Graph coloring

Lecture notes, vol. 2

Basics of graph theory and coloring

Lecturer: Michal Adamaszek Scribe: Giorgia L. G. Cassis

In the next pages, G is always a graph, $V(G)$ its set of vertices and $E(G)$ its set of edges.

Definition 1. *A* walk *in G from u to v ($u, v, \in V(G)$) is a sequence*

$$u = x_1, x_2, \dots, x_k = v,$$

such that $x_i x_{i+1} \in E(G)$ for all i. This walk is called closed if $u = v$. Moreover:

1. *G s called* connected *if there is a walk between any two vertices,*
2. *G is called* disconnected *otherwise.*

Definition 2. *The* distance *between u and v ($u, v, \in V(G)$) is*

$$d(u,v) := \begin{cases} \text{length of the shortest walk from } u \text{ to } v, & \text{if } \exists \text{ a walk between } u \text{ and } v \\ \infty, & \text{if } \nexists \text{ a walk between } u \text{ and } v \ . \\ 0, & \text{if } u = v \end{cases}$$

Moreover the diameter *of G is*

$$diam(G) := \max_{u,v \in V(G)} (d(u,v)).$$

Exercise 3. In the definition of connectedness and distance we can replace *walk* with *path.*

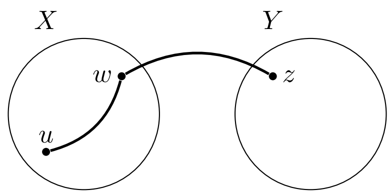
Lemma 4. *G is disconnected $\iff \exists$ non-empty sets X, Y with $V(G) = X \cup Y$, $X \cap Y = \emptyset$ such that there is no edge from X to Y.*

Proof. ($\Leftarrow$) Take $u \in X$ and $v \in Y$, then clearly $d(u,v) = \infty$ and G is disconnected.
($\Rightarrow$) If G is disconnected, then we can find $u, v \in V(G)$ with $d(u,v) = \infty$. We define

$$X := \{x \in V(G) : \ d(u,x) < \infty\},$$

$$Y := V(G) \backslash X.$$

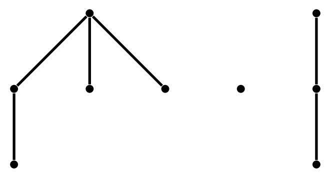
Since $u \in X$ and $v \in Y$, these two sets are non-empty. Moreover there is no edge between these two sets: assume there exist $w \in X, \ z \in Y$ such that $wz \in E(G)$, then $d(u,z) \leqslant d(u,w) + 1 < \infty$, which is a contradiction.



□

Trees

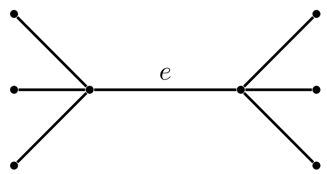
Definition 5. *A* tree *is a connected graph without cycles. A* forest *is a graph, whose every connected component is a tree.*



Theorem 6. *The following are equivalent:*

1. *G is a tree,*
2. *G is connected and $|E(G)| = |V(G)| - 1$,*
3. *Every two vertices of G are connected by exactly one path.*

Proof. ($3 \Rightarrow 1$) G is already connected by assumption, hence we must only show that G has no cycles. Assume that we an find one, then all vertices u, v from that cycle are connected by $\geqslant 2$ paths, which is a contradiction.
($1 \Rightarrow 2$) We work with induction on the number of vertices (let $|V(G)| = n$). For $n = 1$ we clearly have no edges. Assume $n \geqslant 2$. Since G is connected we can choose an edge $e \in E(G)$. Let us take a look at $G - e$ (G without e):



Removing an edge generates at most 2 connected components (call them G_1 and G_2), both without cycles. Since G was a tree, G_1 and G_2 are connected and $G - e$ is a forest with two components. We can compute

$$\begin{aligned} |E(G)| = 1 + |E(G_1)| + |E(G_2)| &\overbrace{=}^{\text{i.a.}} 1 + |V(G_1)| - 1 + |V(G_2)| - 1 = \\ &= |V(G_1)| + |V(G_2)| - 1 = |V(G)| - 1. \end{aligned}$$

