MACHINE LEARNING
UNIT-1

By, Mrutyunjaya S Yalawar

A.Y.2020-2021

Introduction

- Machine Learning algorithms enable the computers to learn from data, and even improve themselves, without being explicitly programmed.
- Any field that needs to interpret and act on data can benefit from machine learning techniques.
- The goal of this course is to present key algorithms and theory that form the core of machine learning with a balanced presentation of both theory and practice.

Prerequisites

- Data Structures
- Knowledge on statistical methods

Course Objectives

- This course explains machine learning techniques such as decision tree learning, Bayesian learning etc.
- To understand computational learning theory.
- To study the pattern comparison techniques.

Course Outcomes

- Understand the concepts of computational intelligence like machine learning
- Ability to get the skill to apply machine learning techniques to address the real time problems in different areas
- Understand the Neural Networks and its usage in machine learning application

JNTUH Syllabus
R18 Regulations

By, Mrutyunjaya S Yalawar, Dept of CSE

UNIT-I

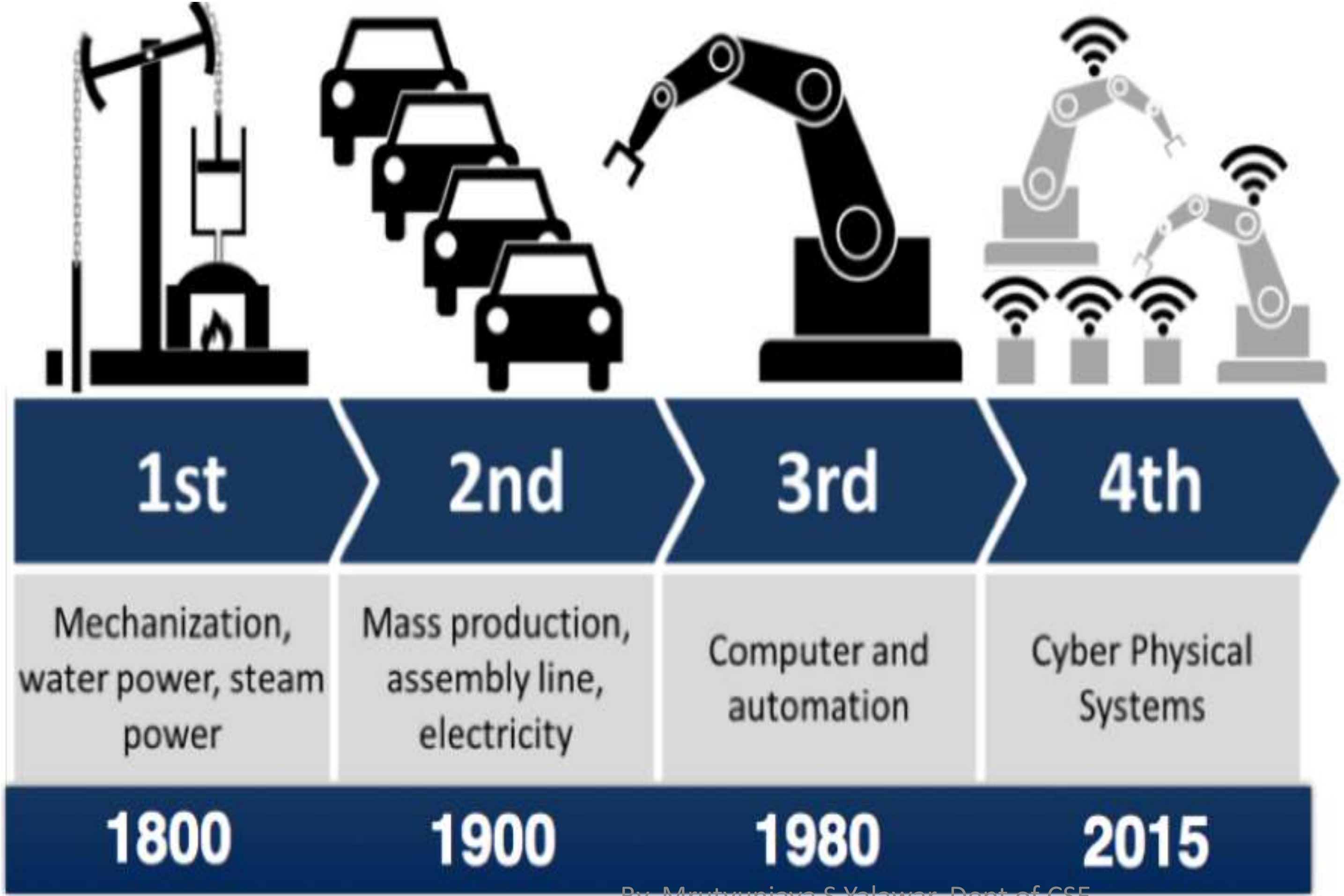
10 Hours

- **Introduction -** Well-posed learning problems, designing a learning system, Perspectives and issues in machine learning
- **Concept learning and the general to specific ordering –** introduction, a concept learning task, concept learning as search, find-S: finding a maximally specific hypothesis, version spaces and the candidate elimination algorithm, remarks on version spaces and candidate elimination, inductive bias.
- **Decision Tree Learning –** Introduction, decision tree representation, appropriate problems for decision tree learning, the basic decision tree learning algorithm, hypothesis space search in decision tree learning, inductive bias in decision tree learning, issues in decision tree learning.

Machine Learning

- **Learning ↔ Intelligence**

  (Def: Intelligence is the ability to learn and use concepts to solve problems.)
- **Machine Learning ↔ Artificial Intelligence**
- **Def:** **AI** is the science of making machines do things that require intelligence if done by men (Minsky 1986)
- **Def:** **Machine Learning** is an area of AI concerned with development of techniques which allow machines to learn.
- **Why Machine Learning? ↔ Why Artificial Intelligence?**


1st
2nd
3rd
4th
Mechanization, water power, steam power
Mass production, assembly line, electricity
Computer and automation
Cyber Physical Systems
1800
1900
1980
2015


NEURAL COMMUNICATION
2,53
miles
130
bpm
