# Algorithms and Experiments for the Approximate Balanced Coloring Problem

## SIAM DS17

DAVID PHILLIPS

United States Naval Academy

Joint work with
Louis M. Pecora (Naval Research Laboratory)
Francesco Sorrentino (University of New Mexico)

May 24, 2017

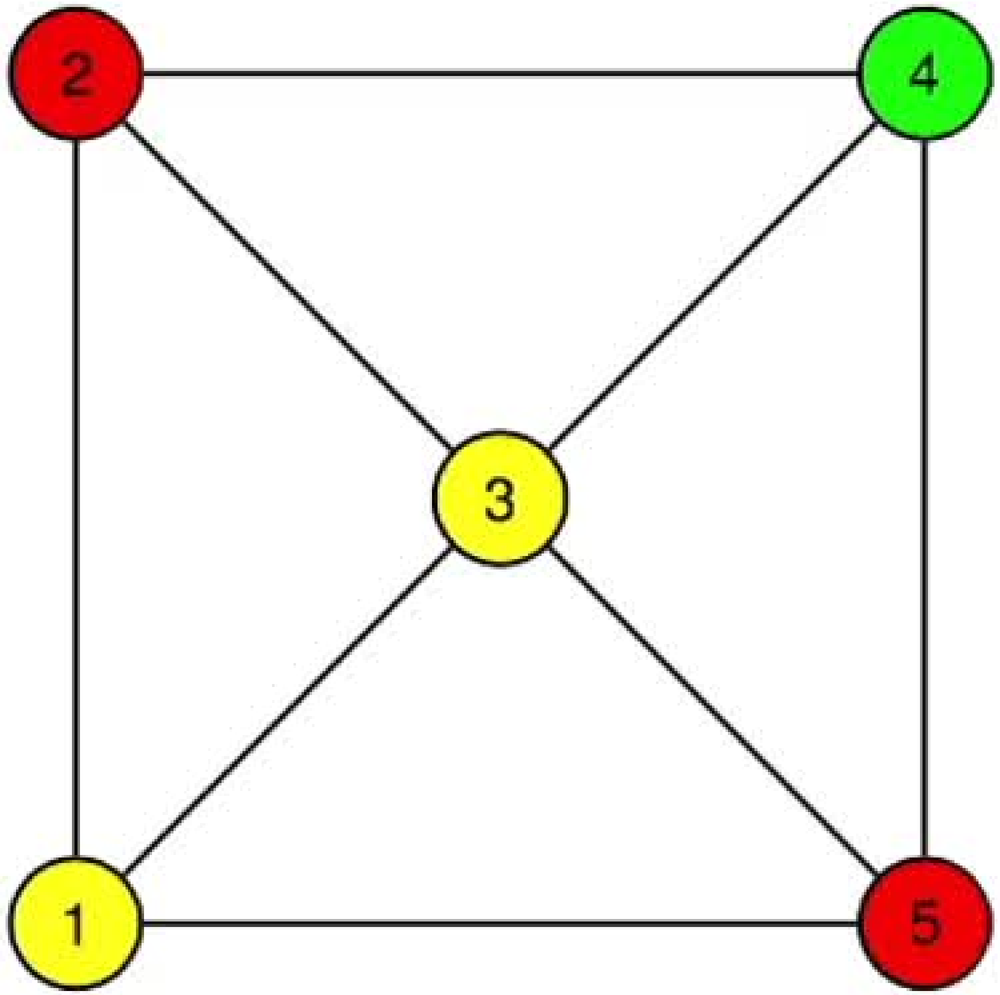
# Overview

- Describe **balanced coloring**
  - Finds synchronous clusters of nodes
  - Algorithm of Belykh and Hasler (2011) finds the **largest** synchronous clustering of nodes
  - How do we generalize this?

