Feedback vertex set

Feedback vertex set problem: input is an undirected graph with nonegative vertex weights. Output is a subset F of the vertices such that $G \setminus F$ is acyclic. Objective is to minimize the sum of the weights of vertices in F.

Theorem 1 *There is a 2-approximation for the feedback vertex set problem.*

Cyclomatic number of G: Let $\text{cyc}(G) = |E| - |V| + \kappa(G)$, where $\kappa(G)$ is the number of connected components of G. For each vertex v, let $\delta_G(v) = \text{cyc}(G) - \text{cyc}(G \setminus v)$.

Lemma 1 *For G connected, we have $\delta_G(v) = degree(v) - \kappa(G \setminus v)$.*

(The proof uses elementary arguments, see textbook.)

Lemma 2 *If H is a subgraph of G, then $\delta_H(v) \leq \delta_G(v)$ for every v.*

(The proof is by elementary induction on the number of edges of $G \setminus H$, see textbook.)

```
Algorithm:
  if G is acyclic,
    output the empty set
  else
    for each vertex compute deltaG(v)
    let c be the minimum value of w(v)/deltaG(v) over all vertices v
    for each vertex v, let w'(v)=w(v)-c deltaG(v)
    let V' be the set of vertices with nonzero w'(v) and G' be the induced graph.
    Recursively compute a minimal FVS F' for G' with weight function w'.
    using vertices of V-V' only, extend F' into a minimal FVS of G
```

(Since there is always at least one vertex such that $w'(v) = 0$, the algorithm terminates. It is easy to check that, using vertices of $V - V'$ only, one can easily extend F' into a minimal FVS of G, so the algorithm is well-defined.)

The key to the analysis is the following graph-theory statement.

Lemma 3 *Let G be a graph. Then*

- *For every FVS F, we have $cyc(G) \leq \sum_{v \in F} \delta_G(v)$.*
- *For every minimal FVS F, we have $\sum_{v \in F} \delta_G(v) \leq 2cyc(G)$.*

To prove the first statement, we consider F, label its vertices $v_1, \ldots, v_f$, and observe that $\text{cyc}(G) = \sum_{i=0}^{k-1} (\text{cyc}(G_i) - \text{cyc}(G_{i+1}))$, where $G_i = G \setminus \{v_1, \ldots, v_i\}$. We rewrite each right hand side term as $\delta_{G_i}(v_{i+1})$, which is at most $\delta_G(v_{i+1})$ by Lemma 2.

To prove the second statement, we can assume that G is connected. Let $F = \{v_1, v_2, \ldots v_f\}$, let $k = \kappa(G \setminus F)$, and let t be the number of connected components C of $G \setminus F$ such that the edges leading out of C all lead to the same vertex of F (first type components, the others being called second type components; see picture in textbook). We have:

$$\begin{aligned}
\sum_{v\in F}\delta_G(v) &= \sum_{i=1}^{k}\text{degree}_G(v_i) - \sum_{i=1}^{k}\kappa(G\setminus v_i) \text{ by Lemma 1}\\
\sum_{i=1}^{k}\text{degree}_G(v_i) &= 2|E| - 2|E(G\setminus F)| - |E(F, V\setminus F)|\\
|E(G\setminus F)| &= |V| - f - k \text{ since } G\setminus F \text{ is a forest with } k \text{ trees}\\
\sum_{i=1}^{k}\kappa(G\setminus v_i) &= f + t
\end{aligned}$$

To lowerbound the number of edges between F and $V\setminus F$, by minimality of F we have that every v_i is on a cycle of G that has v_i as its only vertex from F. By picture, such a cycle must all be inside the same connected component of $G\setminus F$, so v_i must have 2 edges e_i and e'_i leading into that connected component. Take $\{e_1, e_2, \ldots, e_f\}$: that's f edges already. By inspection of a few cases on a picture, in $G\setminus\{e_1, e_2, \ldots, e_f\}$ each first type component must have at least one edge leading out of the component and each second type component must have at least two edges leading out of the component, so that is an additional $1\cdot t + 2\cdot(k-t)$ edges. Thus:

$$|E(F, V\setminus F)| \geq f + t + 2(k-t).$$

Adding all the inequality and equalities and using the definition of $\text{cyc}(G)$ yields the desired statement.

The proof of the Theorem is now easy. Rewrite the algorithm in iterative fashion.

```
Algorithm:
  i <- 0, Gi <- G, Vi <- V, wi <- w
  While Gi has cycles:
    for each vertex of Gi compute deltaGi(v)
    let ci be the minimum value of wi(v)/deltaGi(v) over all vertices v of Gi
    for each vertex v of Gi, let w(i+1)(v)=wi(v)-ci deltaGi(v)
    let V(i+1) be the set of vertices with nonzero w(i+1)(v) and G(i+1) be the induced graph.
    i++
  let k be the final value of i
  Fk <- emptyset of vertices of Gk
  For i=k-1 down to 0
    Using vertices of Vi-V(i+1) only, extend F(i+1) into a minimal FVS Fi of Gi
Output F0
```

The cost of the output is:

$$\begin{aligned}
\sum_{v\in F_0} w(v) &= \sum_i \sum_{v\in F_i\setminus F_{i+1}} w(v) \\
&= \sum_i \sum_{v\in F_i\setminus F_{i+1}} \sum_{j\le i} c_j \delta_{G_j}(v) \\
&= \sum_j \sum_{i\ge j} \sum_{v\in F_i\setminus F_{i+1}} c_j \delta_{G_j}(v) \\
&= \sum_j c_j \sum_{v\in F_j} \delta_{G_j}(v).
\end{aligned}$$

Since F_j is a minimal FVS of G_j, by the first part of Lemma 3 we have $\sum_{v\in F_j} \delta_{G_j}(v) \le 2\text{cyc}(G_j)$. Let F^* be the optimal solution. Since $F^* \cap V_j$ is a FVS of G_j, by the second part of Lemma 3 we have $\text{cyc}(G_j) \le \sum_{v\in F^*\cap V_j} \delta_{G_j}(v)$. Thus

$$\begin{aligned}
\sum_{v\in F_0} w(v) &\le 2\sum_j c_j \sum_{v\in F^*\cap V_j} \delta_{G_j}(v) \\
&= 2\sum_{v\in F^*} \sum_{j:v\in V_j} c_j \delta_{G_j}(v) \\
&= 2\sum_{v\in F^*} w(v) \\
&= 2\text{OPT}.
\end{aligned}$$

This algoirithmic technique - decomposing the graph using a nested sequence, decomposing the weights into pieces, such that the subproblem defined by a subgraph and corresponding piece of the weight can be analyzed easily- is called layering.