WORLD ECONOMIC FORUM
We'll control gadgets with brain signals
The world will be one in which we can communicate our intent directly and instantly to machines and have very complex outcomes

- Learning ↔ Intelligence

  (Def: Intelligence is the ability to learn and use concepts to solve problems.)

- Machine Learning ↔ Artificial Intelligence

  Def: AI is the science of making machines do things that require intelligence if done by men (Minsky 1986)

  Def: Machine Learning is an area of AI concerned with development of techniques which allow machines to learn

- **Why Machine Learning? ↔ Why Artificial Intelligence?**

  To build machines exhibiting intelligent behaviour (i.e., able to reason, predict, and adapt) while helping humans work, study, and entertain themselves.

- Machine Learning ↔ Artificial Intelligence
- Machine Learning ← **Biology** (e.g., Neural Networks, Genetic Algorithms)
- Machine Learning ← **Cognitive Sciences** (e.g., Case-based Reasoning)
- Machine Learning ← **Statistics** (e.g., Support Vector Machines)
- Machine Learning ← **Probability** Theory (e.g., Bayesian Networks)
- Machine Learning ← **Logic** (e.g., Inductive Logic Programming)
- Machine Learning ← **Information Theory** (e.g., used by Decision Trees)

Applications

- The highly complex nature of many real-world problems, though, often means that inventing specialized algorithms that will solve them perfectly every time is impractical, if not impossible.

- Examples of machine learning problems include, "Is this cancer?", "What is the market value of this house?", "Which of these people are good friends with each other?", "Will this rocket engine explode on take off?", "Will this person like this movie?", "Who is this?", "What did you say?", and "How do you fly this thing?".

- All of these problems are excellent targets for an ML project, and in fact ML has been applied to each of them with great success.

Well-posed learning problems

By, Mrutyunjaya S Yalawar, Dept of CSE

Well-posed learning problems.

- Def 1 (Mitchell 1997): A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves by experience E.

- Def 2 (Hadamard 1902): A (machine learning) problem is well-posed if a solution to it exists, if that solution is unique, and if that solution depends on the data / experience but it is not sensitive to (reasonably small) changes in the data / experience.

Continuation…….

- **<u>A checkers learning problem</u>**

  Task T : playing checkers.

  Performance measure P : percent of games won against opponents.

  Training experience E : playing practice games against itself.

