CMPSCI 611 Advanced Algorithms **Fall 2001**

Lecture 5: September 19

Lecturer: Micah Adler *Scribes: Yuichi Kikuchi*

5.1 Greedy Algorithm: Subset System

5.1.1 Subset System

Definition 5.1 *A* **subset system** $S = (E, \mathcal{I})$ *is a finite set* E *with a collection* $\mathcal{I}$ *of subsets of* E *such that*

$$\text{if } i \in \mathcal{I} \text{ and } i' \subseteq i \text{ then } i' \in \mathcal{I}.$$

The above property is commonly stated as "$\mathcal{I}$ is closed under inclusion." We refer to the elements of $\mathcal{I}$ as "independent sets." (Also called "valid" subsets of E)

Example

$$\begin{aligned} E &= \{e_1, e_2, e_3\} \\ \mathcal{I} &= \{\{e_1, e_2\}, \{e_2, e_3\}, \{\}, \{e_1\}, \{e_2\}, \{e_3\}\} \end{aligned}$$

5.1.2 Optimization Problem for subset system

The **subset system** is used for optimization problem on (E,$\mathcal{I}$) of the following general form:

Input: $(E, \mathcal{I})$, and weight function $w : E \to \Re^+$
Output: $i \in \mathcal{I}$ such that $w(i)$ is maximum. $(w(i) = \sum_{e \in i} w(e))$

5.1.2.1 Example: Maximum Weight Forest (MWF)

This is an example of an optimization problem suited for use of subset systems. The MWF problem is defined as follows:

$$\begin{aligned} E &: \textit{edges of graph} \\ \mathcal{I} &: \textit{acyclic subsets of edges} \end{aligned}$$

Translating this into an optimization problem, following the form above, we have E, the edges of the graph, and $\mathcal{I}$, the acyclic subgraphs (i.e., the forests) of the graph G. We know this is closed under inclusion, since removing an edge from a forest always returns a forest.

5.1.2.2 Example: Maximum Weight Matching

$$\begin{aligned} E &: \textit{edges of graph} \\ \mathcal{I} &: \textit{subsets of edges such that no two edges share a vertex.} \end{aligned}$$

5.1.3 Greedy Algorithm

A greedy algorithm for optimization problems on (E,$\mathcal{I}$) is as follows:

$i = \emptyset$
Sort the elements of E by non-increasing weight
For each $e \in E$
  if $i + e \in \mathcal{I}$ **then** $i = i + e$
return i

(Note: It is assumed that we can check if a set of edges is a subset of $\mathcal{I}$) It turns out that for some subset systems, this is very fast and produces optimal output. But for others, it does not. For what subset system does greedy algorithm work?

- MWF: Greedy Algorithm always finds the optimal solution. See the following example.

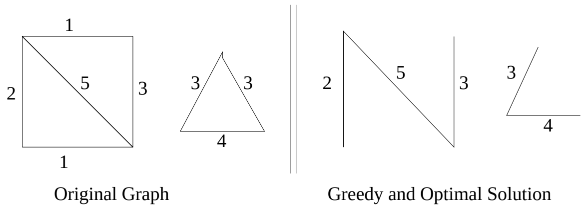

Figure 5.1: Example of MWF Algorithm

- MST: define weight function $w(e) = (w - cost\ of\ edge\ e)$, where w is the maximum cost of any edge to solve MWF.

5.1.4 Maximum Matching Problem

Definition 5.2 : *A matching M in a graph $G = (V, E)$ is a subset of edges $e \in E$ such that $\forall$ vertices $v \in V$, at most one edge in M is incident to v.*

See the following example.

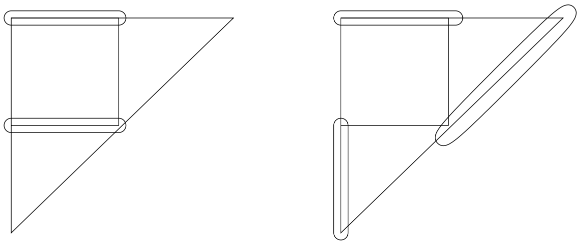

Figure 5.2: Example of two Matching

5.1.4.1 Maximum Matching Problem

Maximum Matching Problem is defined as follows.

