

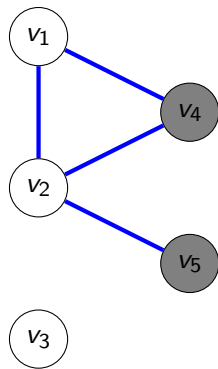
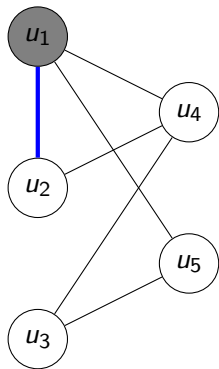
Algorithm Selection for Maximum Common Subgraph

Paulius Dilkas

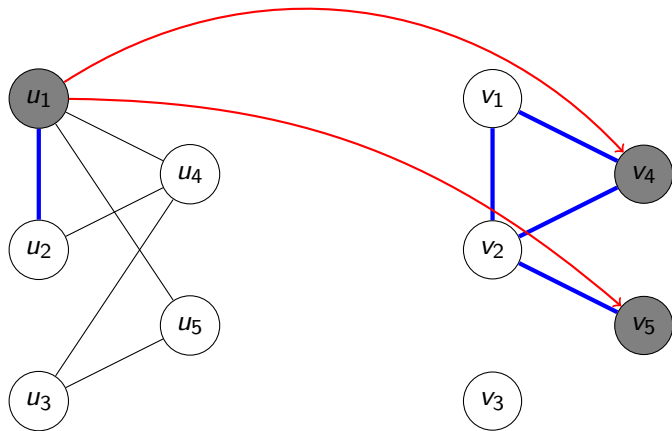
School of Computing Science
University of Glasgow