- **<u>A handwriting recognition learning problem</u>**

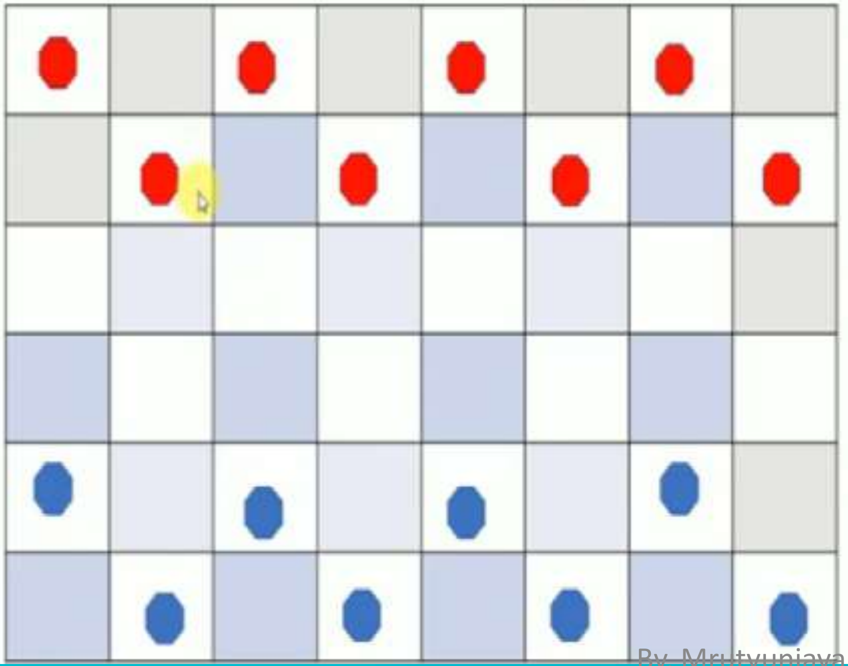
  Task T : recognizing and classifying handwritten words within images.

  Performance measure P : percent of words correctly classified

  Training experience E : a database of handwritten words with given classifications.