($2 \Rightarrow 3$) Assume we already proved Lemma 7, then we can find a leaf $v \in V(G)$. Let w be the only neighbour of v. $G - v$ is again connected and we have

$$|E(G - v)| = |V(G - v)| - 1.$$

Working with induction (whose initial step is obvious) we can say that every two vertices in $G - v$ are connected by exactly one path, hence there is also a unique path x, w, v for every $x \in V(G - v)$. $\square$

Lemma 7. *Any tree has at least 2 leaves, provided that it has at least 2 vertices.*

Proof. Let G be a tree with n vertices and $n - 1$ edges. Since G is connected, then

$$\forall v \in V(G) : \ deg(v) \geqslant 1.$$

Suppose that we only have 1 leaf, then

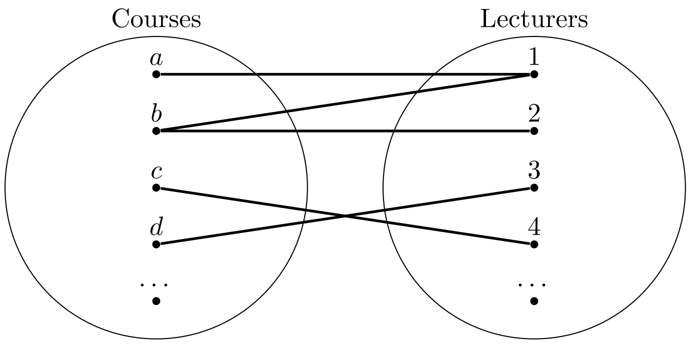
$$\underbrace{2(n-1)}_{\substack{\text{all other vertices} \\ \text{have degree} \geqslant 2}} + \underbrace{1}_{\text{leaf}} \leqslant \sum_{v \in V(G)} deg(v) = 2(n-1),$$

which is a contradiction. The same follows if we allow no leaves. $\square$

Bipartite graphs

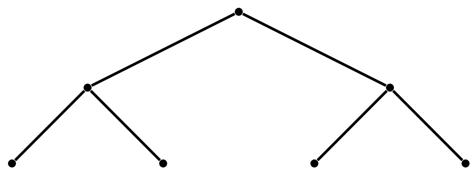
Definition 8. *G is* bipartite *if there are two sets A and B with $V(G) = A \cup B$, $A \cap B = \emptyset$ such that every edge of G has one end in each part.*

Example 9. We can use bipartite graph to illustrate a scheduling problem.



Example 10.

- Complete bipartite graphs ($K_{n,m}$): two sets A and B (with $|A| = n$ and $|B| = m$) with all the possible edges in between.
- Trees:



- n-cubes graphs (Q_n): $V_n := \{(x_1, \ldots, x_n) : \; x_i \in \{0, 1\}\}$ and then define

$$A_n := \{(x_1, \ldots, x_n) \in V_n : 2 \mid \sum_{i=1}^{n} x_i\},$$

$$B_n := \{(x_1, \ldots, x_n) \in V_n : 2 \nmid \sum_{i=1}^{n} x_i\}.$$

  Moving on one edge clearly changes the parity of the sequence, hence $V_n = A_n \cup B_n$ defines a bipartite graph.

- Cycles: C_{2n} is bipartite, C_{2n+1} is not bipartite (for $n \in \mathbb{N}$).

Theorem 11. (König) *G is bipartite iff there is no closed walk of odd length.*

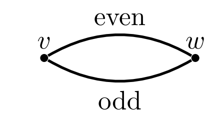
Proof. ($\Leftarrow$) If we start in A, after an odd number of steps we will be in B (since each step brings to the other side of the graph).
($\Rightarrow$) Assume that G is connected (if not, work separately in each connected component). Pick $v \in V(G)$ and define

$$A := \{w \in V(G) : \exists \text{ an even-length walk } v \to w\},$$

$$B := \{w \in V(G) : \exists \text{ an odd-length walk } v \to w\}.$$

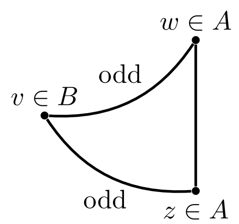
- $A \cup B = V(G)$, since G is connected.

