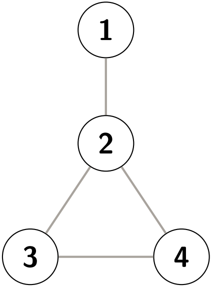
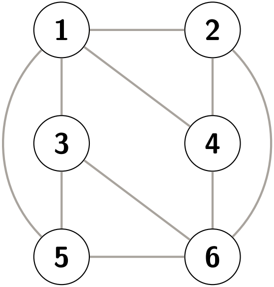
Heuristics and Really Hard Instances for Subgraph Isomorphism Problems

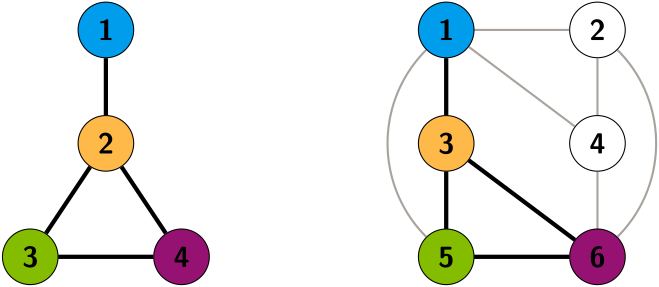
Ciaran McCreesh, Patrick Prosser and James Trimble

Non-Induced Subgraph Isomorphism





Non-Induced Subgraph Isomorphism



Benchmarking

- Based upon chemical and computer vision datasets, we can handle patterns with 1,000 vertices and targets with 10,000 vertices.
- Do these results reflect the worst case, or are they too optimistic?
- Can we create “hard” benchmark instances?

Randomly Selected Subgraphs

- Start with a random target graph.
- Pick vertices at random to make a pattern.
- Shuffle the numbering.

Randomly Selected Subgraphs

- Start with a random target graph.
- Pick vertices at random to make a pattern.
- Shuffle the numbering.
- These instances will always be satisfiable!

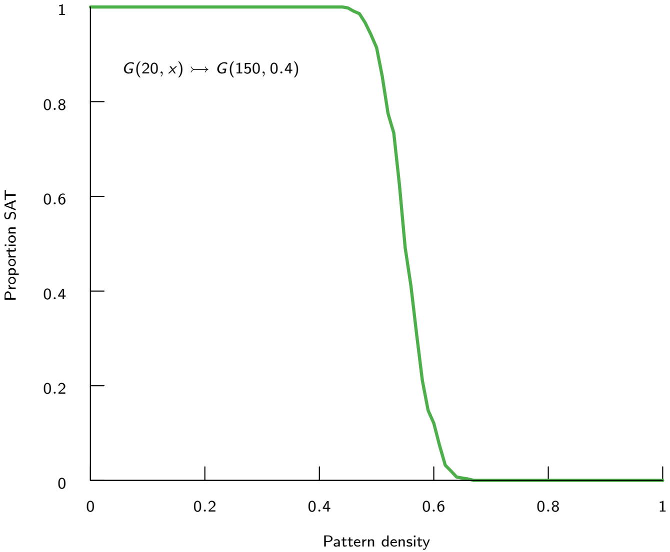
Independently Random Subgraphs

- Make a random target graph.
- Independently, make a random pattern graph.

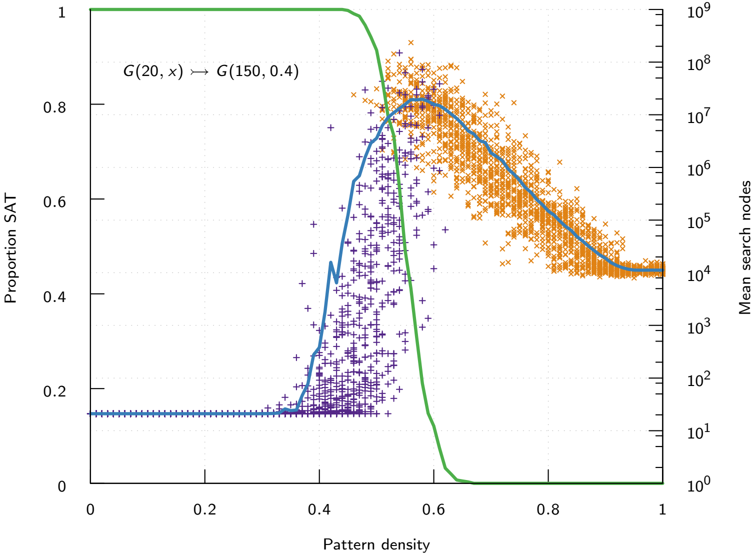
Independently Random Subgraphs

- Make a random target graph.
- Independently, make a random pattern graph.
- Will these instances ever be satisfiable?

