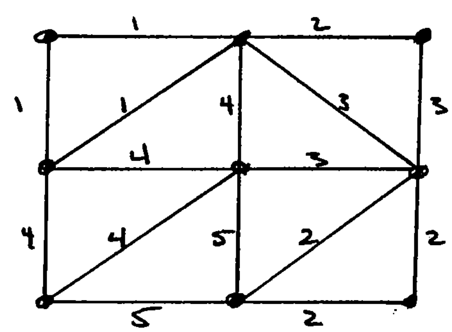
SUPPOSE $G = (V, E)$ IS EQUIPPED WITH A NON-NEGATIVE WEIGHT FUNCTION ON EDGES

$$w : E \longrightarrow \mathbb{R}_+$$

EX.



- THE **WEIGHT** OF A SPANNING TREE IS THE SUM OF THE WEIGHTS OF ITS EDGES.

PROBLEM

DETERMINE A SPANNING TREE OF MINIMUM WEIGHT IN G.

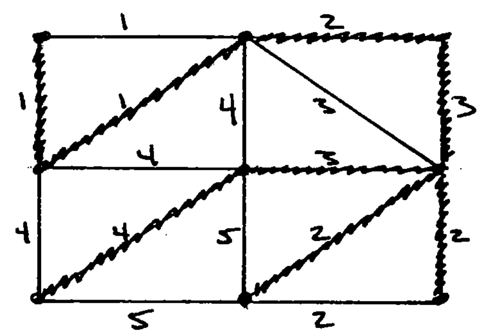
WE EXAMINE TWO FAMOUS GREEDY ALGORITHMS WHICH SOLVE THIS PROBLEM.

IN WHAT FOLLOWS LET $G = (V, E)$, $|V| = n$.

Prim's Algorithm (23.2)

- Choose an initial vertex (which is a tree)
- Amongst all edges incident with the current tree whose addition would not create a cycle, choose one of minimum weight.
- Stop when n-1 edges have been selected.

Ex



$W(T) = 18$

Observe that at each stage of execution, Prim's Algorithm maintains a tree since no cycles are created and only incident edges are added.

When this tree contains n-1 edges it must have n vertices (by previous theorem), hence it is a spanning tree.

Theorem

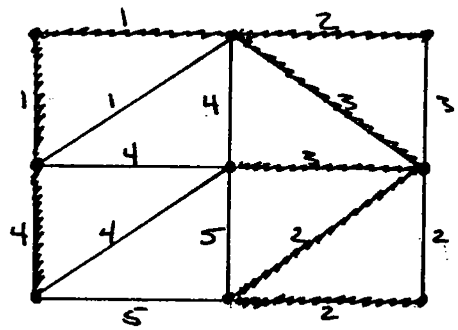
This spanning tree has minimum possible weight.

(See book or take CE 177 for proof.)

KRUSKAL'S ALGORITHM (23.2)

- CHOOSE AN EDGE OF MINIMUM WEIGHT
- AMONGST ALL EDGES WHICH DO NOT CREATE A CYCLE WITH PREVIOUSLY SELECTED EDGES, CHOOSE ONE OF MINIMUM WEIGHT.
- STOP WHEN $n-1$ EDGES HAVE BEEN SELECTED.

EX.



$w(T) = 18$

OBSERVE THAT AT EACH STAGE OF EXECUTION KRUSKAL'S ALGORITHM HAS CREATED A FOREST (UNION OF DISJOINT SUBTREES) SINCE NO CYCLES ARE CREATED.

WHEN THIS FOREST CONTAINS $n-1$ EDGES IT MUST ALSO HAVE n VERTICES. (ANY ~~GRAPH~~ FOREST WITH $n-1$ EDGES HAS AT LEAST n VERTICES. THIS FOREST CAN CONTAIN NO MORE THAN n VERTICES SINCE IT IS A SUBGRAPH OF G.)

THUS THE RESULTING FOREST IS CONNECTED (BY PREVIOUS THEOREM) AND IS A SPANNING TREE IN G.

THEOREM

THIS SPANNING TREE HAS MINIMUM WEIGHT AMONG ALL SPANNING TREES IN G.

PROOF.

LET T BE THE SPANNING TREE IN G CREATED BY KRUSKAL'S ALGORITHM, AND LET S BE ANY OTHER SPANNING TREE. WE MUST SHOW

$$w(T) \leq w(S)$$

LET $e_1, e_2, \ldots e_{n-1}$ BE THE EDGES OF T IN THE ORDER SELECTED BY KRUSKAL'S ALGORITHM. SINCE $S \neq T$ THERE IS A FIRST EDGE e_k WHICH IS NOT IN S, i.e.

$$\{e_1, \ldots, e_{k-1}\} \subseteq E(S)$$
$$e_k \notin E(S)$$

LET H BE THE SUBGRAPH OBTAINED BY ADDING e_k TO S : $H = S + e_k$. BY THE TREENESS THEOREM H CONTAINS A UNIQUE CYCLE WHICH INCLUDES e_k, CALL IT C. NOTE C MUST CONTAIN AN EDGE e OF S WHICH IS NOT IN T, FOR OTHERWISE C IS CONTAINED IN THE ACYCLIC T.