- $A \cap B = \emptyset$, since otherwise we could find the following situation for a $w \in A \cap B$:



This means we would have an odd closed walk.

- There is no edge between two vertices in A (resp. B), otherwise we would have the following situation:



□

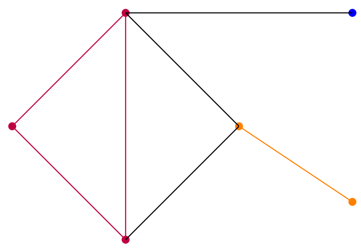
Exercise 12. In König's theorem we can replace *closed walk* with *cycle* (and even *induced cycle*).

Cliques and independent sets

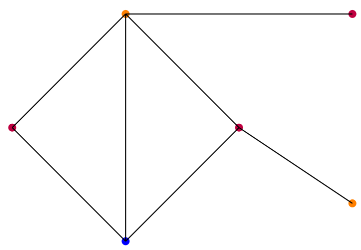
Definition 13.
- *$W \subseteq V(G)$ is a* clique *if for all $x, y \in W$, $x \neq y$, we have $xy \in E(G)$,*
- *$W \subseteq V(G)$ is an* independent set *if for all $x, y \in W$, $xy \notin E(G)$.*

Remark 14. A clique is a subgraph that is a complete graph.

Example 15. The followings are examples of cliques.



Example 16. The followings are examples of independent sets.



Definition 17. • *The* clique number *of G, $\omega(G)$, is the cardinality of the biggest clique in G,*

- *The* independence number *of G, $\alpha(G)$, is the cardinality of the biggest independent set in G.*

Observation 18. *1. $\alpha(G) = \omega(\overline{G})$.*

Proof. A clique in G is an independent set in $\overline{G}$ and vice versa. □

2. If G is bipartite, then $\omega(G) \leqslant 2$.

Proof. If there is a triangle, then two of the vertices have to be in the same part of the bipartite graph, which is impossible by definition. □

3. $\omega(G) \leqslant 2$ does not imply that G is bipartite.

Proof. C_5 (which is not bipartite since it is an odd cycle) has $\omega(C_5) = 2$. □

Definition 19. *Graphs with $\omega(G) \leqslant 2$ are usually called triangle-free.*

Example 20. • $\omega(C_n) = \begin{cases} 3, & \text{if } n = 3 \\ 2, & \text{if } n \geqslant 4 \end{cases}$.

- $\alpha(C_n) = \begin{cases} \frac{n}{2}, & \text{if } 2 \mid n \\ \frac{n-1}{2}, & \text{if } 2 \not\mid n \end{cases}$.
- G bipartite (i.e. $V(G) = A \cup B$), then $\alpha(G) \geqslant max\{|A|, |B|\}$.
- $\omega(K_n) = n$ and $\alpha(K_n) = 1$.

Remark 21. α and ω are hard (meaning NP-hard) to compute.

Lemma 22. $\alpha(G) \geqslant \frac{|V(G)|}{1+\Delta(G)}$.

Proof. We must find an independent set of size at least $\frac{|V(G)|}{1+\Delta(G)}$ and we will work by induction in order to do that. Take an arbitrary vertex $v \in V(G)$ and consider the graph

$$G' := G - v - N_G(V).$$

Clearly $\Delta(G') \leqslant \Delta(G)$. Using the induction assumption, G' has an independent set X of size at least

$$\frac{|V(G')|}{1 + \Delta(G')}.$$

Consider $X \cup \{v\}$, which (by construction) is an independent set in G.

$$|X \cup \{v\}| = |X| + 1 \geqslant \frac{|V(G')|}{1 + \Delta(G')} + 1 \geqslant \frac{|V(G)| - (1 + \Delta(G))}{1 + \Delta(G)} + 1 = \frac{|V(G)|}{1 + \Delta(G)},$$

thanks to $\Delta(G') \leqslant \Delta(G)$ and considering that $1 + \Delta(G)$ is the biggest number of vertices we can take away while defining G'. □

Algorithm 23. (The greedy algorithm) The following loop iterates $\geqslant \frac{|V(G)|}{1+\Delta(G)}$ times.