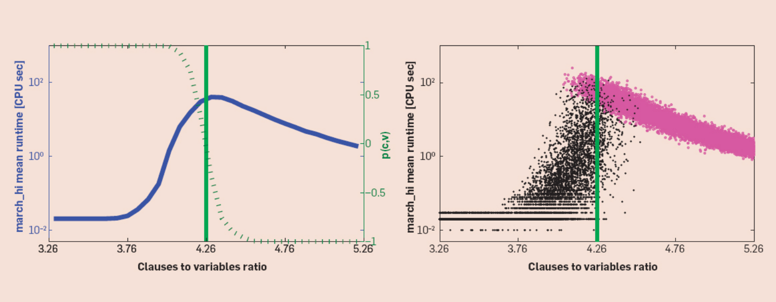
A Phase Transition



A Phase Transition

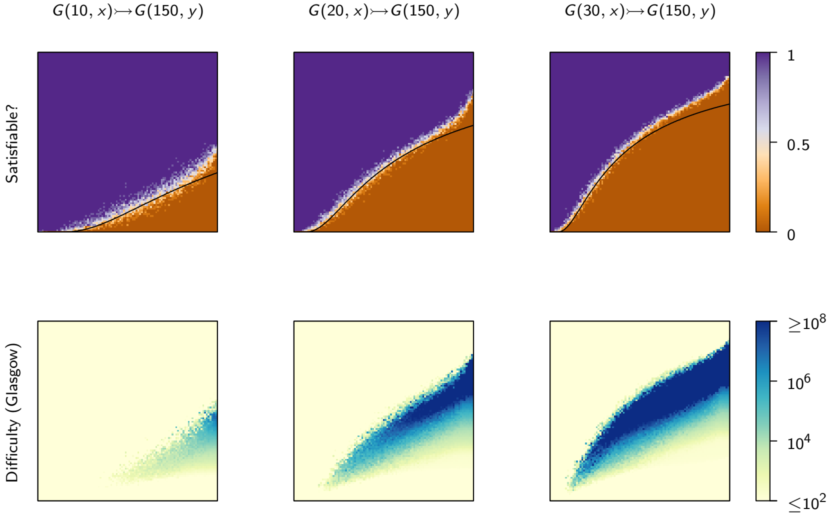


This Looks Familiar. . .



Understanding the Empirical Hardness of NP-Complete Problems.
Kevin Leyton-Brown, Holger H. Hoos, Frank Hutter, Lin Xu.
Communications of the ACM, Vol. 57 No. 5, Pages 98-107

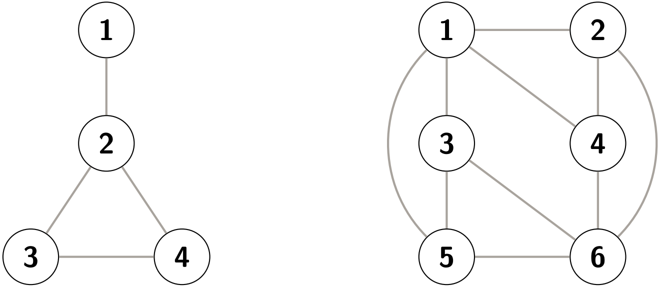
In Two Dimensions?



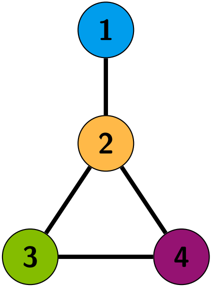
See The Paper For...

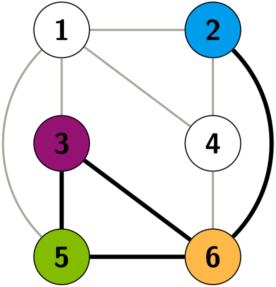
- Is this behaviour solver-independent?
- Estimating the phase transition location.
- Using this to rediscover variable and value ordering heuristics.

