting the choice of b_k.

If $h(C-b_k)=h(C)-1$, (B_{k-1}, b) must be a basis of $C-b_k$. Since B_{k-1} was constructed according to the greedy algorithm and b is an isthmus of $C-b_k$, (B_{k-1}, b) must be w-feasible by the induction hypothesis for $C-b_k$. But this again contradicts the choice of b_k. □

(iii) The greedy algorithm selects an optimal basis.

Proof. Suppose, w.l.o.g., that $Y=(y_1,\dots,y_n)$ is an optimal w-feasible basis of C, and $B=(b_1,\dots,b_n)$ is a basis constructed by the greedy algorithm which is not optimal.

Then there exists an index $k\leqslant n$ such that $w(b_i)\geqslant w(y_i)$, $i=1,\dots,k-1$, and $w(b_k)<w(y_k)$. Consider $B_{k-1}=(b_1,\dots,b_{k-1})$ and $Y_k=(y_1,\dots,y_k)$. By (GS_1), there is a $y\in Y_k$ and $y'\leqslant y$ with $(B_{k-1}, y')\in C$. But then $w(y')\geqslant w(y)\geqslant w(y_k)> w(b_k)$, a contradiction to (ii). □

We remark that with similar methods one can prove:

(iv) If $B_k=(b_1,\dots,b_k)\in C$ is constructed according to the greedy algorithm, B_k not necessarily a basis, then $w(B_k)=\max\{w(I): I\in C, |I|=k\}$. □

A summary of our results is given in

Theorem. *Let $C(P)$ be an independence system on a partially ordered set P. Then $C(P)$ is a generating system if and only if the greedy algorithm works for every natural weighting $w: P\rightarrow \mathbf{R}$.* □

5. Geometries on partially ordered sets

We finally mention the connection between generating systems and matroid-type structures on a partially ordered set P.

A *geometry* $G(P)$ is a pair (P, r), where r is a non-negative integer-valued function defined on the subsets of P so that

(R_0) $r(\emptyset)=0$;

(R_1) for all $S\in P$, $r(S)=r(\bar{S})$, where $\bar{S}$ is the smallest ideal containing S;

(R_2) for all $S\in P$, $r(S)=r(S\cup(p))$ implies, for all $p\in P$, $r(S)\leqslant r(S\cup p)\leqslant r(S)+1$, where $(p)=\{q\in P: q<p\}$;

(R_3) for all $S, T\in P$, $r(S\cup T)+r(S\cap T)\leqslant r(S)+r(T)$.

Then one may show that for each g.s. $C(P)$, there exists a unique geometry $G(P)$ whose "rank" function r is given by for any ideal $S\in P$,

$$r(S)=\max\{|I|: I\subset S,\ I\in C(P)\}.$$

Furthermore, every geometry may be obtained this way.

If P is a trivially ordered set, a generating system is therefore essentially the system of independent sets of some matroid on P.

For details, we refer to [5], especially the proof of Theorem 9.

Acknowledgement

Finally, we take the opportunity here to thank R. Ziegler for many an enjoyable conversation.

References

[1] J. Edmond, Submodular functions, matroids and certain polyhedra, Proc. Int. Conf. Comb. Calgary (Gordon and Breach, New York, 1970) 69–87.
[2] J. Edmonds, Matroids and the greedy algorithm, Math. Programming 1 (1971) 127–136.
[3] R. Euler, Die Bestimmung maximaler Elemente in monotonen Mengensystemen, Univ. Bonn, Report No. 7761-OR (1977).
[4] F.D.J. Dunstan and D.J.A. Welsh, A greedy algorithm for solving a class of linear programmes, Math. Programming 5 (1973) 338–353.
[5] U. Faigle, Geometries on partially ordered sets, to appear in J. Combinatorial Theory B.
[6] D. Gale, Optimal assignments in an ordered set: an application of matroid theory, J. Combinatorial Theory 4 (1968) 176–180.
[7] P.L. Hammer, E.L. Johnson and U.N. Peled, Regular 0–1 programs, Univ. Waterloo, Comb. and Op. Res. Report CORR 73-18, Sept. 73.
[8] J.B. Kruskal, On the shortest spanning subgraph of a graph and the travelling salesman problem, Proc. Amer. Math. Soc. 7 (1956) 48–49.
[9] E.L. Lawler, Combinatorial Optimization: Networks and Matroids (Holt, Rhinehart and Winston, New York, 1976).
[10] R. Rado, Note on independence functions, Proc. Lond. Math. Soc. 7 (1957) 300–320.
[11] R.P. Stanley, Ordered Structures and Partitions, Memoirs Am. Math. Soc. 119 (Providence, RI, 1972).
[12] D.J.A. Welsh, Kruskal's theorem for matroids, Proc. Cambr. Phil. Soc. 64 (1968) 3–4.
[13] D.J.A. Welsh, Matroid Theory (Academic Press, New York, 1976).