# Overview

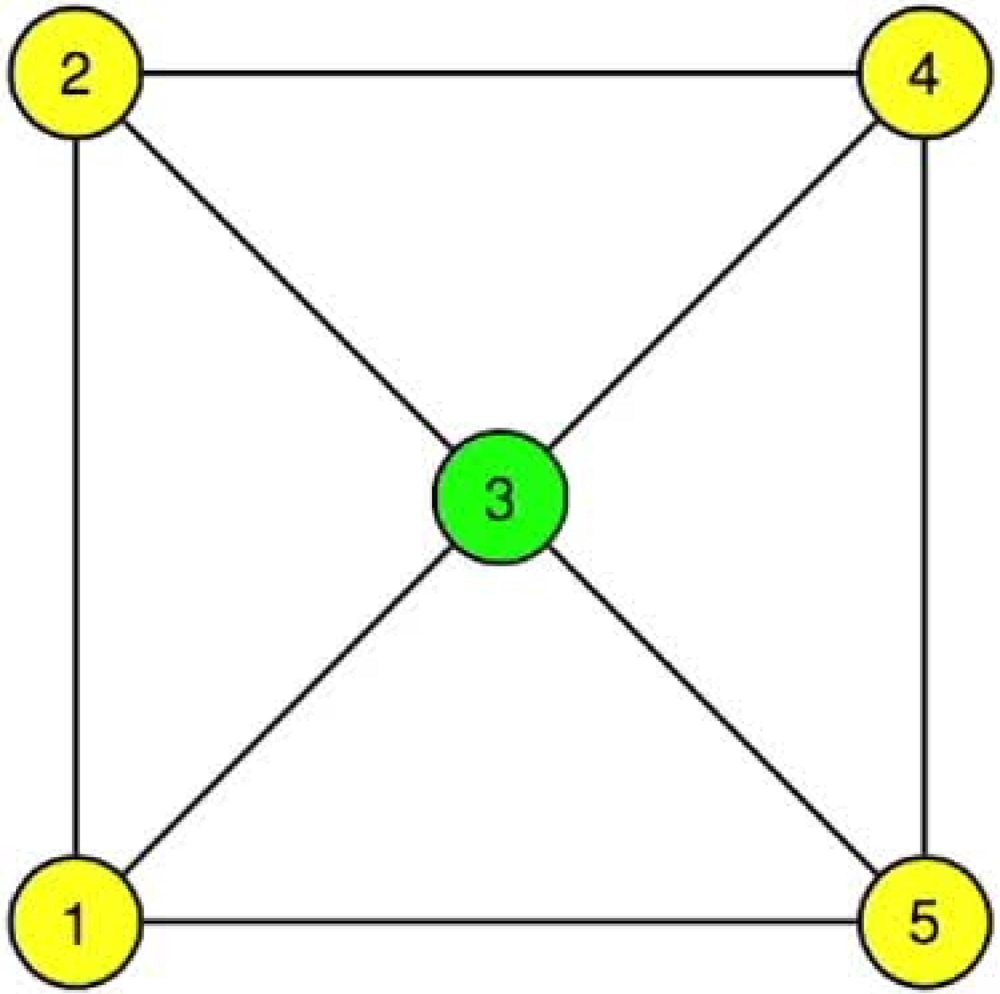
- Describe **balanced coloring**
  - Finds synchronous clusters of nodes
  - Algorithm of Belykh and Hasler (2011) finds the **largest** synchronous clustering of nodes
  - How do we generalize this?
- Apply **integer programming** to problem
  - A technique from **operations research**, and, more specifically, **mathematical programming**
  - Provides a framework to systematically model variations of a given problem
- Cross disciplinary fun with operations research!
  - **Disclaimer:** I'm not a dynamical systems expert
  - Optimization, machine learning, (big) data analytics, stochastics, control, etc
  - **Talk goal: breaking the DS/OR language barrier**

A *coloring* of $G$ is a partition, $\mathcal{C}$, of the nodes in G.



