And

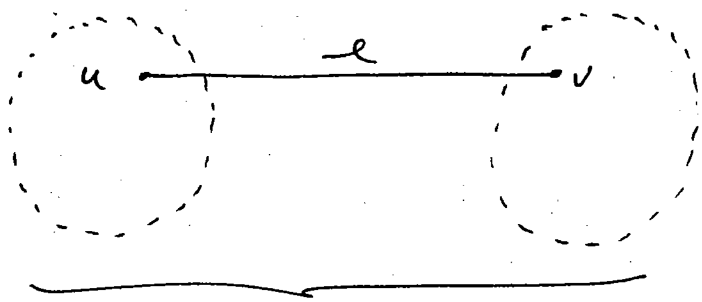
$$|A| = \sum_{i=1}^{m} |E_i| = \sum_{i=1}^{m} (|V_i| - 1) = |V| - m,$$

$\therefore m = |V| - |A|$. Likewise $(V,B)$ contains $|V| - |B|$ trees. // . )

Now $|A| < |B|$ implies $|V| - |B| < |V| - |A|$, so $(V,B)$ contains fewer trees than does $(V,A)$.

Therefore $(V,B)$ contains a tree $T$ which has vertices in two distinct trees of forest $(V,A)$. (Otherwise each tree of $(V,B)$ is a subtree of one in $(V,A)$, which implies $(V,B)$ has at least as many trees as forest $(V,A)$. $\times$ )

Thus $T$, being connected, contains an edge $e = uv$ whose ends lie in different trees of forest $(V,A)$




$e \in E(T) \subseteq B$

Distinct trees of $(V,A)$

THUS BY ADDING $e$ TO $A$ NO CYCLE IS CREATED (INSTEAD TWO TREES ARE JOINED)

$$\therefore \ (V, A \cup \{e\}) \text{ is acyclic}$$

$$\therefore \ A \cup \{e\} \in \mathcal{I},$$

AND THE EXCHANGE PROPERTY HOLDS.

<u>DEFN</u>

GIVEN A MATROID $M = (S, \mathcal{I})$ AND $A \in \mathcal{I}$, WE CALL $x \in S - A$ AN <u>EXTENSION OF $A$</u> IF $A \cup \{x\} \in \mathcal{I}$.

$A \in \mathcal{I}$ IS CALLED <u>MAXIMAL</u> IF IT HAS NO EXTENSIONS, i.e.

$$\forall x \in S - A : A \cup \{x\} \notin \mathcal{I}.$$

A MAXIMAL INDEPENDENT SET IS ALSO CALLED A <u>BASE</u> OF THE MATROID.

IN A MATRIC MATROID A BASE IS JUST A VECTOR SPACE BASIS FOR THE ROW SPACE OF THE UNDERLYING MATRIX.

IN A GRAPHIC MATROID A BASE FORMS THE EDGE SET OF A SPANNING TREE.

DEFN:

A SUBSET $D \subseteq S$ IS CALLED DEPENDENT IF IT IS NOT INDEPENDENT, I.E. $D \notin \mathbb{I}$

A SUBSET $C \subseteq S$ IS CALLED A cycle IF IT IS MINIMAL WITH RESPECT TO THE PROPERTY OF BEING DEPENDENT. I.E. $C$ IS DEPENDENT WHILE EACH OF ITS SUBSETS ARE INDEPENDENT.

OBVIOUSLY A cycle IN A GRAPHIC MATROID IS A cycle IN THE ORDINARY SENSE.

THEOREM

ANY TWO BASES IN A MATROID HAVE THE SAME CARDINALITY.

THIS COMMON CARDINALITY IS CALLED THE RANK OF THE MATROID, DENOTED rank(M).

PROOF:

LET $A, B \in \mathcal{I}$ BE BASES AND SUPPOSE $|A| < |B|$. BY THE EXCHANGE PROPERTY THERE EXISTS $x \in B - A$ SUCH THAT $A \cup \{x\} \in \mathcal{I}$, CONTRADICTING THAT $A$ IS MAXIMAL. $|B| < |A|$ LEADS TO A SIMILAR CONTRADICTION. THEREFORE $|A| = |B|$ AS CLAIMED.

DEFN.

A WEIGHTED MATROID IS A MATROID $M = (S, \mathcal{I})$ EQUIPPED WITH A (STRICTLY POSITIVE) WEIGHT FUNCTION ON ITS UNDERLYING SET $S$.

$$w : S \longrightarrow \mathbb{R}^+$$

WE EXTEND THIS WEIGHT FUNCTION TO SUBSETS $A \subseteq S$ BY SUMMATION.

$$w(A) = \sum_{x \in A} w(x)$$

A SET $A \in \mathcal{I}$ IS CALLED OPTIMAL IF IT HAS MAXIMUM WEIGHT AMONG ALL INDEPENDENT SUBSETS OF $S$.

NOTE THAT AN OPTIMAL SET $A$ IS NECESSARILY MAXIMAL SINCE OTHERWISE WE COULD FIND AN $x \in S - A$ SUCH THAT $A \cup \{x\} \in \mathcal{I}$, AND THEN

$$w(A \cup \{x\}) = w(A) + w(x) > w(A)$$

(USING HERE THAT $w(x) > 0$), WHICH CONTRADICTS THAT $A$ IS OPTIMAL.

PROBLEM

GIVEN A WEIGHTED MATROID $M = (S, \mathcal{I})$, $w: S \to \mathbb{R}^+$, FIND AN OPTIMAL SUBSET OF $S$. i.e. FIND $A \subseteq S$ SUCH THAT

- $A \in \mathcal{I}$
- $w(A) \geq w(B)$ FOR ALL $B \in \mathcal{I}$.