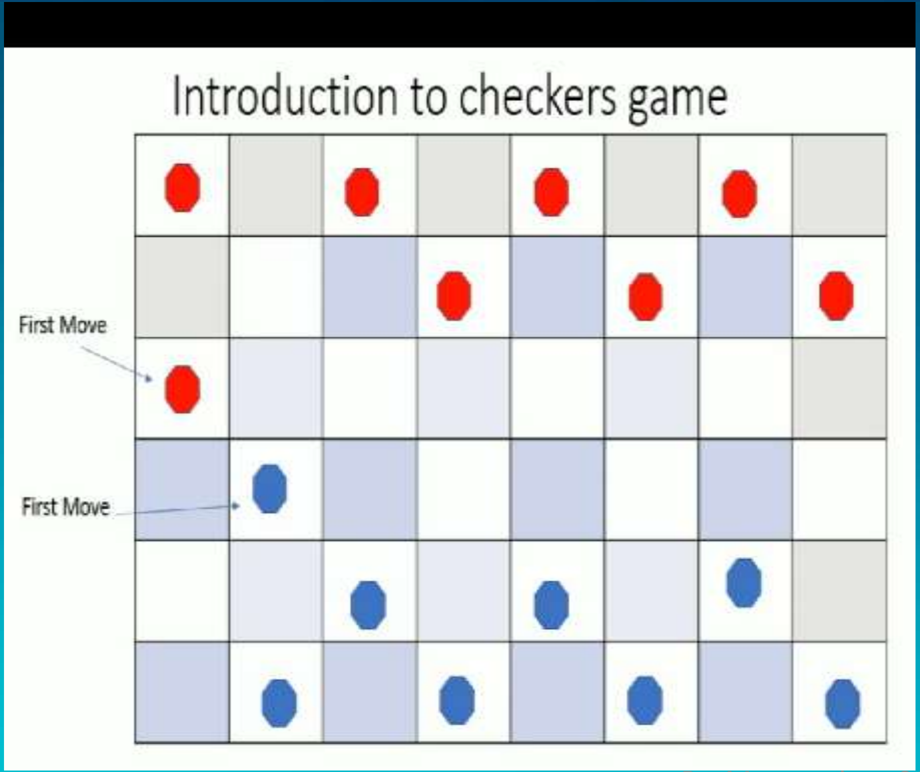
Checker Game

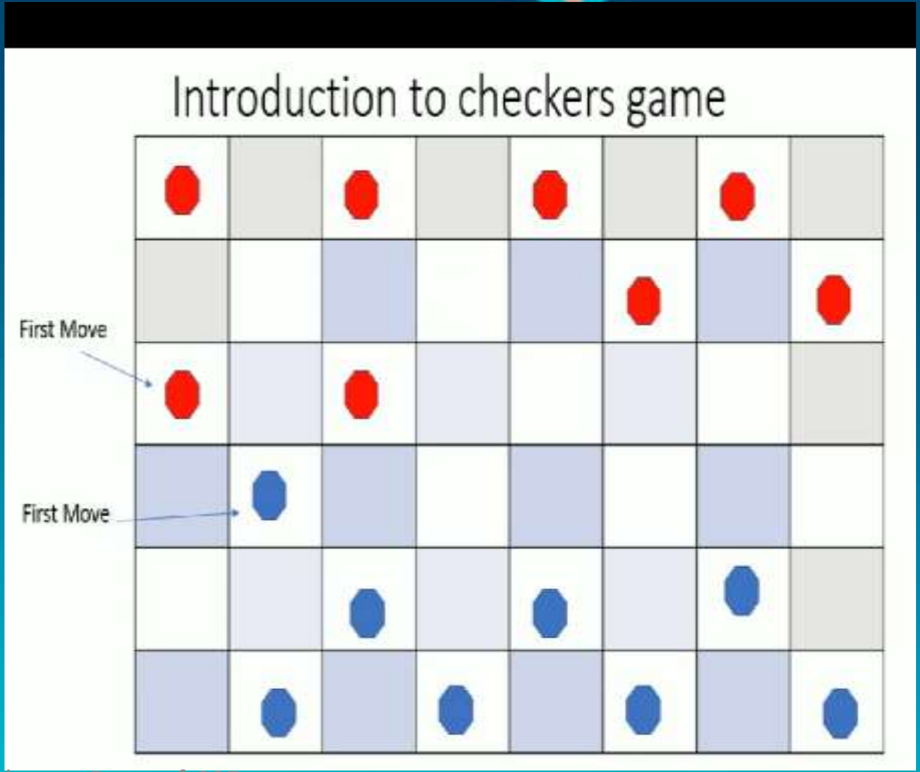
Introduction to checkers game

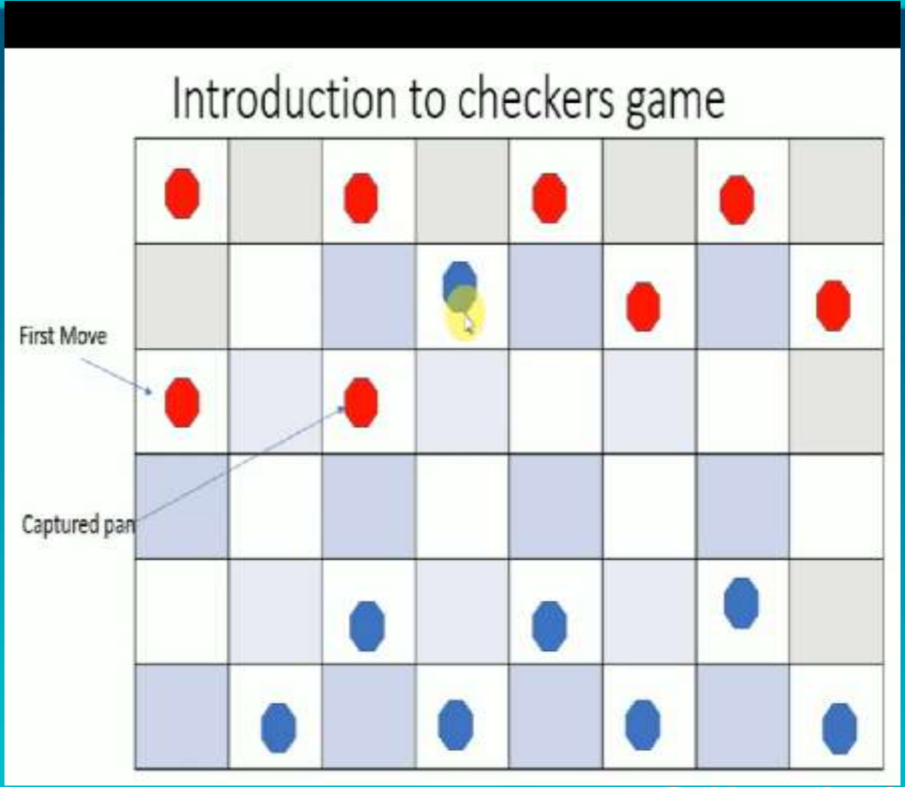


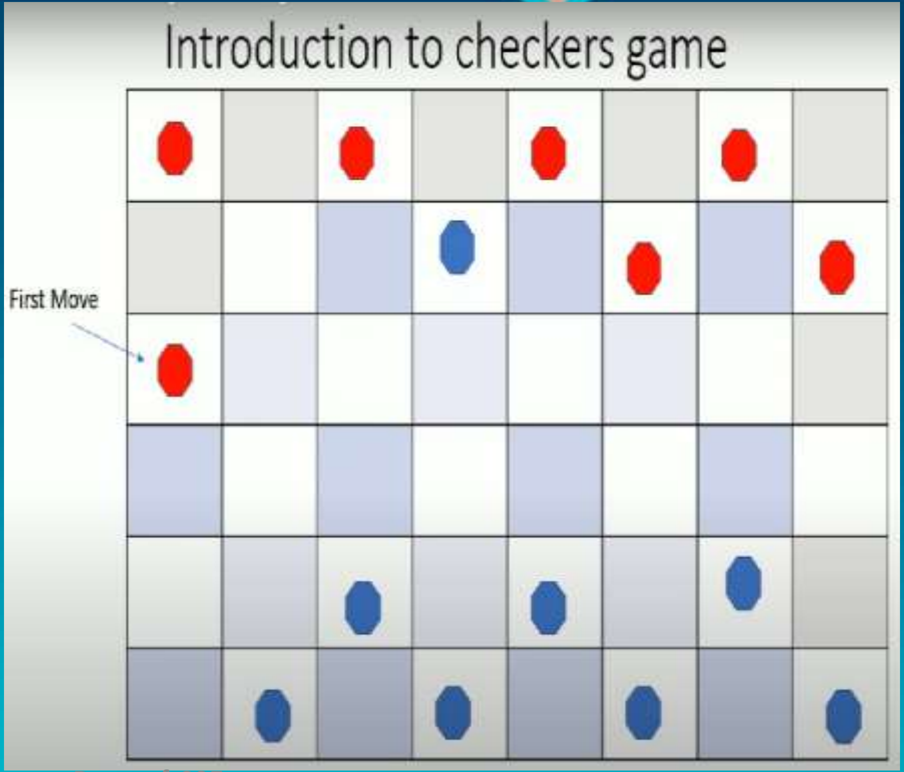
Rules of the checkers game

- Always move the checker diagonally forward toward the opponents side
- Note as soon as the checker moves to the first row of the opponents side the checker becomes the king.
- The king can move both the directions diagonally forward or backward.
