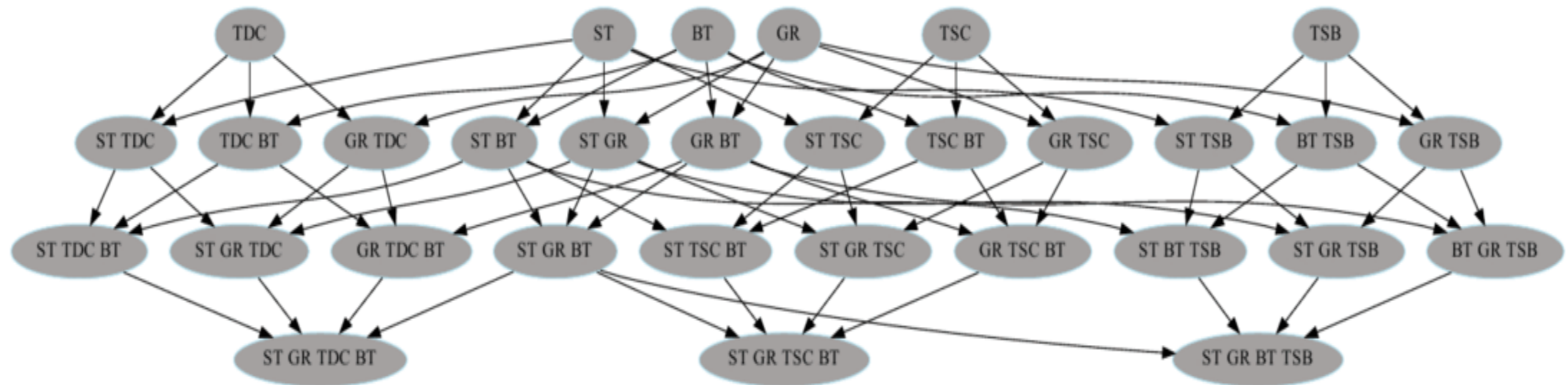# Knowledge Identification

In order to customize the curriculum, for each skill sets, we need to identify whether the trainee has acquired it.
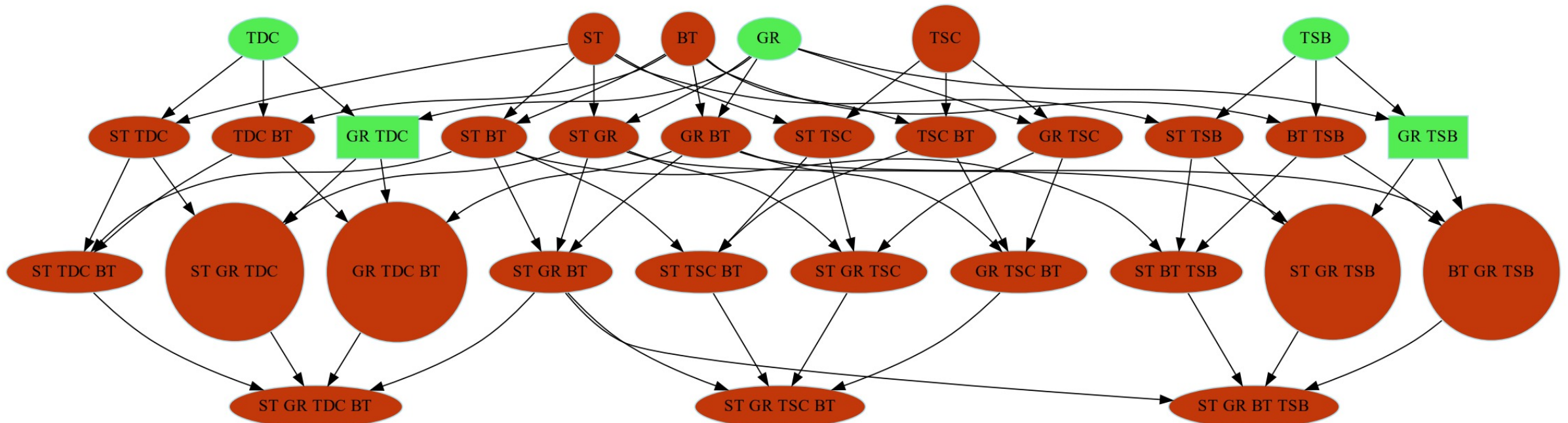
Process of Knowledge Identification

1. Get knowledge graph
2. Choose next sampling point
3. Generate the current knowledge graph as a .pdf

Keep iterating until the graph is fully colored

Why does the order of sampling matter?

- Because we can make use of two assumptions:
  - All children of a red node will be red (fail)
  - All parents of a green node will be green (pass)

1. Get knowledge graph

- The program gets a knowledge graph demonstrating the relationship of different skill sets. The whole graph will be uncolored which means we know nothing about the trainee's knowledge state
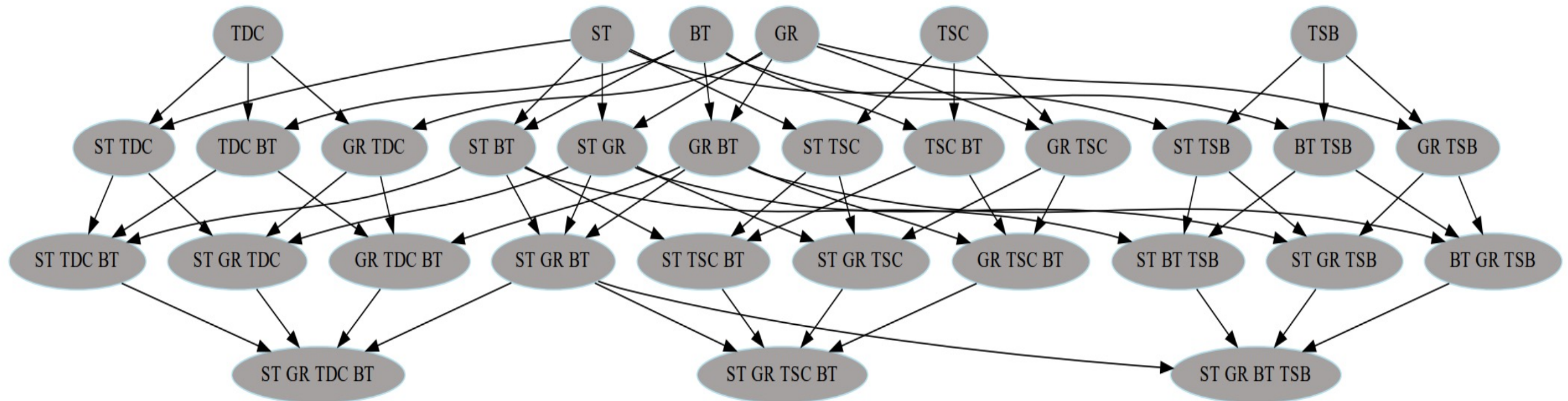
2. Choose next sampling point

- Calculates the potential updated nodes for each uncolored node using one of the functions below:

$$Argmax(n+, n-)$$
$$Argmax(\min(n+, n-))$$
$$Argmax(\max(n+, n-))$$

n+ means the additional number of colored nodes if we color the node green (pass), n- means the additional number of colored nodes if we color the node red (fail)

- Then we will choose the node with the highest potential to sample

3. Generate the current knowledge graph

- Using Graphviz package to create a colored and shaped tree.
  - Newly sampled node will be in rectangle.