23rd March 2018

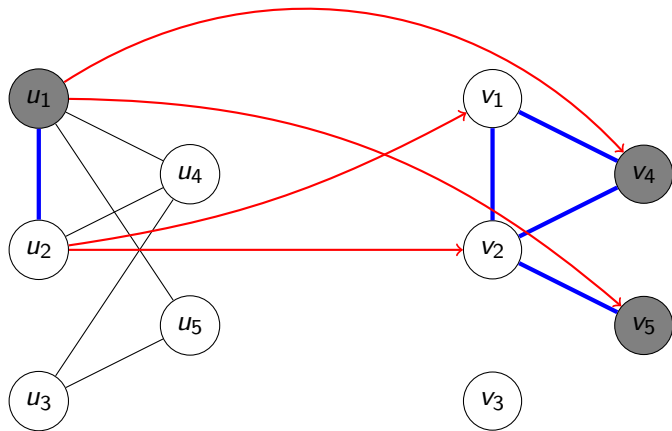
The Problem: Maximum Common Subgraph



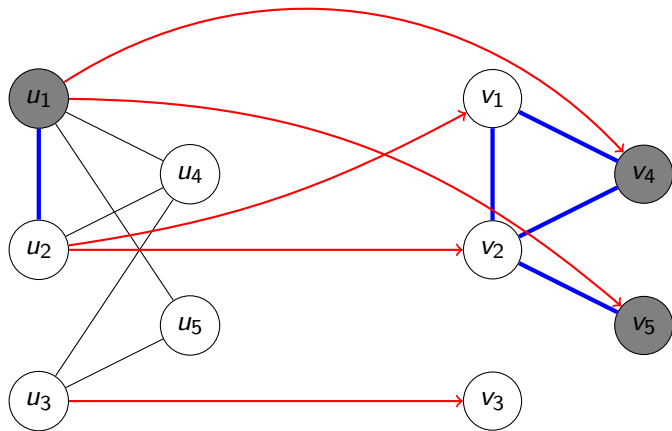
The Problem: Maximum Common Subgraph



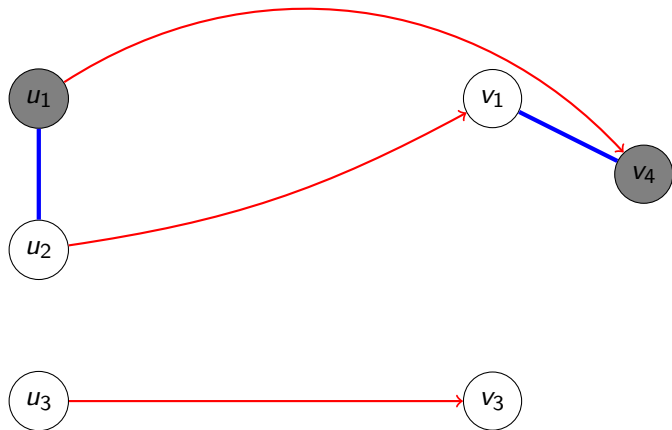
The Problem: Maximum Common Subgraph



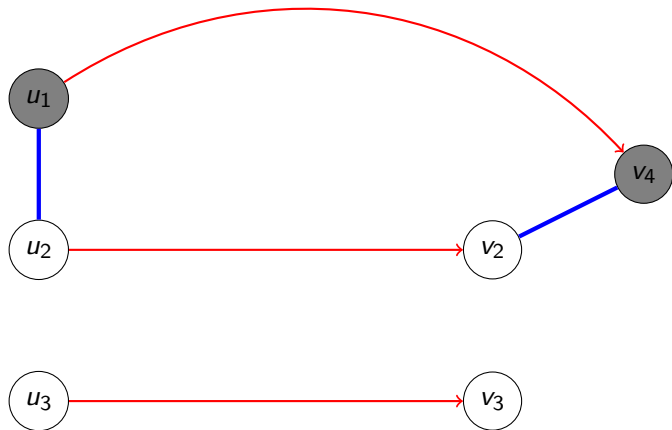
The Problem: Maximum Common Subgraph



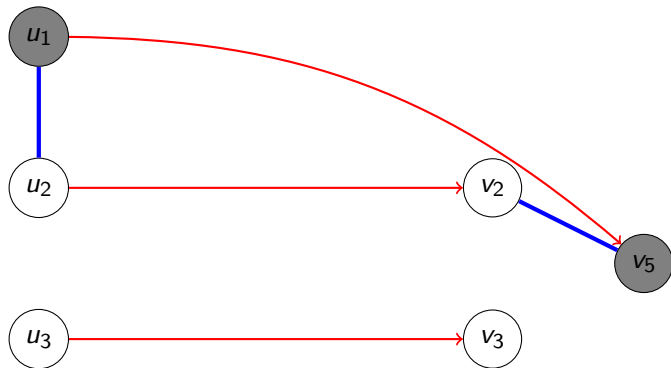
The Problem: Maximum Common Subgraph



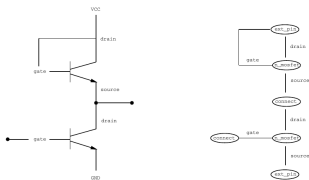
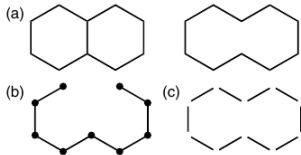
The Problem: Maximum Common Subgraph



The Problem: Maximum Common Subgraph

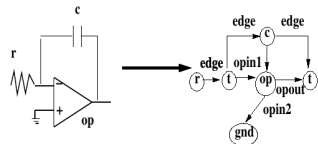


Why Is It Important?



circuit

graph representation



Existing Ways to Solve It

- ▶ MCSPLIT, MCSPLIT↓
 - ▶ McCreesh, Prosser and Trimble 2017
- ▶ clique encoding
 - ▶ McCreesh, Ndiaye et al. 2016
- ▶ k ↓
 - ▶ Hoffmann, McCreesh and Reilly 2017

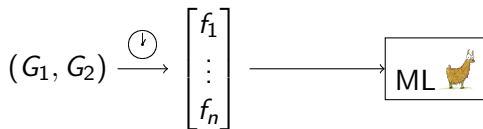
Solution: Algorithm Selection

(G_1, G_2)

