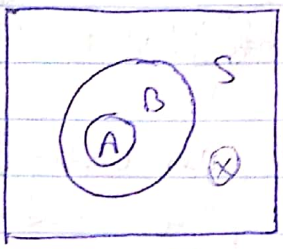
Due - 11/10 CMSC818B HW2 Adheesh C.
116236935

1) $X = \{x_1, x_2, \dots x_n\}$,

a) $f_{max}(A) = \max\limits_{x_i \in A} w_i$

where w_i is the weight of elements in A

So to prove f_{max} is a submodular f^n, first we define 3 sets $\rightarrow$ A, B, S | $A \subset B \subset S$
let $x \in S \setminus B$



since $f_{max}(B)$ is always $\geq f_{max}(A)$

we can have 3 cases
let $f_m = f_{max}$ (for notation)

Case I $\rightarrow$ $f_{max}(B) \geq f_{max}(A) \geq f_{max}(x)$

$f_m(A \cup \{x\}) - f_m(A) \;\square\; f_m(B \cup \{x\}) - f_m(B)$

$f_m(A) - f_m(A) \;\square\; f_m(B) - f_m(B)$

here since LHS = RHS = 0

$\square$ can be =, $\geq$, $\leq$

Case II $\rightarrow$ $f_m(B) \geq f_m(x) \geq f_m(A)$

$f_m(A \cup \{x\}) - f_m(A) \;\square\; f_m(B \cup \{x\}) - f_m(B)$

$f_m(x) - f_m(A) \;\square\; f_m(B) - f_m(B)$

$f_m(x) - f_m(A) \;\square\; 0$

since $f_m(x) \geq f_m(A) \Rightarrow \square$ can be = or $\geq$

Case III → $f_m(x) \geq f_m(B) \geq f_m(A)$

$$f_m(A \cup \{x\}) - f_m(A) \;\square\; f_m(B \cup \{x\}) - f_m(B)$$

$$f_m(x) - f_m(A) \;\square\; f_m(x) - f_m(B)$$

~~Here since~~ $f_m(x) - f_m(A) - f_m(x) + f_m(B) \;\square\; 0$

$$f_m(B) - f_m(A) \;\square\; 0$$

Here $f_m(B) \geq f_m(A)$

$\Rightarrow \cancel{f_m(B) - f_m(A)}\ \square$ has to $=$ or $\geq$.

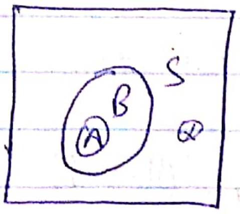
So in each case since $\square$ can only be $\geq$ or $=$, the f^n is submodular

HP

b) $f_{min}(A) = \min\limits_{x \in A} w_i$

using the same logic as part (a),

3 sets → $A, B, S \mid A \subset B \subset S$
$x \in S \setminus B$



Here $f_{min}(B) \leq f_{min}(A)$ always

So again 3 cases

let $f_m = f_{min}$ for notation sake

Case I $\rightarrow$ $f_m(B) \leq f_m(A) \leq f_m(x)$

$f_m(A \cup \{x\}) - f_m(A) \ \square \ f_m(B \cup \{x\}) - f_m(B)$

$\underbrace{f_m(A) - f_m(A)}_{0} \ \square \ f_m(B) - f_m(B)$

$\Rightarrow \cancel{f_m(B) = f_m(A)} \ \square \ 0$

since $\cancel{f_m(B) \leq f_m(A)}$

$\Rightarrow \square$ is $\leq$ or $=$ or $\geq$

Case II $\rightarrow$ $f_m(B) \leq f_m(x) \leq f_m(A)$

$f_m(A \cup \{x\}) - f_m(A) \ \square \ f_m(B \cup \{x\}) - f_m(B)$

$f_m(x) - f_m(A) \ \square \ f_m(B) - f_m(B)$

$\Rightarrow f_m(x) - f_m(A) \ \square \ 0$

since $f_m(x) \leq f_m(A) \Rightarrow f_m(x) - f_m(A) \leq 0$

$\Rightarrow \square$ is either $=$ or $\leq$

Case III $\rightarrow$ $f_m(x) \leq f_m(B) \leq f_m(A)$

$f_m(A \cup \{x\}) - f_m(A) \ \square \ f_m(B \cup \{x\}) - f_m(B)$

$f_m(A) - f_m(A) \ \square \ f_m(x) - f_m(B)$

$\Rightarrow f_m(B) - f_m(A) \ \square \ 0$

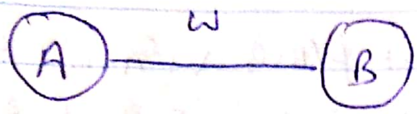
since $f_m(B) \leq f_m(A) \Rightarrow f_m(B) - f_m(A) \leq 0$

$\Rightarrow \square$ is either $=$ or $\leq$

Since for submodularity

$$f(A \cup \{x\}) - f(A) \geq f(B \cup \{x\}) - f(B)$$

& this isnt true for 2 cases, the function is not submodular.

