Theoretical Probability

Lesson 1.1

LISTING OUTCOMES

Example 1

List the sample space of possible outcomes for:

a tossing a coin

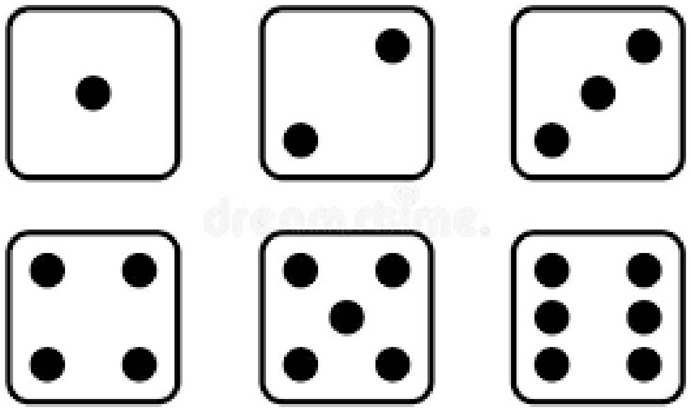
b rolling a die.

a When a coin is tossed, there are two possible outcomes.

$\therefore$ sample space $= \{H, T\}$

b When a die is rolled, there are 6 possible outcomes.

$\therefore$ sample space $= \{1, 2, 3, 4, 5, 6\}$

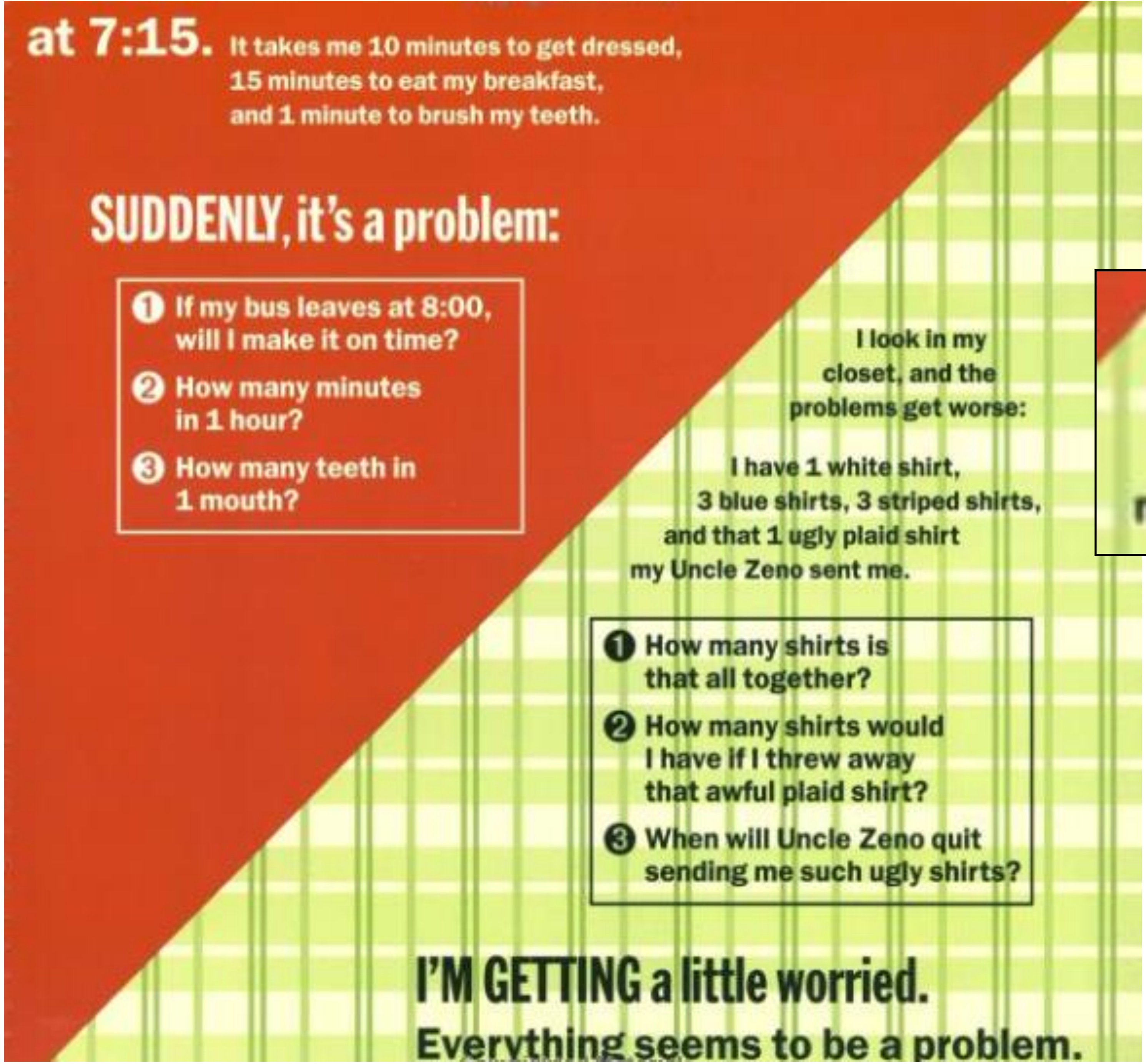


A

B

C

D


at 7:15. It takes me 10 minutes to get dressed, 15 minutes to eat my breakfast, and 1 minute to brush my teeth.