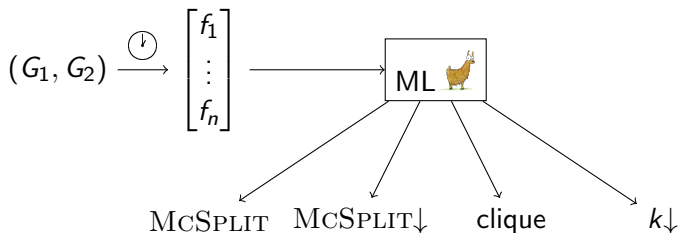
Solution: Algorithm Selection

$$(G_1, G_2) \xrightarrow{\text{⌚}} \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}$$

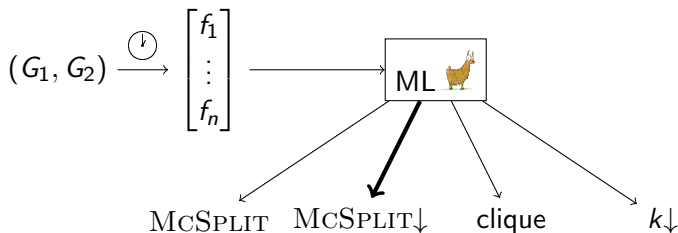
Solution: Algorithm Selection



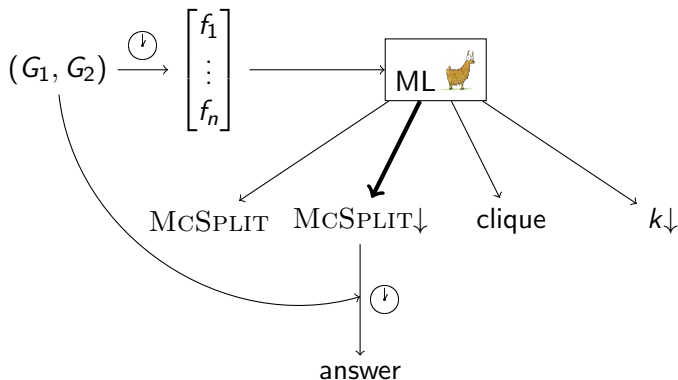
Solution: Algorithm Selection



Solution: Algorithm Selection



Solution: Algorithm Selection



Features (34 in total)

1–8 are from Kotthoff, McCreesh and Solnon 2016

1. number of vertices
2. number of edges
3. mean/max degree
4. density
5. mean/max distance between pairs of vertices
6. number of loops
7. proportion of vertex pairs with distance $\geq 2, 3, 4$
8. connectedness

Features (34 in total)

1–8 are from Kotthoff, McCreesh and Solnon 2016

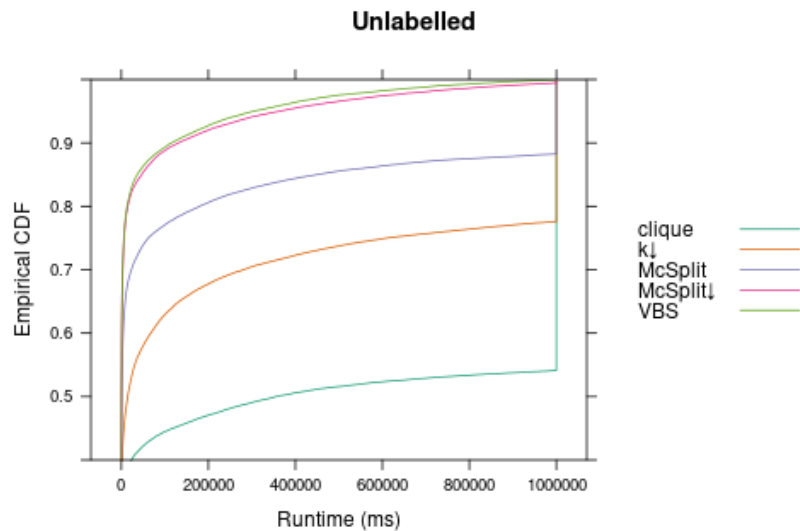
1. number of vertices
2. number of edges
3. mean/max degree
4. density
5. mean/max distance between pairs of vertices
6. number of loops
7. proportion of vertex pairs with distance $\geq 2, 3, 4$
8. connectedness
9. standard deviation of degrees
10. labelling percentage

