

PACE Solver Description: A Heuristic Directed Feedback Vertex Set Problem Algorithm

Andrei Arhire ✉

Alexandru Ioan Cuza University of Iași, Romania

Paul Diac ✉

Alexandru Ioan Cuza University of Iași, Romania

Abstract

A feedback vertex set of a graph is a set of nodes with the property that every cycle contains at least one vertex from the set i.e. the removal of all vertices from a feedback vertex set leads to an acyclic graph. In this short paper, we describe the algorithm for finding a minimum directed feedback vertex set used by the `_UAIC_ANDREIARHIRE_` solver, submitted to the heuristic track of the 2022 PACE challenge.

2012 ACM Subject Classification Theory of computation → Randomized local search; Theory of computation → Graph algorithms analysis

Keywords and phrases directed feedback vertex set, local search

Supplementary Material The source code is available on Zenodo (10.5281/zenodo.6646187) and GitHub (<https://github.com/AndreiiArhire/PACE2022>).

1 Preliminaries

Let $G = (V, E)$ be a directed graph.

- $N_G^-(v) = \{u \in V \mid \exists(u, v) \in E \wedge \nexists(v, u) \in E\}$, $d_G^-(v) = |N_G^-(v)|$;
- $N_G^+(v) = \{u \in V \mid \nexists(u, v) \in E \wedge \exists(v, u) \in E\}$, $d_G^+(v) = |N_G^+(v)|$;
- $N_G^\pm(v) = \{u \in V \mid \exists(u, v) \in E \wedge \exists(v, u) \in E\}$, $d_G^\pm(v) = |N_G^\pm(v)|$;
- $G - v = (V \setminus \{v\}, E \cap \{\{V \setminus \{v\}\} \times \{V \setminus \{v\}\})$;
- $G \circ v = (V \setminus \{v\}, E \cap \{\{V \setminus \{v\}\} \times \{V \setminus \{v\}\} \cup \{\{N_G^-(v) \cup N_G^\pm(v)\} \times \{N_G^+(v) \cup N_G^\pm(v)\}\})$;
- G is a diclique if $\forall x, y \in V, x \neq y$ and $\{(x, y), (y, x)\} \subset E$.

We will refer to $G - v$ operation as vertex v removal (remove v together with all its adjacent edges) and to $G \circ v$ operation as vertex v merger (connect all its predecessors with all its successors and exclude v from the graph).

2 Solver Summary

The algorithm used by our solver has three main stages. In the first stage, it finds an initial solution. The second stage consists of removing redundant vertices contained in the initial solution. In the third stage, several local searches are performed based on the best solution known so far.

3 Reduction Rules

The problem can be viewed in the following way. We have to place each vertex in a set A or a set B such that at the end A is a minimum feedback vertex set and B is the acyclic remainder. A group of vertices can be placed in A or B based on some verification which can be done in a polynomial time.

The following two operations are considered to be reductions:

- If a vertex can be part of set A without affecting the solution's optimality, we remove it and eventually introduce it in the solution set.
- If a vertex can be part of set B without affecting the optimality of the solution then we merge it.

We make usage of 8 reduction rules described in [4], [2] and [3].

► **Reduction Rule 1.** *If there exists a vertex $v \in V$ and an edge $(v, v) \in E$, remove v .*

In this case, vertex v contains a self-loop, it is erased and inserted into the solution.

► **Reduction Rule 2.** *If there exists a vertex $v \in V$, $(v, v) \notin E$ with $d_G^-(v) + d_G^\pm(v) \leq 1$ or $d_G^+(v) + d_G^\pm(v) \leq 1$, merge v .*

Vertex v can have $d_G^+(v) = 0$ or $d_G^-(v) = 0$, then it can not be part of any cycle. Otherwise, $d_G^+(v) = 1$ or $d_G^-(v) = 1$, connecting all its predecessors with all its successors and excluding v from the graph (merge operation) will not affect the optimality. In both scenarios v is not considered in the solution.

► **Reduction Rule 3.** *If there exists a vertex $v \in V$, $(v, v) \notin E$, $\min(d_G^-(v), d_G^+(v)) = 0$ and $N_G^\pm(v)$ forms a diclique, remove $N_G^\pm(v)$ and merge v .*

The key observation here is that in a diclique at most one node will not be part of the solution because any two vertices form a cycle. Thus the diclique can be viewed as a single node and this way Rule 2 can be applied further.

► **Reduction Rule 4.** *If there exist vertices $u, v \in V$, $(u, v) \in E \wedge (v, u) \notin E$ and u and v are in different strongly connected components of the graph $(V, E \setminus \{E \cap \{(x, y), \forall x \in V, y \in N_G^\pm(x)\}\})$, erase (u, v) .*

Edges between vertices in different strongly connected components are not part of any cycle (otherwise the vertices would be in the same component), thus all these edges can be deleted. Furthermore, in a diclique with two vertices at least one node will be removed, so edges that are part of a diclique can be ignored when strongly connected components are computed.