1. Take $v \in V(G)$,
2. Remove v and $N_G(v)$ (at most $1 + \Delta(G)$ vertices),
3. Iterate using $G' := G - v - N_G(V)$.

Lemma 24. *If $f : G \longrightarrow H$ is a graph homomorphism, then $f^{-1}(v)$ is an independent set for all $v \in V(H)$.*

Proof. Suppose that $x, y \in f^{-1}(v)$ and $xy \in E(G)$. Then, since f is a graph homomorphism, we would have $f(x)f(y) \in E(H)$, which is impossible since $f(x) = f(y) = v$. □

Vertex colouring

Goal. Assign colours to the vertices of a graph, such that adjacent vertices have different colours (the problem is optimized by finding the minimal number of colours needed for this process).

Application 25. 1. Colouring a land-map, such that the countries that share a border have different colours. In this case the vertices represent the countries and we draw an edge whenever two countries share a border.

2. Scheduling the timetable for an exam session. In this case the vertices represent the different subjects offered and we draw an edge whenever there is at least a student following both courses. Minimizing the number of colours means programming the minimal number of exam-spots, such that each student can take all the exams he/she applied for.

Example 26. Take a graph G with 81 vertices forming a 9×9 grid, such that the following conditions are satisfied:

- Every row is a clique,
- Every column is a clique,
- If we divide the grid forming 9 3×3 subgrids, each of these is a clique.

Then each possible colouring represents the solutions of a Sudoku.

Definition 27. *A* (vertex-)colouring *of G with colour set C is a function*

$$c : V(G) \longrightarrow C,$$

such that if $xy \in E(G)$, then $c(x) \neq c(y)$.

Definition 28. *The* chromatic number $\chi(G)$ *is the smallest number k for which there is a colouring of G with k colours. If $\chi(G) \leqslant k$, then G is called k*-colourable.

Definition 29. *Each $c^{-1}(i)$ is called a* color class.

Example 30.
- $\chi(P_n) = 2$,
- $\chi(C_n) = \begin{cases} 2, & \text{if } 2 \mid n \\ 3, & \text{if } 2 \nmid n \end{cases}$,
- $\chi(K_n) = n$,
- $\chi(Q_n) = 2$, since Q_n is bipartite. Moreover, χ of every bipartite graph with at least one edge is 2.

Lemma 31. *1. $\chi(G) \leqslant |V(G)|$,*

2. $\chi(G) = |V(G)| \iff G$ is complete,

3. $\chi(G) = 1 \iff G \simeq \overline{K_n}$,

4. $\chi(G) = 2 \iff G$ is bipartite and has at least one edge.

Proof. 1. $c : |V(G)| \longrightarrow |V(G)|$ with $c(v) = v$ is a colouring with exactly $|V(G)|$ colours.

2. One direction is obvious. Let us assume G is not complete, then we can find $x, y \in V(G)$ with $xy \notin E(G)$. Set $c(x) = c(y)$ and colour all the other $|V(G)| - 2$ vertices with different colours: this is a colouring with $|V(G)| - 1$ colours.

3. One direction is obvious. For the other look at $c : V(G) \longrightarrow \{1\}$ with $c(x) = c(y)$ for all $x, y \in V(G)$. This implies $xy \notin E(G)$ for all $x, y \in V(G)$, hence $G \simeq \overline{K_n}$.

4. One direction is obvious. For the other, if $\chi(G) = 2$ then it suffices to define $A := c^{-1}(1)$ and $B := c^{-1}(2)$ to find the partition of the set $V(G)$.

□

Remark 32. Already determining whether $\chi(G) \leqslant 3$ or $\chi(G) \geqslant 4$ is hard.