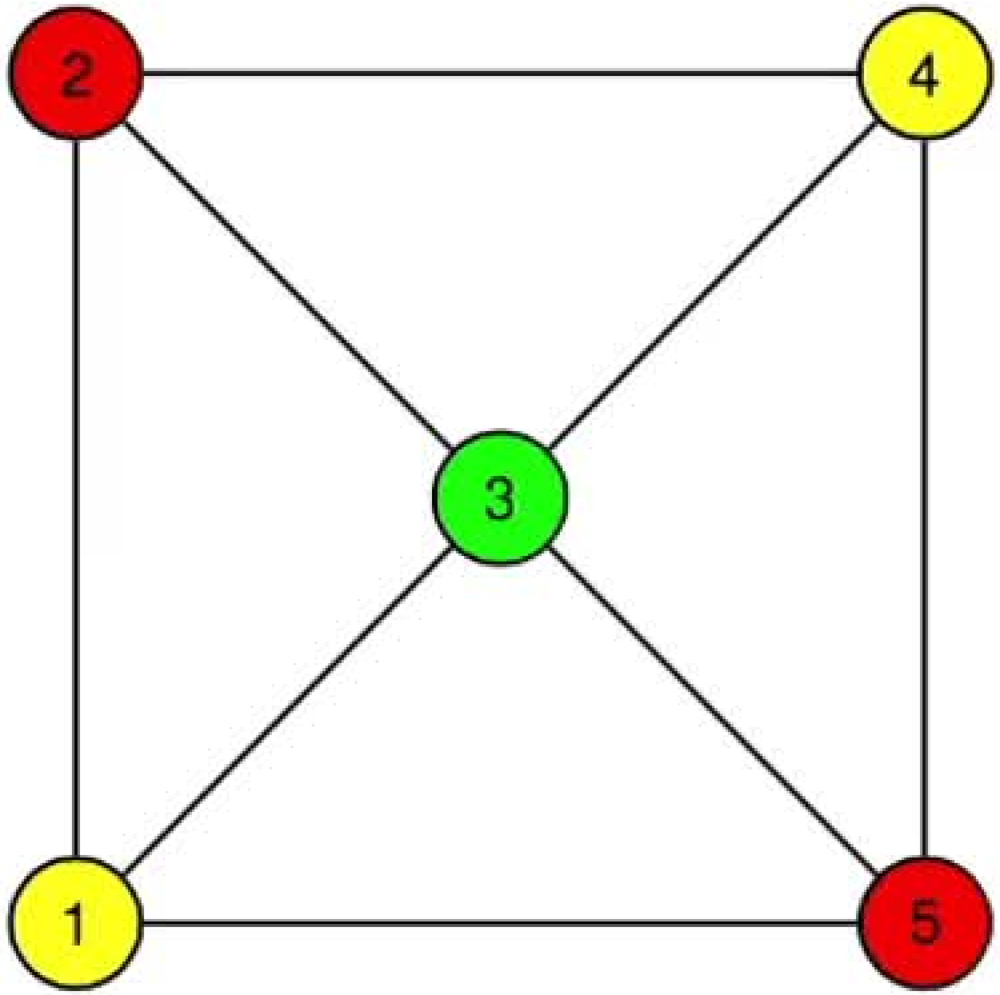

$$\mathcal{C} = \{\{1,3\},\{2,5\},\{4\}\}$$

A *balanced coloring*: for every $S, T \in \mathcal{C}$ and $u, v \in S$, the number of edges from $u$ and $v$ to nodes in $T$ is equal.

