DMU1

October 14, 2015

```
In [1]: function findNextParent(G, nodes, data, i, n, parents)
            bestParent = 0;
            bestScore = -Inf;
            for possibleParent = 1:n # check candidates to add as parent
                if !any(possibleParent.==parents) # if not already contained in parents list
                    addEdge!(G, nodes[possibleParent], nodes[i]);
                    if isValid(G) # if valid network
                        possibleScore = logBayesScore(G,data);
                        if possibleScore > bestScore # if best candidate for new parent
                            bestParent = deepcopy(possibleParent);
                            bestScore = deepcopy(possibleScore);
                        end
                    end
                    removeEdge!(G, nodes[possibleParent], nodes[i]); # restore network to how it was
                end
            end
            return bestScore, bestParent
        end

Out[1]: findNextParent (generic function with 1 method)
```

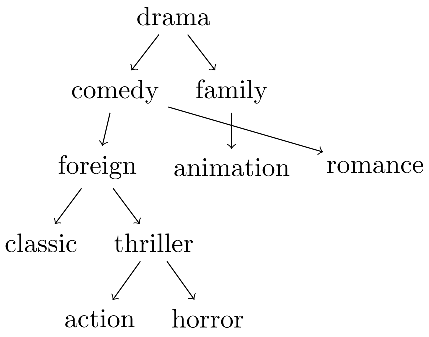
```
In [2]: function saveFile(Gbest, nodes)
            f = open("medium.gph","w")

            for i = 1:length(nodes)
                parentNodes = parents(Gbest,nodes[i])
                for j = 1:length(parentNodes)

                    #text = "parent1,child1\n"

                    write(f,parentNodes[j])
                    write(f,",")
                    write(f,nodes[i])
                    write(f,"\n")

                end
            end

            close(f)
        end

Out[2]: saveFile (generic function with 1 method)
```

```
In [4]: using BayesNets
        using DataFrames
        data = readtable("medium.csv");
        nodes = names(data);

        n = length(nodes);
        restarts = 10;
        bestScoreOverall = -Inf;
        bestG = BayesNet(nodes);
        oldScore = 0;
        bestScoreThisRestart = 0;

        for r = 1:restarts

            # order nodes
            nodes = shuffle(nodes);
            G = BayesNet(nodes);

            for i = 1:n; # which child node to work with
                parents = [0]; # start with no parents

                while true # try adding parents
                    oldScore = logBayesScore(G,data)
                    newScore, bestParent = findNextParent(G, nodes, data, i, n, parents)
                    if newScore > oldScore && bestParent != 0
                        addEdge!(G, nodes[bestParent], nodes[i]);
                        parents = hcat(parents, [bestParent])
                    else
                        break
                    end
                end

            end

            bestScoreThisRestart = oldScore;
            println(bestScoreThisRestart)

            if bestScoreThisRestart > bestScoreOverall
                bestScoreOverall = deepcopy(bestScoreThisRestart);
                bestG = deepcopy(G);
            end

        end

        println()
        println(bestScoreOverall)

        saveFile(bestG, nodes)

        return bestG
```

```
-52153.23458763232
-52164.23798175722
-52169.23590860684
-52159.74169203544
```

```
-52158.163375637596
-52137.043947180384
-52156.499954059436
-52149.89305436439
-52149.56094475941
-52120.08215124806

-52120.08215124806
```

Out[4]:


drama
comedy
family
foreign
animation
romance
classic
thriller
action
horror