- To start ,move the checker diagonally one place ahead.
- To capture the opponent jump over an opponents checker to capture it.
- When all the opponents checker is captured or not able to move then the player has won.


Introduction to checkers game
First Move
First Move


Introduction to checkers game
First Move
First Move

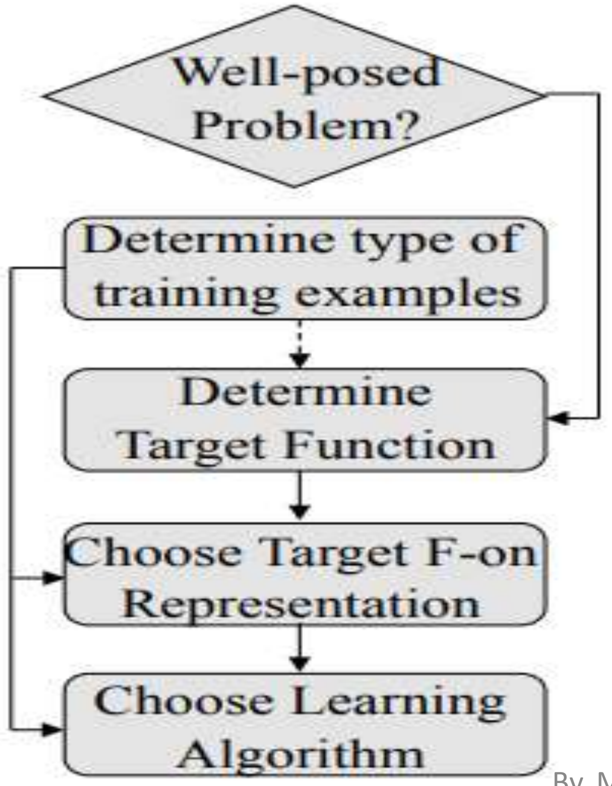

Introduction to checkers game
First Move
Captured pan


Introduction to checkers game
First Move

Designing a learning system.