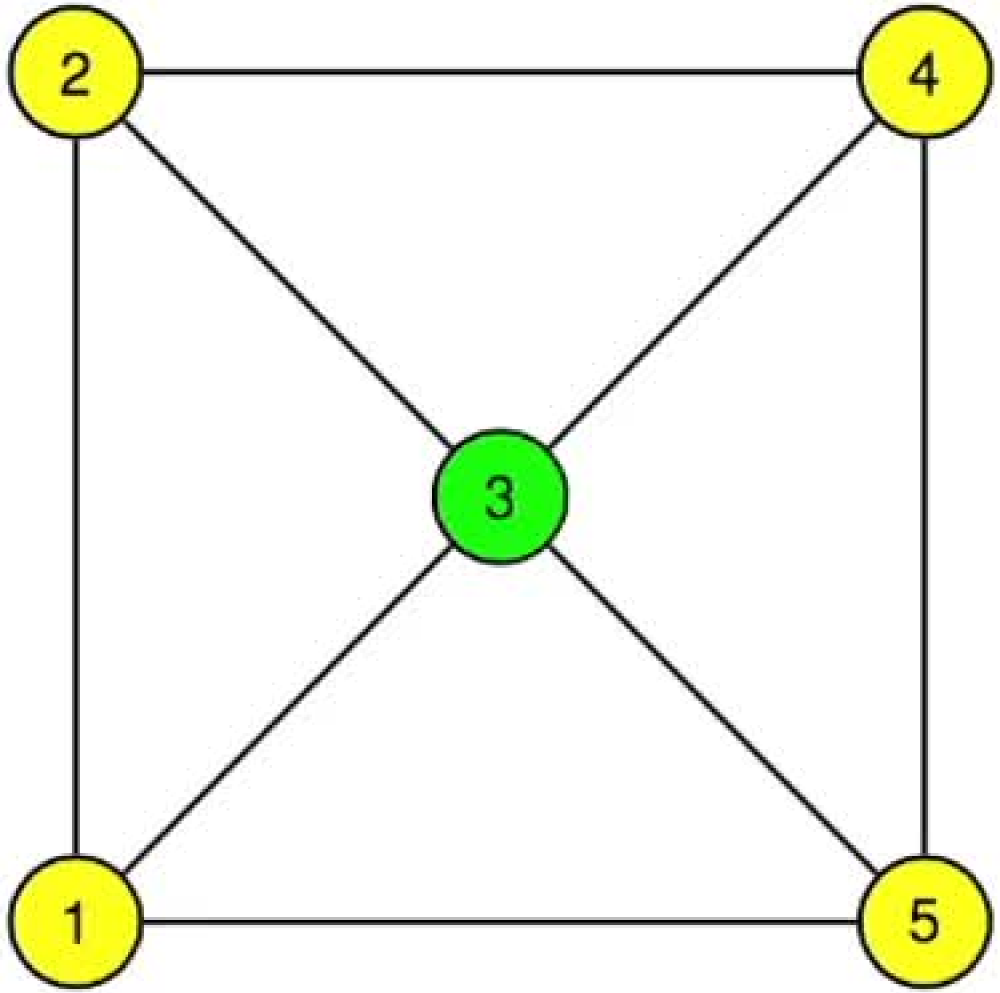



Balanced

A *minimal balanced coloring* is the balanced coloring using as few colors as possible.




Still balanced, not minimal




**Minimal balanced coloring finds the largest synchronous clusters**

- A greedy algorithm by Belykh and Hasler (2011) finds the minimal balanced coloring in polynomial time.

- What about other generalizations?
  - Approximate edge weights, isolating nodes, directed graphs, etc.

Mathematical programming:

$$\begin{array}{ll} \max & f(x) \\ \text{subject to} & x \in S \end{array}$$

$f : S \subset \mathbf{R}^n \to \mathbf{R}$

Linear programming: $f$ is linear, $S$ is a polyhedron, i.e.,
$f(x) = \sum_i c_i x_i$,
$S = \{x \in \mathbf{R}^n : Ax \leq b\}, A \in \mathbf{R}^{m \times n}, b \in \mathbf{R}^m$.
$x$ are decision variables
**Goal: find the optimal** $x$

(Linear) Integer programming: $f$ linear, $S$ is a polyhedron intersected with $Z^n$.

Implementations exist to solve linear and integer programs.

$G = (V, E), C =$ set of colors

$G = (V, E)$, $C =$ set of colors

**Decision variables**

$t \quad = \text{number of colors used}$

$z_c \quad = \text{indicator of color } c \text{ used}, c \in C$

$y_{ic} \quad = \text{indicator of } i \text{ colored } c, i \in V, c \in C$

$x_{ijcd} \quad = \text{indicator of } i \text{ colored } c, j \text{ colored } d, ij \in E, c, d \in C$

