Means - Ends Analysis : (Mixture of Backwar & forward search tech)

→ Search strategies either forward or Backward.

→ Mixed Strategy: Solve major probe then solve the smaller then combine them.

→ process centers around finding the diff bet current state & goal state.

→ solve recursively.

MEA Algo:

Step1: Compare current to goal, if no diff then return success & exit.

Step2: Select the most significant difference & reduce it

→ Select operator O else failure.

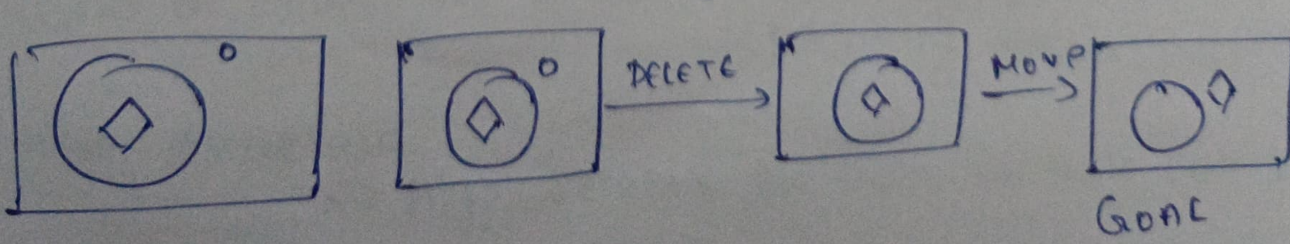
→ operator O to CURRENT (O-START), (O-RESULT)

→ if (FIRST- PART ← ---- MEA(Current, O-START)

AND

(LAST- PART ← ---- MEA (O- RESULT, GOAL)

Return the result combining FIRST-PART, O & LAST-PART



PLANNING (FULLY)

→ <u>Planning</u>: → the task of coming up with a sequence of actions that will achieve a goal is called planning.

→ <u>planning</u> <u>environment</u>: (P.E)

classical P.E → fully observable

Non-Classical P.E → partially.

→ <u>Difficulty</u> in <u>Prob. Solving</u> <u>agent</u>:

→ perform irrelevant actions → No prob decomposition

→ finding heuristic solution.

→ <u>Planning</u> prob

→ Domain Model (Defines Action)

→ initial state (actions yet to take place)

→ final state (plan to achieve)

→ find the sequence of actions, achieve the goal from a gn initial world state

→ <u>Simple</u> <u>Planning</u> <u>Agent</u>:

→ An agent interacts with the real world with perceptions & Actions.

perceptions ↓ sense the world & assess the situation

Actions → wat agent does to the Domain

STRIPS:

→ planner used in Shakey (1st robots in AI)
→ Action - centric representation (each action to effect)

Strips planning prob:

Specifies
→ An initial state S
→ A goal state G
→ A set of strips actions

STRIPS representation:

→ precondition: set of assignment of values to features that must be true for actions to occur

→ Effect: set of resulting assignment of values to change the result of action.

Reinforcement Learning Systems

- Reinforcement learning systems have 4 main elements:
  - Policy
  - Reward signal
  - Value function
  - Optional model of the environment

KAMAPURAM SRM

Policy

- A policy is a mapping from the perceived states of the environment to actions to be taken when in those states
- A reinforcement learning agent uses a policy to select actions given the current environment state

RAMAPURAM SRM

On-policy versus Off-policy

- An on-policy agent learns only about the policy that it is executing
- An off-policy agent learns about a policy or policies different from the one that it is executing

RAMAPURAM SRM

Reward Signal

- The reward signal defines the goal
- On each time step, the environment sends a single number called the reward to the reinforcement learning agent
- The agent's objective is to maximise the total reward that it receives over the long run
- The reward signal is used to alter the policy

RAMAPURAM SRM

Value Function (1)

Value Function (1)

- The reward signal indicates what is good in the short run while the value function indicates what is good in the long run
- The value of a state is the total amount of reward an agent can expect to accumulate over the future, starting in that state
- Compute the value using the states that are likely to follow the current state and the rewards available in those states
- Future rewards may be time-discounted with a factor in the interval [0, 1]

Value Function (2)

- Use the values to make and evaluate decisions
- Action choices are made based on value judgements
- Prefer actions that bring about states of highest value instead of highest reward
- Rewards are given directly by the environment
- Values must continually be re-estimated from the sequence of observations that an agent makes over its lifetime

Model-free versus Model-based

- A model of the environment allows inferences to be made about how the environment will behave

RAMAPURAM SRM

Model-free versus Model-based

- A model of the environment allows inferences to be made about how the environment will behave
- Example: Given a state and an action to be taken while in that state, the model could predict the next state and the next reward
- Models are used for planning, which means deciding on a course of action by considering possible future situations before they are experienced
- Model-based methods use models and planning. Think of this as modelling the dynamics p(s' | s, a)
- Model-free methods learn exclusively from trial-and-error (i.e. no modelling of the environment)
- This presentation focuses on model-free methods

RAMAPURAM SRM

Advantages of Reinforcement Learning

- It can solve higher-order and complex problems. Also, the solutions obtained will be very accurate.
- The reason for its perfection is that it is very similar to the human learning technique.
- Due to it's learning ability, it can be used with neural networks. This can be termed as **deep reinforcement learning**.
- Since the model learns constantly, a mistake made earlier would be unlikely to occur in the future.
- Various problem-solving models are possible to build using reinforcement learning.
- When it comes to creating simulators, object detection in automatic cars, robots, etc., reinforcement learning plays a great role in the models.
- The best part is that even when there is no training data, it will learn through the experience it has from processing the training data.
- For various problems, which might seem complex to us, it provides the perfect models to tackle them.