NOW REMOVE e FROM H TO OBTAIN A SUBGRAPH R, WHICH IS CONNECTED SINCE e BELONGS TO A CYCLE IN H.

$$R = H - e = S + e_k - e$$

SINCE R IS CONNECTED AND HAS $n-1$ EDGES, IT IS ANOTHER SPANNING TREE OF G, BY TREENESS THEOREM.

THE NATURE OF KRUSKAL'S ALGORITHM GUARANTEES THAT $w(e_k) \le w(e)$.

__Pf__: e DOES NOT FORM A CYCLE WITH $\{e_1, \dots, e_{k-1}\}$ SINCE $\{e_1, \dots, e_k, e\} \subseteq E(S)$. THUS IF $w(e) < w(e_k)$, THEN KRUSKAL WOULD HAVE CHOSEN e ON THE K^{TH} ITERATION OF THE GREEDY LOOP INSTEAD OF e_k. //

THUS R IS A SPANNING TREE OF G WITH ONE MORE EDGE IN COMMON WITH T THAN S, AND SATISFYING $w(R) \le w(S)$.

IF $R = T$ WE ARE DONE, OTHERWISE WE MAY PERFORM THIS SAME CONSTRUCTION WITH R IN PLACE OF S.

i.e. CONSTRUCT ANOTHER SPANNING TREE R_2 WITH ONE MORE EDGE IN COMMON WITH T THAN R, AND SATISFYING $W(R_1) \leq W(R)$.

CONTINUING IN THIS FASHION WE CONSTRUCT A SEQUENCE OF SPANNING TREES WHICH MUST EVENTUALLY REACH T:

$$W(T) \leq \ldots \leq W(R_1) \leq W(R) \leq W(S),$$

SO $W(T) \leq W(S)$ AS REQUIRED. ///

MATROIDS AND THE GREEDY ALGORITHM

A MATROID IS AN ABSTRACT MATHEMATICAL STRUCTURE WHICH GENERALIZES MANY EXAMPLES WHERE A GREEDY STRATEGY APPLIES.

DEFN.

A <u>MATROID</u> IS AN ORDERED PAIR $M = (S, \mathcal{I})$ SATISFYING:

1.) S IS A FINITE NON-EMPTY SET, AND $\mathcal{I} \subseteq \mathcal{P}(S)$. THE MEMBERS OF $\mathcal{I}$ ARE CALLED THE <u>INDEPENDENT</u> SUBSETS OF S

2.) <u>HEREDITARY PROPERTY</u>
IF $B \in \mathcal{I}$ AND $A \subseteq B$, THEN $A \in \mathcal{I}$.

3.) <u>EXCHANGE PROPERTY</u>
IF $A \in \mathcal{I}$, $B \in \mathcal{I}$, AND $|A| < |B|$, THEN THERE EXISTS $x \in B - A$ SUCH THAT $A \cup \{x\} \in \mathcal{I}$.

NOTE THAT (2) IMPLIES THAT $\emptyset \in \mathcal{I}$ (PROVIDED $\mathcal{I}$ IS ITSELF NON-EMPTY.)

<u>EX.</u> (MATRIC MATROIDS)
LET P BE A (RECTANGULAR) MATRIX AND LET $S = \{\text{rows of } P\}$, CONSIDERED AS VECTORS.
LET

$$\mathcal{I} = \{A \subseteq S \mid A \text{ is linearly independent}\}$$

OBVIOUSLY S IS FINITE AND NON-EMPTY, AND $\mathcal{I} \subseteq \mathcal{P}(S)$. PROPERTIES (2) & (3) ARE ELEMENTARY FACTS OF LINEAR ALGEBRA.

SIMILARLY WE COULD LET S BE THE COLUMNS OF D!

<u>EX.</u> (GRAPHIC MATROIDS)
LET $G = (V, E)$ BE AN UNDIRECTED GRAPH. LET $S = E$ AND

$$\mathcal{I} = \{ A \subseteq S \mid \text{SUBGRAPH } (V, A) \text{ is Acyclic} \}.$$

(1) IS CLEARLY SATISFIED. (2) HOLDS SINCE ANY SUBSET OF AN ACYCLIC SET OF EDGES IS ACYCLIC. (BY REMOVING EDGES WE CANNOT CREATE CYCLES.)

WE PROVE THE EXCHANGE PROPERTY (3):

LET $A, B \in \mathcal{I}$ AND SUPPOSE $|A| < |B|$. THEN THE FOREST (V, A) CONTAINS EXACTLY $|V| - |A|$ TREES AND (V, B) CONTAINS $|V| - |B|$ TREES.

(<u>Pf</u>: SUPPOSE (V, A) CONTAINS m TREES: $T_i = (V_i, E_i)$, $1 \leq i \leq m$. THEN $|E_i| = |V_i| - 1$