MANY OPTIMIZATION PROBLEMS FOR WHICH A GREEDY STRATEGY IS APPLICABLE CAN BE SHOWN, BY REFORMULATION, TO BE EQUIVALENT TO THE ABOVE PROBLEM.

MORE IMPORTANTLY, ANY PROBLEM WHICH CAN BE REFORMULATED AS ABOVE CAN BE SOLVED BY A GREEDY STRATEGY.

Ex

FIND A MINIMUM WEIGHT SPANNING TREE IN A WEIGHTED GRAPH $G = (V, E, w)$.

CONSIDER THE GRAPHIC MATROID $M_G$ OF $G$ WITH WEIGHT FUNCTION

$$w'(e) = w_0 - w(e)$$

WHERE $w_0$ IS A FIXED NUMBER GREATER THAN ALL EDGE WEIGHTS IN $G$. THUS $w'(e) > 0$ FOR ALL $e \in E$.

RECALL THAT A MAXIMAL INDEPENDENT SET (BASE) $A \subseteq E$ CORRESPONDS TO A SPANNING TREE IN $G$. OBSERVE

$$\begin{aligned} w'(A) &= \sum_{e \in A} w'(e) \\ &= \sum_{e \in A} (w_0 - w(e)) \\ &= |A| \cdot w_0 - w(A) \\ &= (|V| - 1) w_0 - w(A). \end{aligned}$$

THUS ANY SPANNING TREE WHICH MAXIMIZES $w'$ NECESSARILY MINIMIZES $w$, AND CONVERSELEY.

THE FOLLOWING ALGORITHM (SOMETIMES CALLED "THE GREEDY ALGORITHM") SOLVES THE OPTIMAL SUBSET PROBLEM IN A WEIGHTED MATROID.

GREEDY$(M, w)$

```
A ← ∅
SORT S[M] IN DESCENDING ORDER BY WEIGHT
for each x ∈ S[M]      // TAKEN IN ORDER
      if A ∪ {x} ∈ I[M]
            A ← A ∪ {x}
return A
```

THEOREM

GREEDY$(M, w)$ RETURNS AN INDEPENDENT SET OF MAXIMUM WEIGHT IN $M$.

PROOF

OBSERVE THAT $A \in \mathcal{I}$ AT EACH STAGE OF EXECUTION SO THE SET RETURNED IS CERTAINLY INDEPENDENT. IT IS ALSO MAXIMAL BY ITS VERY CONSTRUCTION.

LET $B$ BE ANY MAXIMAL INDEPENDENT SET OF $M$. THEN $|A| = |B| = r = \text{rank}(M)$. WE MUST SHOW THAT $w(A) \geq w(B)$, WHENCE $A$ IS OPTIMAL.

WRITE $A = \{x_1, \ldots, x_r\}$ AND $B = \{y_1, \ldots, y_r\}$ WHERE THE ELEMENTS ARE INDEXED BY DECREASING WEIGHTS. IN PARTICULAR, THE ELEMENTS OF A ARE LISTED IN THE ORDER SELECTED BY GREEDY(M,w).

IF $A = B$ THEN $w(A) = w(B)$ AND THERE IS NOTHING TO PROVE, SO ASSUME $A \neq B$.

LET $x_k$ BE THE FIRST ELEMENT OF A NOT IN B, i.e.

- $x_i = y_i$ FOR $1 \leq i < k$
- $x_k \neq y_k$

LET $A' = A - \{x_k\}$. THEN $A' \in \mathcal{I}$ AND $|A'| = |A| - 1 < |B|$. BY THE EXCHANGE PROPERTY THERE EXISTS $y \in B - A'$ SUCH THAT $A' \cup \{y\} \in \mathcal{I}$.

<u>CLAIM</u>: $w(x_k) \geq w(y)$

<u>PROOF</u>: OBSERVE $\{x_1, \ldots, x_{k-1}, y\} \subseteq A' \cup \{y\}$ SO THAT $\{x_1, \ldots, x_{k-1}, y\} \in \mathcal{I}$. IF $w(x_k) < w(y)$ THEN GREEDY(M,w) WOULD HAVE CHOSEN y ON THE $k^{TH}$ ITERATION OF LOOP 3–5 RATHER THAN $x_k$. $\therefore w(x_k) \geq w(y)$.

LET $A_1 = A' \cup \{y\} = (A - \{x_k\}) \cup \{y\}$. THEN $A_1 \in \mathcal{I}$ AND THE ABOVE CLAIM SHOWS $w(A) \geq w(A_1)$. ALSO NOTE $A_1$ HAS ONE MORE ELEMENT IN COMMON WITH $B$ THAN $A$ DOES, NAMELY $y$.

IF $A_1 = B$ THEN $w(A) \geq w(B)$ AND WE ARE DONE. OTHERWISE WE REPEAT THIS PROCESS WITH $A_1$ IN PLACE OF $A$ TO OBTAIN $A_2 \in \mathcal{I}$ WHERE $w(A_1) \geq w(A_2)$ AND $A_2$ HAS MORE IN COMMON WITH $B$ THAN $A_1$.

CONTINUING IN THIS FASHION WE CONSTRUCT A SEQUENCE OF INDEPENDENT SETS STARTING AT $A$ AND ENDING AT $B$ WITH DECREASING WEIGHTS:

$$w(A) \geq w(A_1) \geq w(A_2) \geq \ldots \geq w(B)$$

SHOWING THAT $w(A) \geq w(B)$ AS REQUIRED. ///

<u>RUN TIME</u>: LET $n = |S|$. IF THE TEST ON (4) TAKES $\Theta(f(n))$ TIME AND THE SORT ON (2) TAKES $\Theta(n \lg n)$, THEN GREEDY RUNS IN TIME $\Theta(n \lg n + n f(n))$.