Features (34 in total)

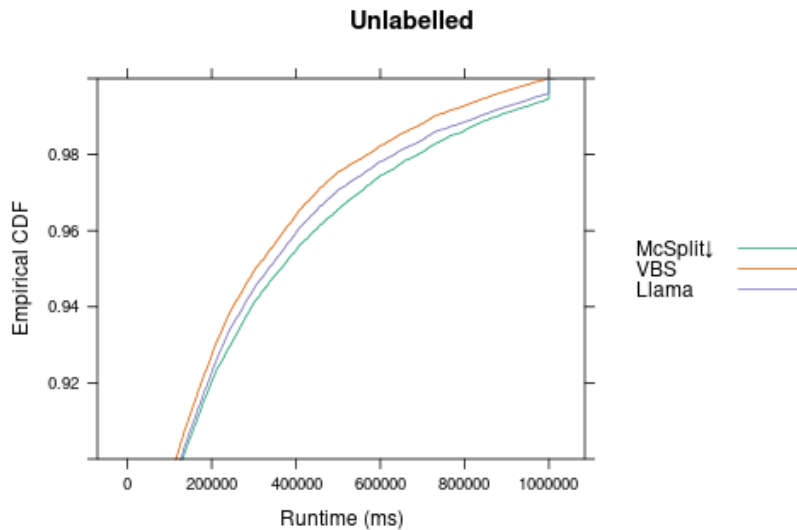
1–8 are from Kotthoff, McCreesh and Solnon 2016

1. number of vertices
2. number of edges
3. mean/max degree
4. density
5. mean/max distance between pairs of vertices
6. number of loops
7. proportion of vertex pairs with distance $\geq 2, 3, 4$
8. connectedness
9. standard deviation of degrees
10. labelling percentage
11. ratios of features 1–5