$1 + |C| + |V||C| + |E||C|^2 = O(|E||C|^2)$ variables

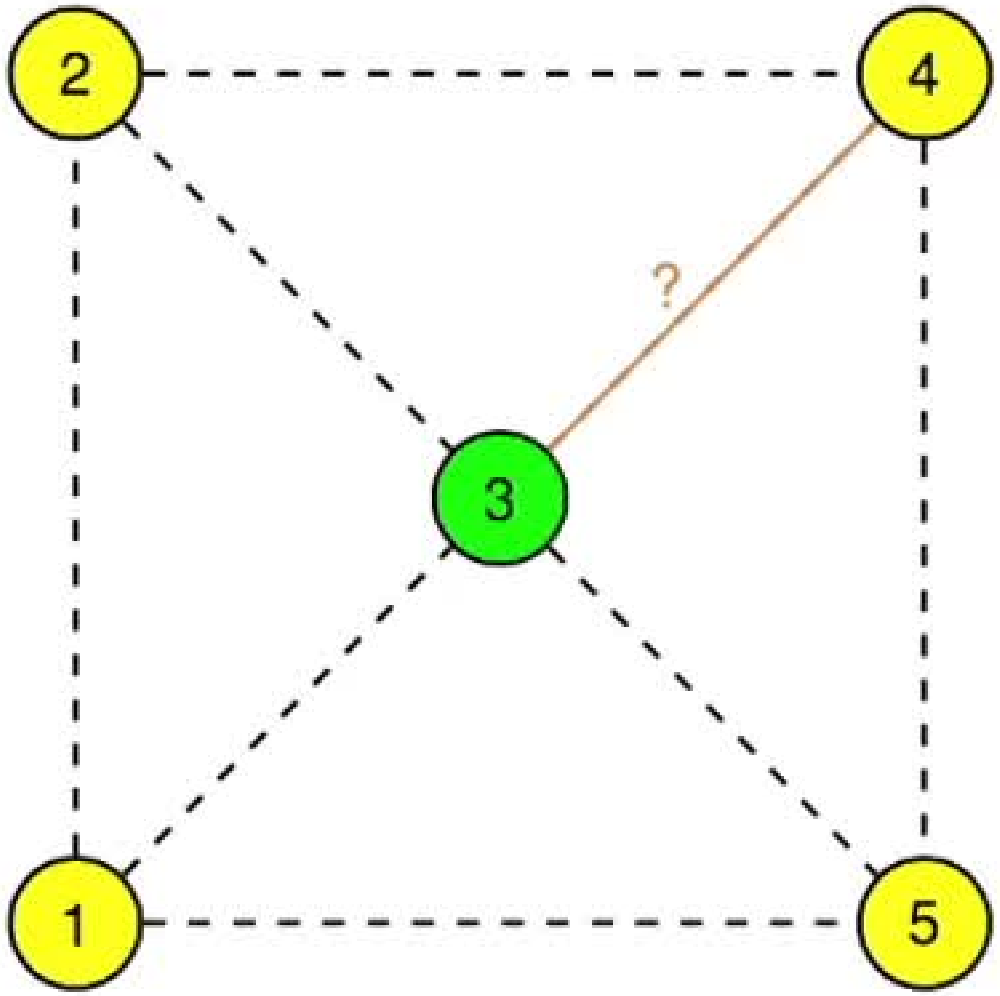
Idea: Use linear equations and inequalities to create the set of all feasible balanced colorings.

$$
\begin{array}{llll}
\min & t & & \\
\text{s.t.} & \sum_{c \in C} y_{ic} = 1 & i \in V & \text{(a)} \\
& \sum_{c \in C, d \in C} x_{ijcd} = 1 & ij \in E & \text{(b)} \\
& x_{ijcd} \leq y_{ic} & ij \in E, c, d \in C & \text{(c1)} \\
& x_{ijcd} \leq y_{jd} & (i,j) \in E, c, d \in C & \text{(c2)} \\
& \sum_{ij \in E} x_{ijcd} \leq \sum_{pq \in E} x_{pqcd} + M(1 - y_{ik}) + M(1 - y_{pk}) & i, p \in V, c, d \in C & \text{(d1)} \\
& \sum_{ji \in E} x_{jicd} \leq \sum_{qp \in E} x_{qpcd} + M(1 - y_{ik}) + M(1 - y_{pk}) & i, p \in V, c, d \in C & \text{(d2)} \\
& y_{ic} \leq z_c & i \in V, c \in C & \text{(e)} \\
& t = \sum_{c \in C} z_c & & \text{(f)} \\
& x_{ijcd}, y_{ic}, z_c \in \{0, 1\} & \forall i, j, c, d &
\end{array}
$$

$$|V| + |E| + 2|E||C|^2 + 2|V|^2|C|^2 + |V||C| + 1$$

$= O(|V|^2|C|^2)$ constraints.

