Exercise 2.1. A bipartite graph is a simple graph $G = \langle V, E \rangle$, where $V$ can be partitioned into two sets $X, Y$ (i.e., $X \cup Y = V$, and $X \cap Y = \emptyset$), such that every edge $\{a, b\} \in E$ coincides with both $X$ and $Y$, i.e., $\{a, b\} \cap X \neq \emptyset$ and $\{a, b\} \cap Y \neq \emptyset$.

Show that the following are equivalent:

1. $G = \langle V, E \rangle$ is bipartite.
2. $G$ is 2-colourable, i.e., there is a map $c : V \rightarrow \{0, 1\}$ such that no two adjacent vertices $a, b$ have the same colour, i.e., $c(a) \neq c(b)$ for all $\{a, b\} \in E$.
3. $G$ does not contain a cycle $v_1 \xleftarrow{e_1} v_2 \xrightarrow{e_2} \cdots \xleftarrow{e_{n-1}} v_n \xrightarrow{e_n} v_1$ of odd length.

Exercise 2.2. Write a program that reads a graph in N-Triples format and checks whether the graph is bipartite. Use this program to decide whether `authorship.nt.gz` and `coauthors.nt.gz` are bipartite.

Hint: each of the uncompressed graphs is roughly 4 GiB in size. In Python, you can use `gzip.GzipFile` to process the compressed file without decompressing it first. There is also `authorship-snippet.nt.gz`, a small part of the graph that you can use during development.

Exercise 2.3. From the `coauthors.nt.gz` graph, extract the connected component containing `<http://dblp.uni-trier.de/pers/s/Studer:Rudi>`, i.e, extract the induced subgraph that

- contains `<http://dblp.uni-trier.de/pers/s/Studer:Rudi>`,
- contains all nodes reachable from `<http://dblp.uni-trier.de/pers/s/Studer:Rudi>` by some path, and
- contains all edges that are present in the full graph between these nodes.

Hint: `authorship-snippet.nt.gz` contains `<http://dblp.uni-trier.de/pers/s/Studer:Rudi>` and can be used for testing during development.

---

1https://github.com/knowsys/Course-Knowledge-Graphs/tree/master/test-data/dblp
2https://docs.python.org/3/library/gzip.html