Q2) 

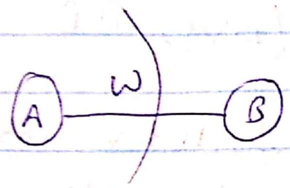
$S = \{a, b\}$

$A = \{a\}$

$B = \{b\}$

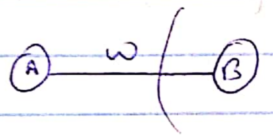
This question ~~is the~~ makes use of graph cuts. as having edge with one endpoint in S & other is $\bar{S}$ splits graph.
So graph cuts can be easily proved as shown

using cut function for A

1) 

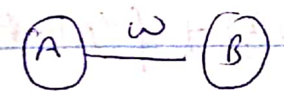
$f(a) = w$

2) cut f^n for B



$f(b) = w$

3) cut f^n for S



$f(S) = 0$

~~4) Cut f^n~~

$\therefore$ by property

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$$

sub. val.

$$w + w \geq f(S) + 0$$

$$w + w \geq 0$$

so as long as $w \geq 0$

it will always be submodular

HP

To generalize, we can consider more than 2 sets so instead of ~~A~~ & B, we have a, b ... n sets, the formula essentially becomes →

$$F(S) = \sum_{(i,j) \in E} F_{ij}(S \cap \{i, j\})$$

where i & j are 2 vertices, S is universal set

since we know that sum of submodular f^n's is also a submodular f^n.

we'll get $F(S) = n\,\omega$

where n will always be ≥ 0

& as long as ω is $\omega \geq 0$, it will be submodular.

HP

3) n = papers = $\{x_1, x_2 \ldots x_n\}$
k = reviewers
V_k = quality/value

we want to maximize quality while minimizing reviewers per paper.
This is ideally a very similar scenario to the Knapsack algorithm which is a combinatorial optimization problem.
There are certain diff btw the general setup of the 2 problems, but both essentially deal with optimization in resource allocation depending on 2 ~~sets~~ independent sets.

1

Scanned by CamScanner

So drawing parallels with the knapsack problem which is defined as →

$$\boxed{\begin{aligned}&\text{maximize } \sum_{i=1}^{n} v_i x_i \\ &\text{subject to } \sum_{i=1}^{n} w_i x_i \leq W\end{aligned}}$$

where n is no. of items, each weight is w_i,
~~Illy for us to reviewer~~ value is v_i
& ~~to~~ max weight capacity is W.

Illy for us reviewers → k, no. of papers → n
~~quality~~/value/~~of~~ paper → v_i ; $x_i \to \{x_1, x_2 \ldots x_n\}$ papers

So
$$\boxed{\begin{aligned}&\text{maximize } \sum_{i=1}^{n} v_i x_i \\ &\text{subject to } \sum_{i=1}^{n} w_i x_i \geq \frac{k}{n}\end{aligned}}$$

w_i is the weightage given every time a paper is reviewed.

clearly as reviewers/paper ↑ value of paper ↓

As ~~&~~ we see it is very similar to the knapsack algo. in formulation.
Infact the knapsack algo is a submodular ~~We~~ ~~can~~ maximization technique.
Since we have the same properties as the knapsack algo, our algo will also be submodular.

It is of the form

maximize $f(S)$
subject to $S \in F$
for subset $S \subseteq [n]$ that satisfies a constraint represented

by a feasible set.

Here for us $f(S) = \sum_{i \in S} v_i x_i$

& $\mathcal{F} = \{ S \subseteq [n] \mid c(S) \geq K/N \}$

where $c(S) = \sum_{i \in S} w_i x_i$

Here f is submodular & monotone.

$\boxed{\text{Reference}}$ → ~~Swoh~~ Submodular function → swoh.web.engr.illinois.edu

We can implement a greedy approach to return ~~opt~~ a near optimal sol.

$$f(S) \geq (1 - 1/e) \, OPT$$

Hence our greedy will return 63% atleast.

~~The greedy approach will be →~~

~~Choose Each researcher, assign one paper~~

The greedy approach will be —

1) Choose one researcher from set $R = \{R_1, R_2 .. R_K\}$
2) Assign him 1 paper from set $x = \{x_1, x_2 ... x_n\}$
3) Repeat step 1 & 2 with diff researchers until set $\{R\} = \emptyset$
then ~~repeat~~ reset $R = \{R_1, R_2, .., R_K\}$ & continue.

Although this is a very naive solution to this, it may prove to be inefficient in the long run as there can be n! assignments.

One such solution would be to use the simplex method by breaking it down to an LP problem.

When an additional diversity constraint is added of m research labs, m<k

we still wish to maximize quality of each paper, but minimize the no. of reviewers per paper as long as they are from same research lab.

So as for our greedy approach,
here
→ first we find max no. of researchers from every lab.
→ assign first set of papers to said researches
→ Here we note, that giving ~~this~~ set back to same researchers will not improve quality, so this set is then distributed among next set of researchers
→ This process continues greedily & we should get an optimal sol. to this problem

HP

Scanned by CamScanner