# Constraint examples:

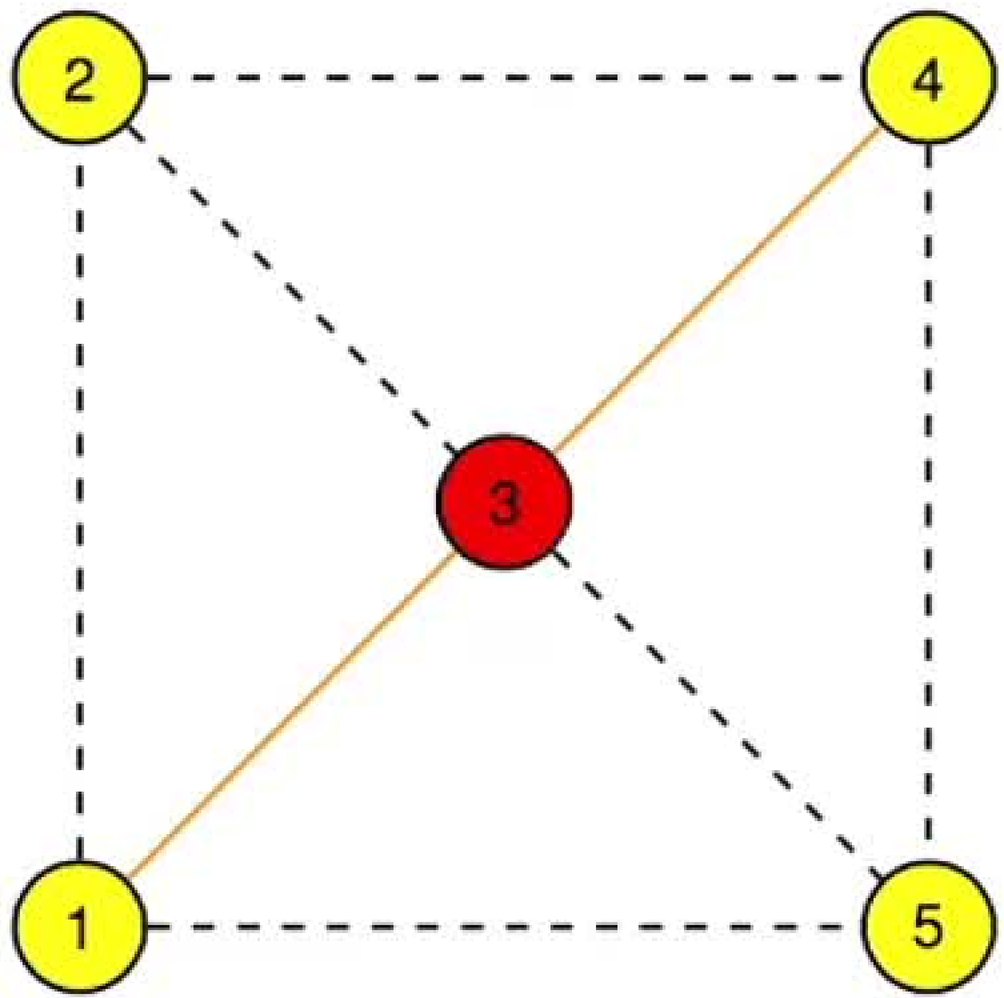



$$x_{34GR} \leq y_{3G}$$

$$x_{34GR} \not\leq y_{4R}$$

Forces $x_{34GR} = 0$ with this color setting

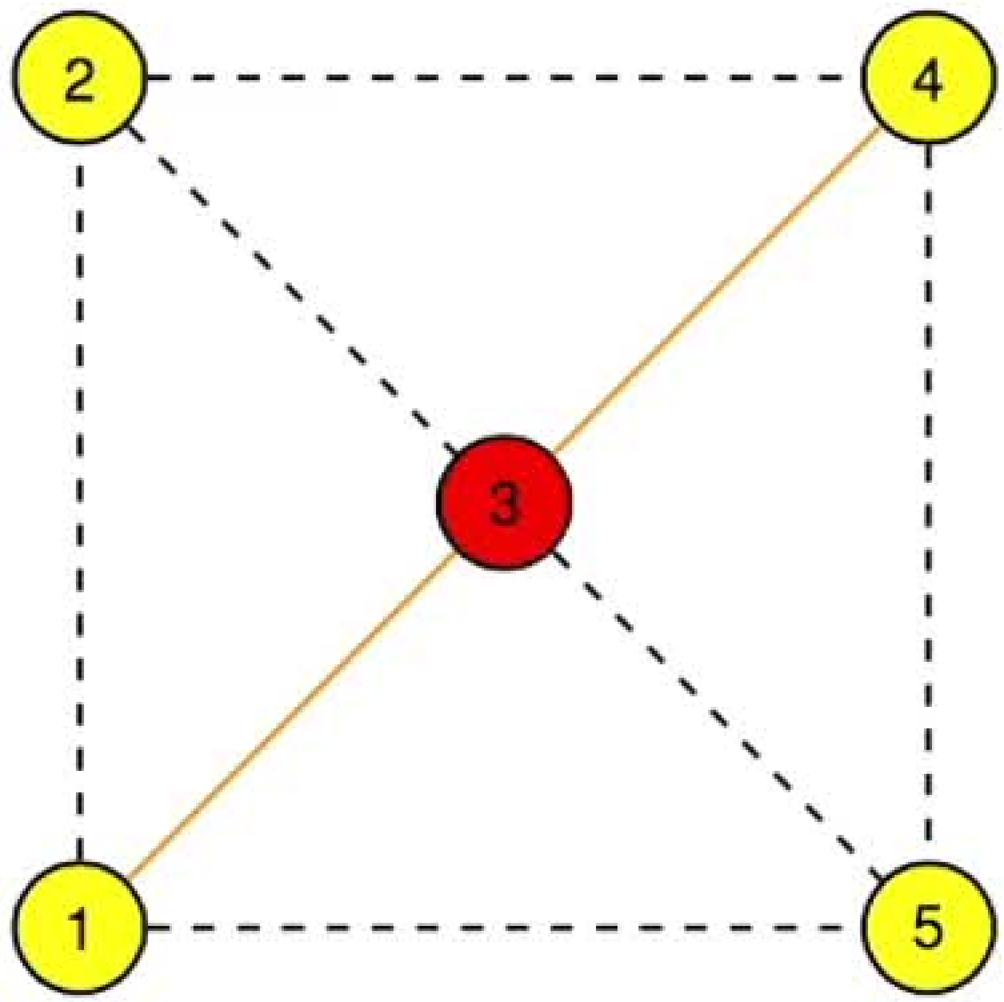
# Constraint examples:




$$x_{12YR} + x_{13YR} + x_{15YR} \leq x_{42YR} + x_{43YR} + x_{45YR} + M(1 - y_{1Y}) + M(1 - y_{4Y})$$

Symmetric constraint results in $x_{13YR} = x_{43YR}$ for this color setting

## Constraint examples:




$$x_{12YR} + x_{13YR} + x_{15YR} \leq x_{42YR} + x_{43YR} + x_{45YR} + M(1 - y_{1Y}) + M(1 - y_{4Y})$$

Symmetric constraint results in $x_{13YR} = x_{43YR}$ for this color setting

# Ongoing work

- Improve the runtime for solving the integer program
  - Can solve up to 50 node graphs tractably 😕
  - Adding families of valid inequalities
  - $O(|E||C|^2)$: Reducing the size of the problem, specifically bounding the number of colors to consider
  - More sophisticated techniques
- Extensions to model
  - Fixing the color number to fine additional colorings
  - Approximate coloring: weighting the edges
  - Solving the Laplacian case