By, Mrutyunjaya S Yalawar, Dept of CSE

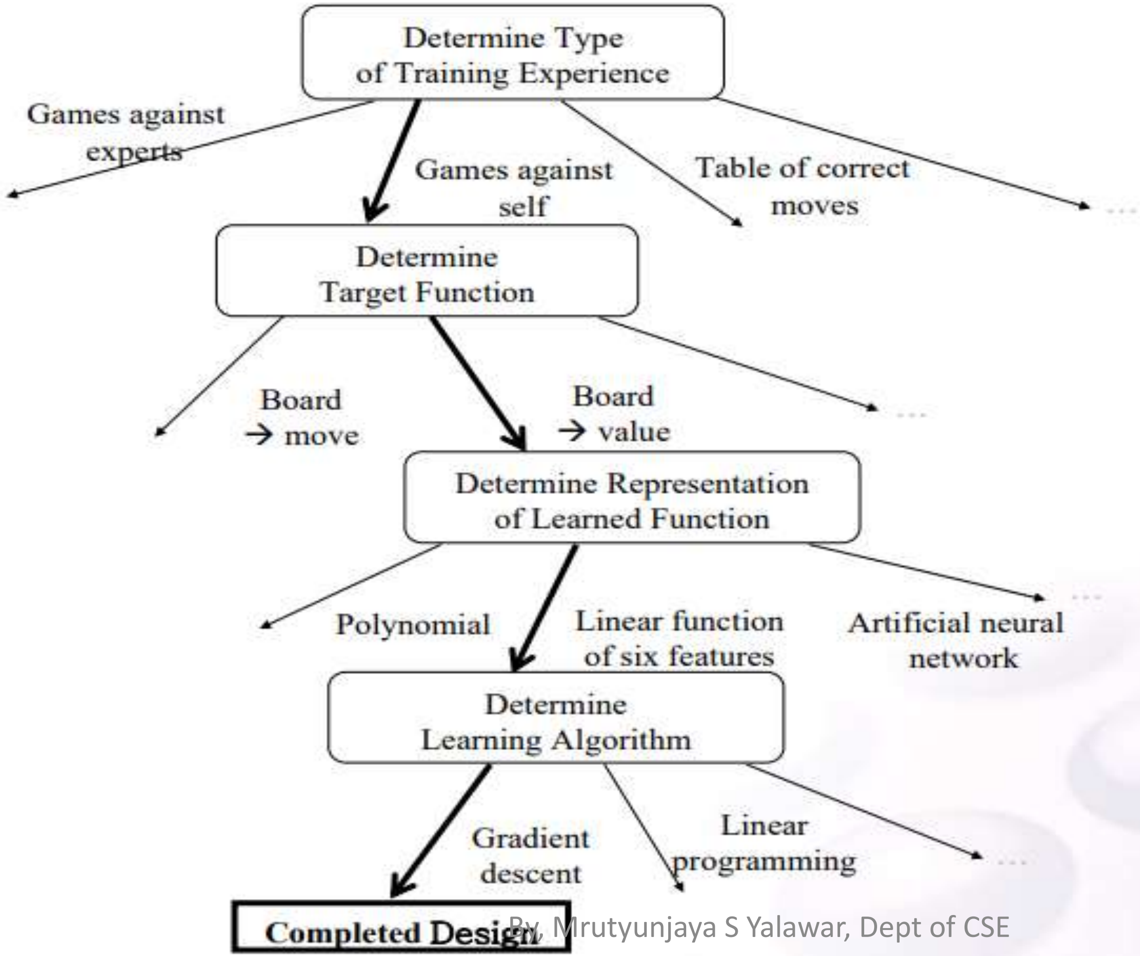
Designing a learning system.



- **In other ways like……..**

1) Choosing the Training Experience.
2) Choosing the Target Function.
3) Choosing a Representation for the Target Function.
4) Choosing a Function Approximation Algorithm.
5) The Final Design.

Designing a learning system (Conti….)



Perspectives and Issues in Machine Learning

- What algorithms exist for learning general target functions from specific training examples ?

