

Tree from my program (output continue)

## Disadvantages:

- You have to implement everything by yourself
- Pre-created tools contain A LOT of features which by yourself would take months to implement
- You can Google or Stackoverflow questions about the tool
- You can get support if you have any troubles with the tool
- The tool must likely will have maintenance and new versions with improvements
- Pre-created tools have more confidence than student created programs

Another part of the lab was to look for a dataset to use as input in WEKA and our program and compare the trees generated by both.

The criteria I chose for choosing this dataset was:

- I must have the task of classification
- The data should be something interesting
- The number of variables shouldn't be too much because WEKA visual representation could get pretty complicated
- Number of variables from 4 to 10
- Test cases could be any

Following this criteria, I ended up founding a small dataset about an experiment on the capability of inflating a balloon given different conditions. The *arff* file can be found in my repo but the attributes of this data set were:

@attribute color: {YELLOW, PURPLE}

@attribute size: {LARGE, SMALL}

@attribute act: {STRETCH, DIP}

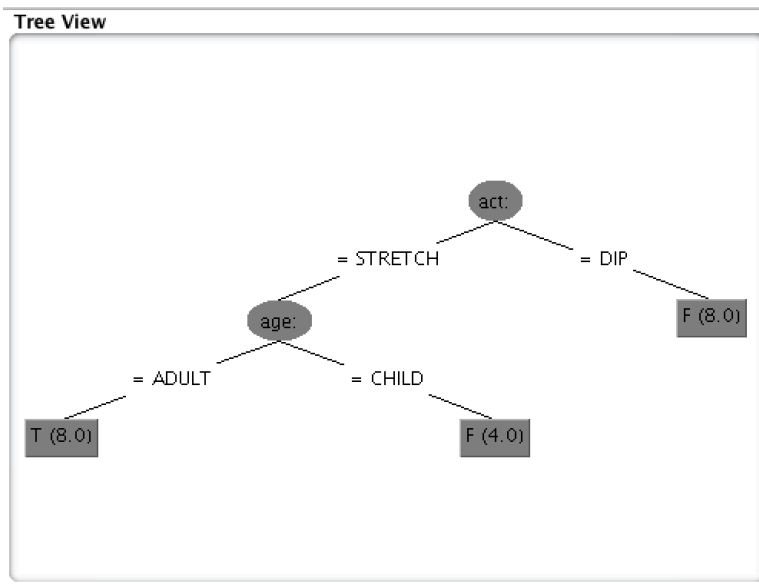
@attribute age: {ADULT, CHILD}

@attribute inflated: {T, F}

The data has a size of 20 statements.

The representations obtained were:

## From WEKA



## From my program

```
act: STRETCH
  age: ADULT
    ANSWER: T
  age: CHILD
    ANSWER: F
act: DIP
```

```
ANSWER: F
```

You can notice that both represent the same decision tree.

## **Conclusion Where would you use decision trees?**

After this lab I've noticed just the start of the huge power of using decision trees. It is a very widely use method pro classification and prediction. This has been very interesting for me because decision trees allow you to approach a problem in a structured way to arrive at a logical decision. I would say you could use decision trees in mostly everything, from deciding plans for the weekend given known and unknown factors to very complex business problems like predicting risks for stock drops. Also another area I would say it would be very interesting to see would be psychology, because there are a lot of attributes in human behavior and a decision tree would be an amazing tool to predict possible depressive disorders for example.