► **Reduction Rule 5.** *If there exist vertices $u, v \in V$, $(u, v) \in E \wedge (v, u) \notin E$ and $(N_G^-(u) \subset \{N_G^-(v) \cup N_G^\pm(v)\}) \vee (N_G^+(v) \subset \{N_G^+(u) \cup N_G^\pm(u)\})$, erase (u, v) .*

The idea with this rule is to delete a set of edges with the property that there is no minimal cycle using them since only minimal cycles need to be broken to compute the feedback vertex set.

The following three Reduction Rules are obtained based on the Reduction Rule 3 together with the idea that at most one node in a diclique is not part of the solution.

► **Reduction Rule 6.** *If there exists a node $v \in V$, $(v, v) \notin E$ such that $\{N_G^+(v) \cup N_G^\pm(v)\}$ or $\{N_G^-(v) \cup N_G^\pm(v)\}$ forms a diclique, merge v .*

► **Reduction Rule 7.** *If there exists a node $v \in V$, $(v, v) \notin E$ and $\{N_G^-(v) \cup N_G^+(v) \cup N_G^\pm(v)\}$ can be split in two sets, $N_G^\pm(v)$ is included in one of them and each set forms a diclique, merge v .*

► **Reduction Rule 8.** *If there exists a node $v \in V$, $(v, v) \notin E$, $d_G^\pm(v) = 0$ and $\{N_G^+(v) \cup N_G^-(v)\}$ can be split in at most three sets and each set forms a diclique, merge v .*

4 Stage one

In first part of the algorithm reductions rules are applied until the graph no more supports any. If graph is not empty a promising vertex is selected to be part of the solution and it is removed. These two operations are performed until the graph becomes empty. A vertex v is considered the best candidate in the selection process if it maximizes among all the other vertices either $(d_G^+(v) + d_G^\pm(v)) \cdot (d_G^-(v) + d_G^\pm(v))$ or $d_G^\pm(v) \cdot \infty + d_G^-(v) \cdot d_G^+(v)$.

5 Stage two

After many vertex selections, some of them may become redundant in the solution i.e. they are not part of any cycle so they can be excluded from the feedback vertex set. To maximize the number of excluded vertices, we take them in the reversed order of their insertion and we introduce them in the acyclic graph. At a fixed vertex we check if the resulting graph is acyclic or not using an optimized version of the Tarjan algorithm for computing strongly connected components. If the graph is acyclic, the vertex will be erased from the solution and the graph will keep the changes. This step is presented in [3] and [1].

6 Stage three

In the last part of the algorithm, some local searches are performed [3]. A local search consists of running the same process as the one presented in stage one but on a subgraph of G . Subgraphs are obtained by removal of a certain subset uniformly random from the best feedback vertex set found so far together with a certain subset from the acyclic remainder. Additionally, a slightly modified version of the algorithm from stage two is applied to the new solution depending on the remaining time.

7 Complexity Analysis

Both selection criteria are simple but not always optimal, and add significant benefits. Implying only the number of adjacent neighbors allows some reductions to be performed efficiently using hash sets keeping in memory most of the data. After each update of the graph, reductions 1 and 2 can be applied in $\mathcal{O}(1)$. Reduction 3 is implemented in $\mathcal{O}((|V| + |E|) \cdot \log(|V|))$ and reduction 4 in $\mathcal{O}(|V| + |E|)$. The rest are ran only for vertices bounded by a small constant and we assume the complexity is $\mathcal{O}(|V| + |E|)$. The first two reductions are ran after each graph change. The others are ran after each loss of around 5% – 25% of the number of edges. For every iteration among T ones, we choose to restore around 30% of the vertices into the acyclic remainder and run the local search. As an instance can be reduced in linear time to a graph with $|V| < |E|$, the final complexity of the program is $\mathcal{O}(T \cdot (|V| + |E| \cdot \log|E|))$, based on Master's Theorem [5], where T is the number of local searches. T is a parameter bounded by the time limit.

References

- 1 Berend Hasselman. An efficient method for detecting redundant feedback vertices. CPB Discussion Paper 29, CPB Netherlands Bureau for Economic Policy Analysis, April 2004. URL: <https://ideas.repec.org/p/cpb/discus/29.html>.
- 2 Hen-Ming Lin; Jing-Yang Jou. On computing the minimum feedback vertex set of a directed graph by contraction operations. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 19:295–307, 2000. URL: <http://doi.org/10.1109/43.833199>, doi: 10.1109/43.833199.
- 3 Mile Lemaic. *Markov-Chain-Based Heuristics for the Feedback Vertex Set Problem for Digraphs*. PhD thesis, Universität zu Köln, 2008.
- 4 Hanoch Levy; David W Low. A contraction algorithm for finding small cycle cutsets. *Journal of Algorithms*, 9:470–493, 1988. URL: [http://doi.org/10.1016/0196-6774\(88\)90013-2](http://doi.org/10.1016/0196-6774(88)90013-2), doi: 10.1016/0196-6774(88)90013-2.
- 5 R.M. Verma. A general method and a master theorem for divide-and-conquer recurrences with applications. *Journal of Algorithms*, 16:67–79, 1994. URL: <http://doi.org/10.1006/jagm.1994.1004>, doi: 10.1006/jagm.1994.1004.