- How does the number of training examples influence accuracy ?

- When and how can prior knowledge held by the learner guide the process of generalizing from examples ?

Issues in Machine Learning (cont.)

- What is the best strategy for choosing a useful next training experience, and how does the choice of this strategy alter the complexity of the learning problem ?

- What is the best way to reduce the learning task to one or more function approximation problems ?

- How can the learner automatically alter its representation to improve its ability to represent and learn the target function ?

Concept Learning

Concept learning

Concept learning

- supervised, eager learning
- target problem: whether something belongs to the target concept or not.
- target function: V: D → {true, false}

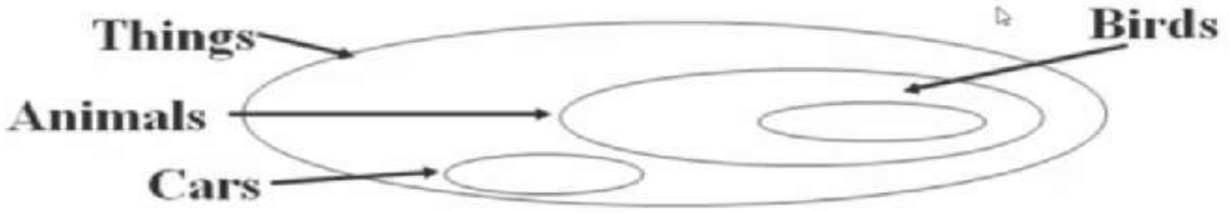
- Underlying idea: Humans acquire general concepts from specific examples (e.g., concepts: beauty, good friend, well-fitting-shoes) (note: each concept can be thought of as Boolean-valued function)

- Concept learning is inferring a Boolean-valued function from training data

→ concept learning is the prototype binary classification.

WHAT IS A CONCEPT?

A Concept is a subset of objects or events defined over a larger set. For example, We refer to the set of everything (i.e. all objects) as the set of things. Animals are a subset of things, and birds are a subset of animals.



WHAT IS A CONCEPT?

In more technical terms, a concept is a boolean-valued function defined over this larger set.

For example, a function defined over all *animals* whose value is *true* for birds and *false* for every other animal.

WHAT IS A CONCEPT LEARNING?

Given a set of examples labeled as members or non-members of a concept, *concept-learning* consists of automatically inferring the general definition of this concept.

In other words, *concept-learning* consists of approximating a boolean-valued function from training examples of its input and output.

Example of a Concept Learning Task

⇒ **Concept:** Good Days for Watersports (values: Yes, No)

⇒ **Attributes/Features:**

- Sky (values: Sunny, Cloudy, Rainy)
- AirTemp (values: Warm, Cold)
- Humidity (values: Normal, High)
- Wind (values: Strong, Weak)
- Water (Warm, Cool)
- Forecast (values: Same, Change)

⇒ **Example of a Training Point:**

<Sunny, Warm, High, Strong, Warm, Same, Yes>

Example of a Concept Learning Task

Day	Sky	Airtemp	Humidity	Wind	Water	Forecast	WaterSport
1.	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2.	Sunny	Warm	High	Strong	Warm	Same	Yes
3.	Rainy	Cold	High	Strong	Warm	Change	No
4.	Sunny	Warm	High	Strong	Cool	Change	Yes

Chosen Hypothesis Representation:

Conjunction of constraints on each attribute where:

- "?" means "any value is acceptable"
- "0" means "no value is acceptable"

Example of a hypothesis: <?,Cold,High,?,?,?>

(If the air temperature is cold and the humidity high then it is a good day for water sports)

Example of a concept learning task

Goal: To infer the "best" concept-description from the set of all possible hypotheses ("best" means "which best generalizes to all (known or unknown) elements of the instance space").

⇒ **Most General Hypothesis:** Everyday is a good day for water sports

<?,?,?,?,?,?>

⇒ **Most Specific Hypothesis:** No day is a good day for water sports <∅, ∅,∅,∅,∅,∅>

Concept Learning as search

⇒ Concept Learning can be viewed as the task of searching through a large space of hypotheses implicitly defined by the hypothesis representation.

⇒ Selecting a Hypothesis Representation is an important step since it restricts (or *biases*) the space that can be searched. [For example, the hypothesis "If the air temperature is cold **or** the humidity high then it is a good day for water sports" cannot be expressed in our chosen representation.]