RAMAPURAM SRM

Disadvantages of Reinf

RAMAPURAM SRM

Need for Multi agent Learning

- A single agent cannot handle learning in case of applications
- A team or group of agents possesses the potential to overcome the limitation of single agent and work in in coordination to accomplish a task.
- This can be two cases in multi agent based learning:

RAMAPURAM SRM

Multi agent based learning

1) Where the agent tries to maximize its own utility

- Eg: Consider a manufacturing industry domain. The task are built and assigned where each agent works in co-operation to build the end product. This is the case to achieve a common goal.

2) Where they worked in collaboration to achieve some common goals.

- Eg: Game playing. In that gaming environment, multiple agents are in Operation to select the best strategy.
- Can be related with the reinforcement learning, Where for each strategy of the agent some reward is achieved this is where each agent tries to maximize his own utility function.

RAMAPURAM SRM

Speedup learning

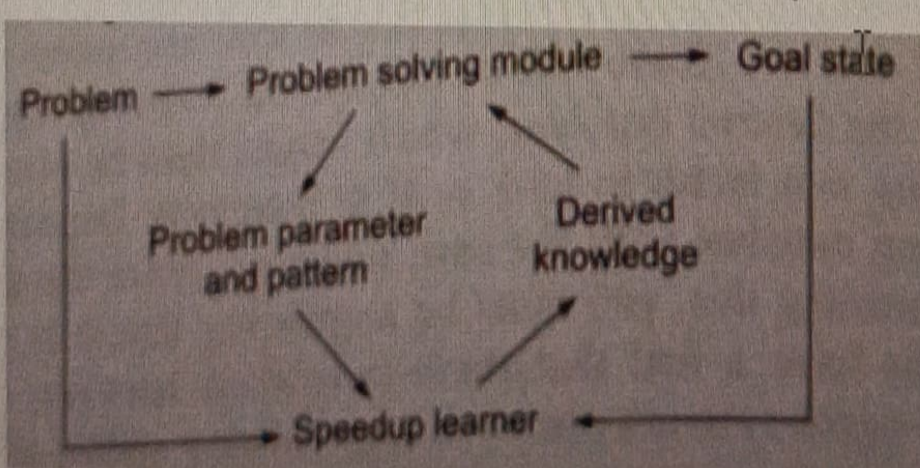
- Speed learning typically deals with speeding up problem solving by effective use of problem solving experience. Hence, Prayer problem solving experience is an input for speed of learning.
- In this learning,

1) There is no option with the environment.

2) New problems cannot be solved.

So, speed up learning accelerates the process experiences And prior observations.

Speedup learning (Contd..)



Speedup learning modules.

Hierarchical planning is also called as plan decomposition. Generally plans are organized in Hierarchical format.

Pop one level planner:

If you are planning to take a trip, then first you have to decide the location. To decide location we can search for various good locations from internet based on, whether conditions, travelling expenses, etc. This is level one planning.

Hierarchy of actions:

In terms of major and minor or actions, hierarchy of actions can be decided. Minor activities would cover more precise activities to accomplish the major activities.

Major steps are given more importance. Once major steps are decided we attempt to solve the minor detailed actions.

Planner:

i. First identify a hierarchy of major conditions.

ii. Construct a plan in levels.

iii. Patch major level details.

iv. Finally demonstrate

	Search	**Planning**
States	data structures	Logical sentences
Actions	code	Preconditions/outcomes
Goal	code	Logical sentence (conjunction)
Plan	Sequence from S_0	Constraints on actions

Answering to your question, planning only do predictions and just tells us the outcomes but searching need coding that's why it has 100 actions.

RAMAPURAM SRM

Adaptive learning

- No learning method is complete in itself. So
- Need to select the learning method based on the requirements.
- Need to develop a combination of some of the existing methods based on requirements.
- Adaptive machine learning algorithms are the machine learning models, where the changes in the environment help in selecting the algorithm or learning method.

RAMAPURAM SRM

Adaptive learning (Contd..)

- As per the scenario, most suitable algorithm is selected.
- Moreover the development of especially fast adapting algorithms poses many different issues like selection of choices, handling equilibrium states and so on.
- The adaptive learning solves some of the complex problems for which a single learning method is not enough.
- This method is even more appropriate when the environment is continuously changing and real time response is expected.

Ensemble Learning (Contd...)

Ensemble learning method is the one where multiple learners or learning algorithms are trained.

- In most of the learning algorithms a single hypothesis drives the learning.
- In ensemble learning method the whole collection or ensemble of hypothesis is selected from the hypothesis space and their predictions are combined.
- In this approach, the learners or referred to as base learners.
- The most commonly used ensemble learning methods are

1) Boosting
2) Bagging.

Ensemble Learning (Contd..)

Boosting:

- Boosting can probably be defined as the method for generating accurate predictions by combining the rules that are comparatively inaccurate.

Ensemble Learning (Contd..)

Boosting:

- Boosting can probably be defined as the method for generating accurate predictions by combining the rules that are comparatively inaccurate.
- Boosting works on the weighted training sets. The weights of the training example reflects the importance of training examples.

Bagging:

- In Bagging, the training data is resampled. This is referred to as ***bootstrap sampling***, where the training data with replacement is taken in the learning approaches.