Induced Subgraph Isomorphism

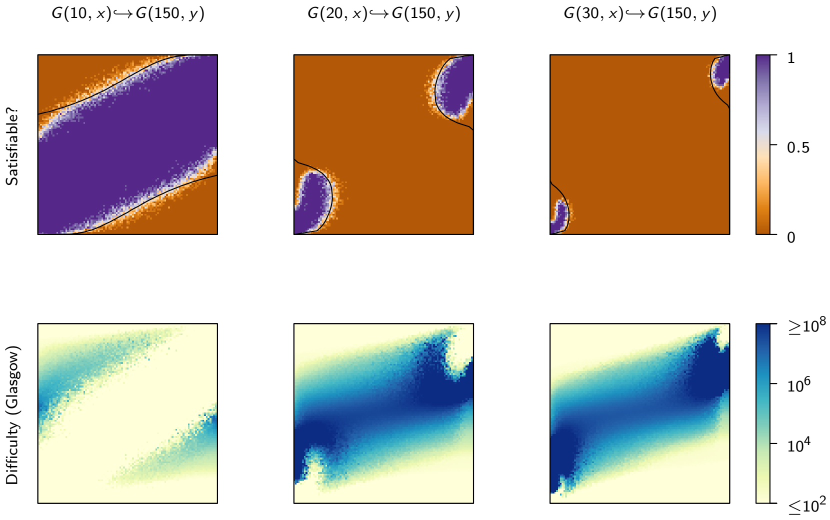


Induced Subgraph Isomorphism

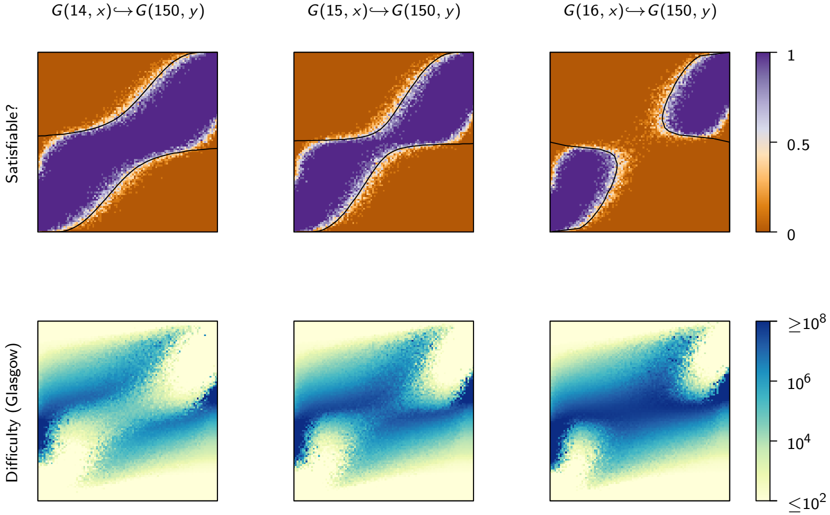




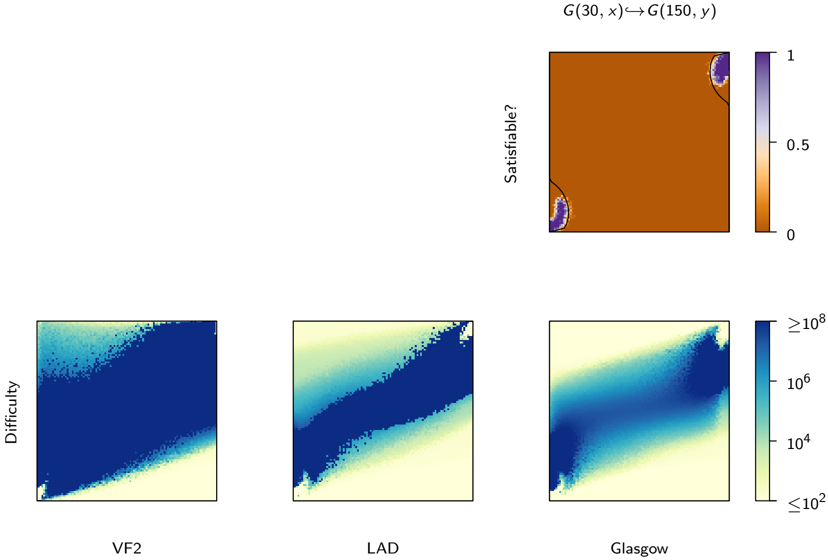
Induced in 2D



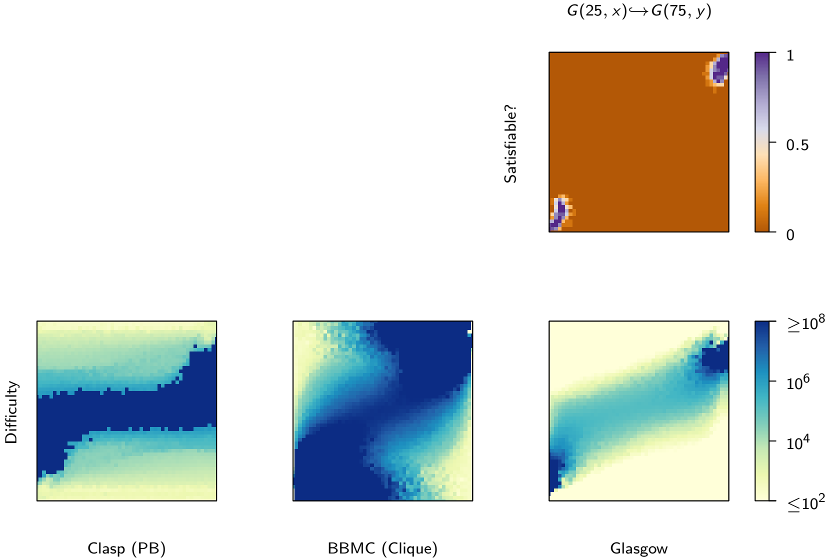
What Changes Between 10 and 20?



Is The Central Region Really Hard?



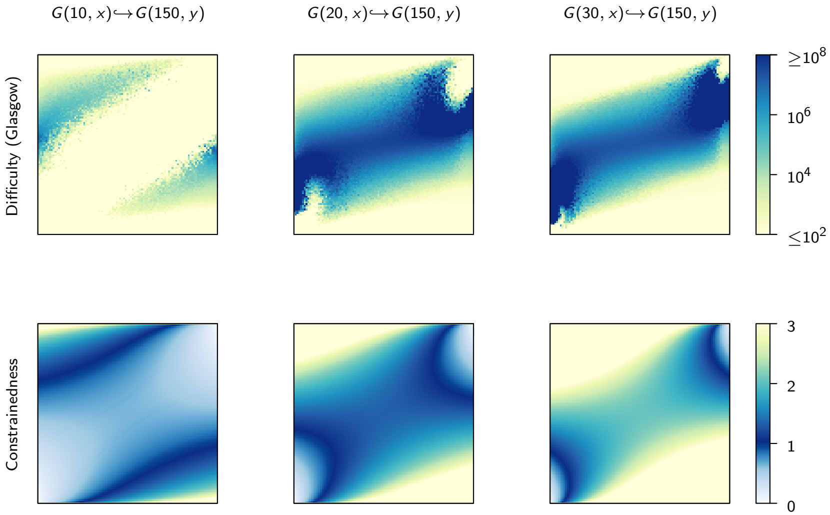
What About Encodings or Reductions?



Constrainedness

$$\kappa = 1 - \frac{\log\left(t^{\underline{p}} \cdot {d_t}^{d_p \cdot \binom{p}{2}} \cdot (1 - d_t)^{(1 - d_p) \cdot \binom{p}{2}} \right)}{\log t^{\underline{p}}}$$

Constrainedness versus Difficulty



See The Paper For...

- More on solver-independence and reductions.
- Estimating the phase transition location.
- Using this to invent new variable and value ordering heuristics.
  - But something unexpected happens this time!

Future Work

- Other randomness models (bounded degree, regular, scale-free).
- Better estimates of the phase transition location for very sparse or very dense patterns.
  - This needs a horrible variance calculation. Please get in touch if you like doing this sort of thing.
- Dynamic heuristics?

http://www.dcs.gla.ac.uk/~ciaran

c.mccreesh.1@research.gla.ac.uk