Concept Learning Task – Notation

- Concept learning task:
  - target concept: Girls who Simon likes
  - target function: $c: D \rightarrow \{0, 1\}$
  - data $d \in D$: Girls, each described in terms of the following attributes
    - $a_1 \equiv$ *Hair* (possible values: blond, brown, black)
    - $a_2 \equiv$ *Body* (possible values: thin, average, plump)

instances

    - $a_3 \equiv$ *likesSimon* (possible values: yes, no)
    - $a_4 \equiv$ *Pose* (possible values: arrogant, natural, goofy)
    - $a_5 \equiv$ *Smile* (possible values: none, pleasant, toothy)
    - $a_6 \equiv$ *Smart* (possible values: yes, no)

error rate

  - target f-on representation: $h \equiv c'$: $\langle a_1, a_2, a_3, a_4, a_5, a_6 \rangle \rightarrow \{0, 1\}$
  - training examples D: positive and negative examples of target function c

- **Aim**: Find a hypothesis $h \in H$ such that $(\forall d \in D)\ h(d) - c(d) < \varepsilon \approx 0$, where H is the set of all possible hypotheses $h \equiv \langle a_1, a_2, a_3, a_4, a_5, a_6 \rangle$, where each a_k, $k = [1..6]$, may be '?' (≡ any value is acceptable), '0' (≡ no value is acceptable), <u>or a specific value</u>.

h ≡ ⟨?, ?, ?, ?, ?, ?⟩ h ≡ ⟨0, 0, 0, 0, 0, 0⟩ h ≡ ⟨?, ?, yes, ?, ?, ?⟩

Concept Learning as Search

- Concept learning task:
  - target concept: Girls who Simon likes
  - target function: c: $D \rightarrow \{0, 1\}$
  - data $d \in D$: Girls, each described in terms of the following attributes
    - $a_1 \equiv$ *Hair* (possible values: blond, brown, black) + '?'
    - $a_2 \equiv$ *Body* (possible values: thin, average, plump) + '?'
    - $a_3 \equiv$ *likesSimon* (possible values: yes, no) + '?' $\rightarrow |H| = 1 + 4 \cdot 4 \cdot 3 \cdot 4 \cdot 4 \cdot 3 = 2305$
    - $a_4 \equiv$ *Pose* (possible values: arrogant, natural, goofy) + '?'
    - $a_5 \equiv$ *Smile* (possible values: none, pleasant, toothy) + '?'
    - $a_6 \equiv$ *Smart* (possible values: yes, no) + '?'

    instances

    $h \equiv \langle 0,0,0,0,0,0 \rangle$

    error rate
  - target f-on representation: $h \equiv c'$: $\langle a_1, a_2, a_3, a_4, a_5, a_6 \rangle \rightarrow \{0, 1\}$
  - training examples D: positive and negative examples of target function c

- **Aim**: Find a hypothesis $h \in H$ such that $(\forall d \in D)\ h(d) - c(d) < \varepsilon \approx 0$, where H is the set of all possible hypotheses $h \equiv \langle a_1, a_2, a_3, a_4, a_5, a_6 \rangle$, where each a_k, $k = [1..6]$, may be '?' ($\equiv$ any value is acceptable), '0' ($\equiv$ no value is acceptable), or a specific value.

concept learning $\equiv$ searching through H

Find-S Algorithm – Example

1. Initialise $h \in H$ to the most specific hypothesis: $h \leftarrow \langle a_1,\ldots,a_n \rangle$, $(\forall i)\ a_i = 0$.
2. FOR each positive training instance $d \in D$, do:

    FOR each attribute a_i, $i = [1..n]$, in h, do:

        IF a_i is satisfied by d

        THEN do nothing

        ELSE replace a_i in h so that the resulting $h' >_g h$, $h \leftarrow h'$.

3. Output hypothesis h.

	c(d)	*hair*	*body*	*likesSimon*	*pose*	*smile*	*smart*
1	1	blond	thin	yes	arrogant	toothy	no
2	0	brown	thin	no	natural	pleasant	yes
3	1	blond	plump	yes	goofy	pleasant	no
4	0	black	thin	no	arrogant	none	no
5	0	blond	plump	no	natural	toothy	yes

$h \leftarrow \langle 0,0,0,0,0,0 \rangle$ → $h \equiv d1$ → $h \leftarrow \langle$blond, ?, yes, ?, ?, no$\rangle$

Thank You