CS F364
Design & Analysis of Algorithms

Algorithm Design Techniques - GREEDY

Matroids – A Theoretical Framework for Greedy Algorithms

- **Definition and Examples**
- **Properties**
- **Greedy Algorithms**
    - **Correctness and Efficiency**

2/24/2015 Sundar B. CSIS, BITS, Pilani

Matroids: Motivation

- **Matroids** provide a ***sound*** but ***incomplete*** theory of the ***greedy method*** i.e.
  - The theory of matroids can be used to derive **optimal solutions** for many problems / scenarios but
  - the theory **does not cover all cases** for which a greedy method applies.

Matroids: Definitions

- Definitions (Hereditary Property and Exchange Property)
  - Given a finite set S and a nonempty family I of subsets – referred to as ***independent subsets*** – of S:
    - I is said to be ***hereditary*** if it satisfies the following property:
      - if B is in I, and A is subset of B, then A is in I
    - I is said to satisfy the ***exchange*** property if the following is true:
      - if A is in I, B is in I, and |A| < |B|,
      - then there exists x is in B-A such that (A U {x}) is in I

MATROIDS: DEFINITIONS [2]

- Definition:
  - A ***matroid*** is an ordered pair M = <S, I> such that
    1. S is a finite set
    2. I is a nonempty family of subsets of S, satisfying
       a. the hereditary property and
       b. the exchange property
- Example:
  - Let S be any finite set and
  - I_k be the set of all subsets of S of size at most k, where k<=|S|
    - Then $\langle S, I_k \rangle$ is a matroid.
- Example:
  - Linearly independent columns of a matrix.

GRAPHIC MATROID

- Given, an undirected graph G=(V,E), consider M_G = <S_G,I_G> defined as:
  - $S_G = E$
  - If $A \subseteq S$, then A is in I_G iff A is acyclic
    - i.e. a set of edges A is independent iff the subgraph $G_A = (V, A)$ forms a forest (of trees).
- Theorem:
  - If G=(V,E) is an undirected graph, then M_G=<S_G,I_G> is a matroid.
- Given G=(V,E), **M_G=<S_G,I_G>** is referred to as the ***Graphic Matroid.***

Graphic Matroid

Theorem (Graphic Matroid):

If $G=(V,E)$ is an undirected graph, then $M_G=\langle S_G, I_G \rangle$ is a matroid.

Proof:

- S_G is finite.
- I_G is hereditary: *the subset of a forest is a forest*
- I_G satisfies the exchange property:
  - Let $G_A = (V,A)$ and $G_B = (V, B)$ be forests of G and let $|B| > |A|$
  - Then G_A and G_B contain $|V|-|A|$ and $|V|-|B|$ trees resp.
  - Therefore there must be a tree T in G_B
    1. whose vertices are from two different trees in G_A and
    2. there must be an edge (u,v) in T such that u and v are from different trees in G_A
    - Adding (u,v) to G_A preserves acyclicity:

      i.e. (G_A U {(u,v)}) is in I_G
    - i.e. I_G satisfies the exchange property

Matroids : Extensions to Subsets

- Definition (**Extension**)*:*
  - Given a matroid G = <S,I>, an element x not in A is said to be an extension of A is in I if (AU{ x }) is in I:
    - i.e. *if addition of x to A preserves independence*.
- Example:
  - Consider a graphic matroid M_G :
    - Let A be independent set of edges.
    - Then edge e is an extension of A iff
      1. e is not in A and
      2. adding e does not induce a cycle

MATROIDS : MAXIMAL INDEPENDENT SUBSET

- Definition (**MIS)**:
  - Given a matroid G = <S,I>, an independent subset A is in I is said to be maximal if it has no extensions:
    - i.e. *if A is not a subset of any B* in I.
- Example:
  - Consider a graphic matroid M_G :
    - What would be an MIS for M_G ?

Matroids : Property of MISs

- Theorem (Size of **MISs)*:***
  - All maximal independent subsets of a matroid are of equal size.
- Proof (by contradiction):
  - Let A be an MIS of a given matroid M.
  - Suppose there is another independent subset B of M that is maximal and |B|>|A|
    - Then by the exchange property
      - there exists an x in B-A such that (A U {x}) is in I
        - i.e. A is not maximal.

