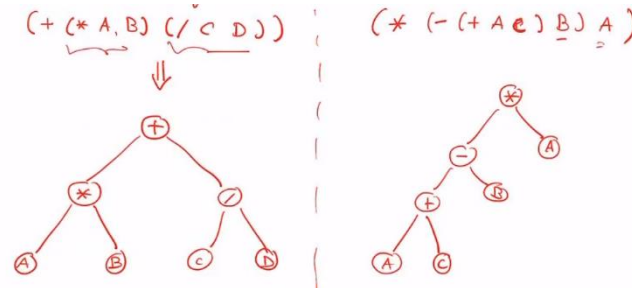


22- Genetic Programming: tree representation, initialization, crossover, mutation, and bloat.

Tree representation: -> it's an S-expression in a visual form, easier to understand the order of operations

(S-expression: $(+ x 2) \rightarrow x+2$)



INITIALIZATION:

- As always, the population is generated randomly with values from the terminal set as well as function set. With this tree representation, the terminal sets are the leaves, and the function set is the nodes
- We can set a probability p , to draw from F set, and $1-p$ to draw from T set
- Probability p will influence the depth of the tree
- It is wise to have a dynamic probability as we go deeper in the tree, at the beginning favouring a draw from F set, and later favouring a draw from T set

CROSSOVER:

- We exchange at random the sub-trees of the parents at random

MUTATION:

- We chose a random node and rebuild randomly the sub-tree from this node
- We can change the mutation probability depending how deep we are in the tree

BLOAT:

By doing crossover and mutation our functions may become more and more complex

In practice we should prevent this effect by having max length parameters in place

Sometimes it is possible that a subtree can be simplified such as $(- x x)$