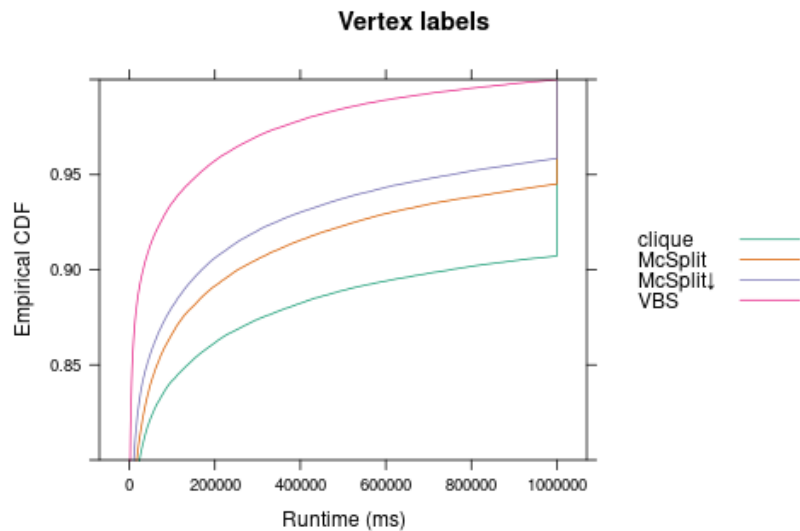
Results



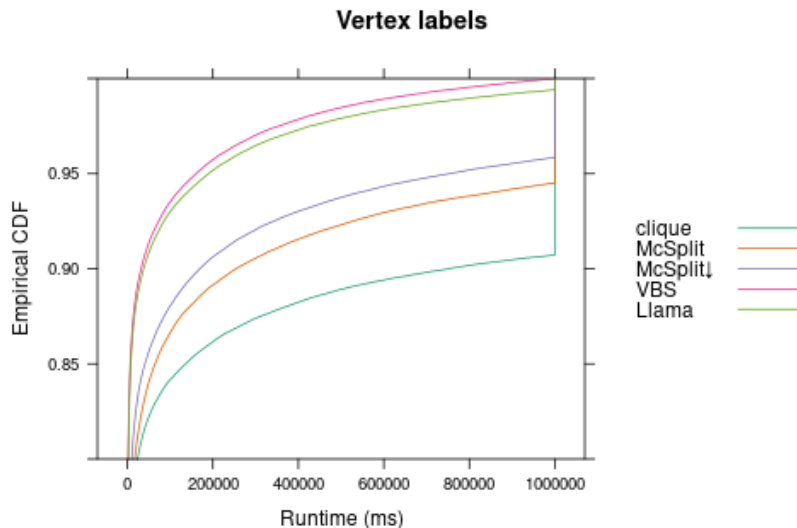
Results (27%)



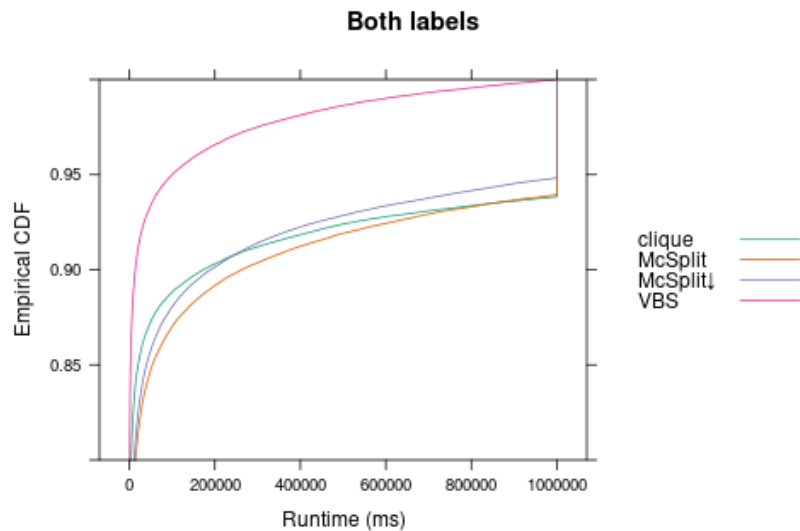
Results



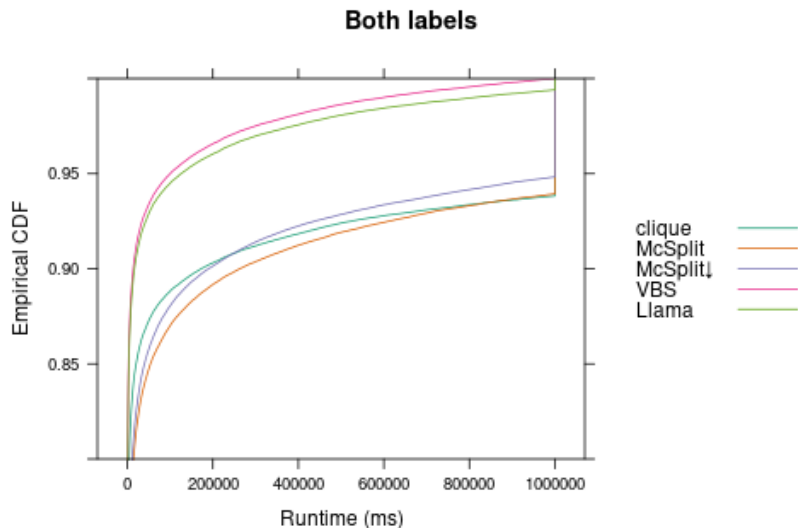
Results (86%)



Results



Results (88%)



Other Accomplishments

- ▶ Identified important features:
 - ▶ labelling, standard deviation of degrees
- ▶ Extended $k\downarrow$ to support vertex labels
 - ▶ using neighbourhood degree sequences
- ▶ Defined and developed new algorithms capable of switching between MCSPLIT and the clique encoding