Input:	$G(V, E)$, $w : E \to \Re^+$
Output:	a matching of maximum weight.

The following figure shows that for this problem, the greedy approach does not find the optimal solution.

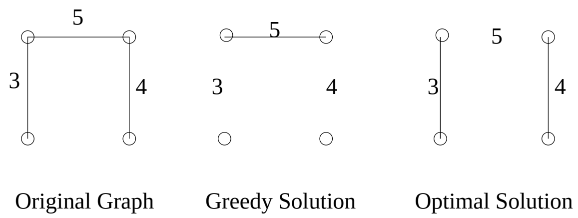

Figure 5.3: Comparison of Greedy and Optimal solutions for a Matching Problem

(Note: If no edges can be added to a set of edges, we say that the set is **maximal**. If no other solution has larger weight, the solution is maximum thus considered optimal in this problem.)

5.2 Matroids

Definition 5.3 *A matroid is a subset system $M = (E, \mathcal{I})$ that satisfies the exchange property: if $i, i' \in \mathcal{I}$ such that $|i| < |i'|$, then there is some $e \in i' - i$, such that $i + e \in \mathcal{I}$.*

Example: The subset system for MWF is a matroid.

Proof: We need to show that if i and i' are forests, and $|i| < |i'|$ then i' has edge e such that $i + e$ is a forest. We will use, without proving, the knowledge that the number of connected components in a forest with K edges is $|V| - K$. Then

$|i| < |i'|$
$\Rightarrow$ i has more connected components than i'
$\Rightarrow$ i' has an edge e between connected components of i
$\Rightarrow$ $i + e$ is a forest. ■

Now we'll prove a fact about any subset system that has the matroid property, and use that fact to prove that any subset system with this property has an optimal greedy solution.

Theorem 5.4 Cardinality Theorem: *A subset system $(E, \mathcal{I})$ is a matroid if and only if $\forall A \subseteq E$, if i and i' are maximal independent subsets of A, then $|i| = |i'|$.*

Proof: ($\Longrightarrow$: i.e., if a matroid, then for all $A \subseteq E$, if i and i' are maximal independent subsets of A, then $|i| = |i'|$)

Let (E, I) be a matroid, and i, i' are maximal independent subsets of A.

If $|i| < |i'|$, then by the exchange property (which we have, since we're assuming we have a matroid) we can increase the size of i by $e \in i' - i$. And $i + e$ still belongs to A, which is a contradiction to the claim that i is a maximal independent subset.

Thus the only way i and i' can be maximal independent subsets of A and conform to the exchange property is if $|i| = |i'|$.

($\Longleftarrow$: i.e., if not a matroid, then not true for some A...)
If (E, I) is not a matroid, then $\exists i, i' \in I$, such that $|i| < |i'|$ but there is no $e \in i' - i$ that satisfies $i + e \in I$.
Let $A = i \cup i'$, then i is a maximal independent subset in A because there is no $e \in A$ that satisfies $i + e \in A$
There is some i'', such that $i' \subseteq i''$ and i'' is maximal independent subset in A.
Since $|i''| \geq |i'| > |i|$, we get i and i'' are both maximal independent subsets, and they satisfy $|i''| \neq |i|$. ■

Now, we can prove that MWF is a matroid using Cardinality Theorem. Consider any $A \subseteq E$. Let $\mathcal{C}$ be the number of connected components in A. Any maximal acyclic subset of A has $|V| - C$ edges. Each has $\mathcal{C}$ connected components.

$$C = |V| - k.$$

5.3 Reference

- Scribe notes for lecture 4 of Computer Science 611, University of Massachusetts at Amherst, 2000.