Directory Structure of the Visual Graph Arena Dataset

This document illustrates the directory structure of the Visual Graph Arena (VGA) dataset. The VGA dataset is structured to facilitate easy access to various graph analysis tasks including isomorphism, path problems, and cycle detection. Each category is further subdivided into training and testing datasets with detailed sample counts and label files. For each input data, 'image_{i}.png' corresponds to the i-th row of it's 'label.txt'. For more information and to access the dataset, please visit the VGA dataset website at vga.csail.mit.edu.

Graph Arena
- isomorphism
  - easy
    - inputs
      - train —— *140,000 samples*
      - test —— *15,718 samples*
    - labels
      - train —— *labels.txt*
      - test —— *labels.txt*
  - hard
    - inputs
      - train —— *127,374 samples*
      - test —— *14,298 samples*
    - labels
      - train —— *labels.txt*
      - test —— *labels.txt*
- path
  - hamiltonian
    - inputs
      - train —— *25,000 samples*
      - test
        - kawai —— *2,480 samples*
        - planar —— *2,480 samples*
    - labels
      - train —— *labels.txt*
      - test —— *labels.txt*
  - shortest
    - inputs
      - train —— *80,000 samples*
      - test
        - random —— *8,672 samples*
        - kawai —— *8,672 samples*
        - planar —— *8,672 samples*
    - labels
      - train —— *labels.txt*
      - test —— *labels.txt*
- cycle
  - hamiltonian
    - inputs
      - train —— *69,935 samples*
      - test
        - kawai —— *7,740 samples*
        - planar —— *7,740 samples*
    - labels
      - train —— *labels.txt*
      - test —— *labels.txt*
  - cordless
    - inputs
      - train —— *80,000 samples*
      - test
        - kawai —— *6,484 samples*
        - planar —— *6,484 samples*
    - labels
      - train —— *labels.txt*
      - test —— *labels.txt*