QED

WEIGHTED MATROIDS

- Definition (**Weighted Matroids**):
  - A matroid M = <S,I> is weighted if there is a weight function w: S --> Z^+
- Definition (Weight of independent sets):
  - The weight function w can be extended to the members of I:
    - For any A is in I, w(A) = $\Sigma_{x \text{ in } A}$ w(x)

Weighted Matroids – Optimal Subsets

- Given a weighted matroid M = <S, I, w>, an independent subset A with maximum possible weight is said to be ***optimal***.
  - Claim:

    Since the weight function is positive on elements of S, an optimal subset is always a maximal independent subset.
  - Terminology:
    - To avoid confusion, we may refer to **optimal subsets** as **maximum weight subset**.

Weighted Matroids – Optimal Subsets

Recall:

Given, an undirected graph G=(V,E), define M_G = <S_G,I_G> as:

- $S_G = E$
- If $A \subseteq S$, then A is in I_G iff A is acyclic

- Then an MIS of M_G is a spanning tree of G

- Consider the ***minimum spanning tree*** problem:
  - Given a weighed graph G = (V,E,w) extend M_G above by the weight function w' :
    - $w'(e) = w_m - w(e)$ where $w_m >= \max_{e \text{ in } E}(w(e))$
  - Then an optimal subset of the weighted matroid is a minimum spanning tree:
    - $w'(A) = \Sigma_{e \text{ in } A}\, w'(e) = (|V|-1)*w_0 - \Sigma_{e \text{ in } A}\, w(e)$

      $= (|V|-1)*w_0 - w(A)$ for any MIS.

WEIGHTED MATROIDS – GREEDY ALGORITHM

- GreedyWM(M) // M is a weighted matroid: <S,I,w>
  1. A = {} // A is the optimal subset being constructed
  2. let N=|M.S|
  3. sort elements of M.S in decreasing order by weight w
  4. for i = 1 to N

     if A U { M.S[i] } in M.I then A = A U { M.S[i] }
  5. return A
- Theorem:
  - Given a weighted matroid M, GreedyWM(M) returns an optimal subset.
- Complexity:
  - Time taken = Time for sorting + Time for N iterations

    = O(N*logN + N*f(N))

    where f(N) is time taken for testing whether a subset is independent.

Correctness of GreedyWM: Greedy Choice

- Lemma (MAT_CHOICE):
  - Let M = <S,I,w> be a weighted matroid with S sorted in decreasing order by weight.
  - Let x be the first element of S such that {x} is in I, if it exists:
    - then there is an optimal subset A of S such that x is in A.
- Proof (by cases):
  - If no such x exists then we are done.
  - else, let B be a nonempty optimal subset.
    - If x is in B, then let A = B; we are done.
    - else then for any element y of B, w(y) <= w(x). Why?
      - Let A = { x }
      - Using the exchange property find a y in B and add it to A until |A| = |B| i.e. A = B - { y } U { x } for some y in B.
        - w(A) >= w (B).
      - But B is optimal i.e. A must also be optimal.

Correctness of GreedyWM: Order of Choice

- Lemma (CHOICE_ORDER):
  - Let M = <S,I> be a matroid.
  - If x is in S and x is an extension of some A in I,
    - then x is also an extension of {}
- Proof:
  - Since x is an extension of A, A U {x} is in I.
  - Since I is hereditary, {x} must be in I.

- Corollary:
  - Let M = <S,I> be a matroid.
  - If x is in S and x is not an extension of {},
    - then x is not an extension of any A in I.

Correctness of GreedyWM: Optimal substructure

- Given a weighted matroid M = <S, I, w>, and an element x in S such that { x } is in I, , define a contraction of M by x as the weighted matroid M' = <S', I', w'>:
  - S' = { y in S | { x, y } is in I }
  - I' = { B subset of S – { x } | B U { x } is in I }
  - w' is w restricted to S'.
- Theorem (OPTIMAL_SUBSTRUCTURE):
  - Let x be the first element of S chosen by GreedyWM for the weighted matroid M = <S,I,w>.
  - The remaining problem of finding a maximum-weight independent subset containing x reduces to
    - finding the maximum-weight independent subset of M', the contraction of M by x.
- Proof: (omitted).