SUDDENLY, it's a problem:

1. If my bus leaves at 8:00, will I make it on time?
2. How many minutes in 1 hour?
3. How many teeth in 1 mouth?

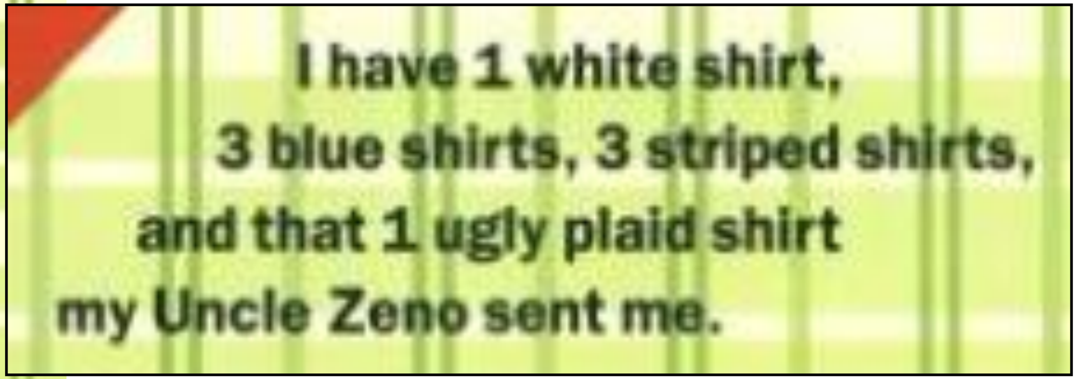
I look in my closet, and the problems get worse:

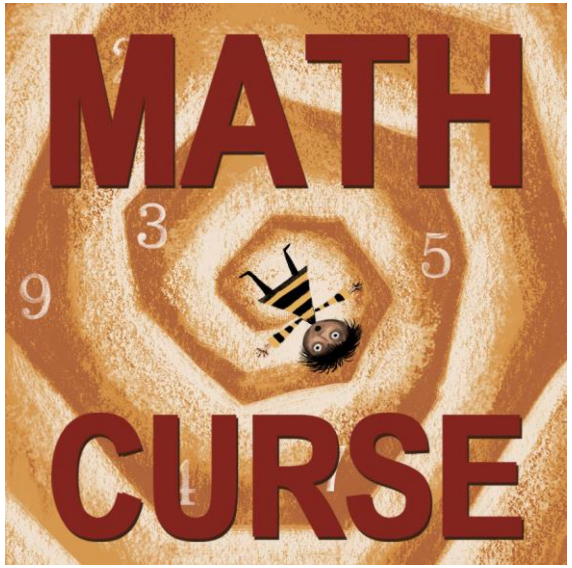
I have 1 white shirt, 3 blue shirts, 3 striped shirts, and that 1 ugly plaid shirt my Uncle Zeno sent me.

1. How many shirts is that all together?
2. How many shirts would I have if I threw away that awful plaid shirt?
3. When will Uncle Zeno quit sending me such ugly shirts?

I'M GETTING a little worried.

Everything seems to be a problem.





We can indicate the likelihood of an event happening in the future by using a percentage.

0%	indicates we believe the event **will not occur**.
100%	indicates we believe the event **is certain to occur**.

All events can therefore be assigned a percentage between 0% and 100% (inclusive).

A number close to 0% indicates the event is **unlikely** to occur, whereas a number close to 100% means that it is **highly likely** to occur.

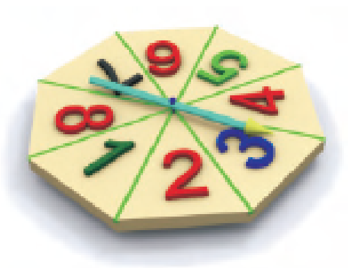
In mathematics, we usually use either decimals or fractions rather than percentages for probabilities. However, as $100\% = 1$, comparisons or conversions from percentages to fractions or decimals are very simple.

An **impossible** event which has 0% chance of happening is assigned a probability of 0.

A **certain** event which has 100% chance of happening is assigned a probability of 1.

All other events can then be assigned a probability between 0 and 1.

C THEORETICAL PROBABILITY



Consider the **octagonal spinner** alongside.

Since the spinner is symmetrical, when it is spun the arrowed marker could finish with equal likelihood on each of the sections marked 1 to 8.

The likelihood of obtaining a particular number, for example 4, would be:

1 chance in 8, $\frac{1}{8}$, $12\frac{1}{2}\%$ or 0.125.

This is a **mathematical** or **theoretical** probability and is based on what we theoretically expect to occur. It is a measure of the chance of that event occurring in any trial of the experiment.

If we are interested in the event of getting a result of *6 or more* from one spin of the octagonal spinner, there are three favourable results (6, 7 or 8) out of the eight possible results. Since each of these is equally likely to occur, $\text{P}(6 \text{ or more}) = \frac{3}{8}$.

In general, for an event E containing **equally likely** possible results, the probability of E occurring is

$$\text{P}(E) = \frac{\textbf{the number of members of the event } E}{\textbf{the total number of possible outcomes}} = \frac{n(E)}{n(U)}.$$

Theoretical probability

$$P(E) = \frac{\text{the number of members of the event } E}{\text{the total number of possible outcomes}} = \frac{n(E)}{n(U)}.$$

Example 4

A ticket is *randomly selected* from a basket containing 3 green, 4 yellow and 5 blue tickets. Determine the probability of getting:

a a green ticket

b a green or yellow ticket

c an orange ticket

d a green, yellow or blue ticket

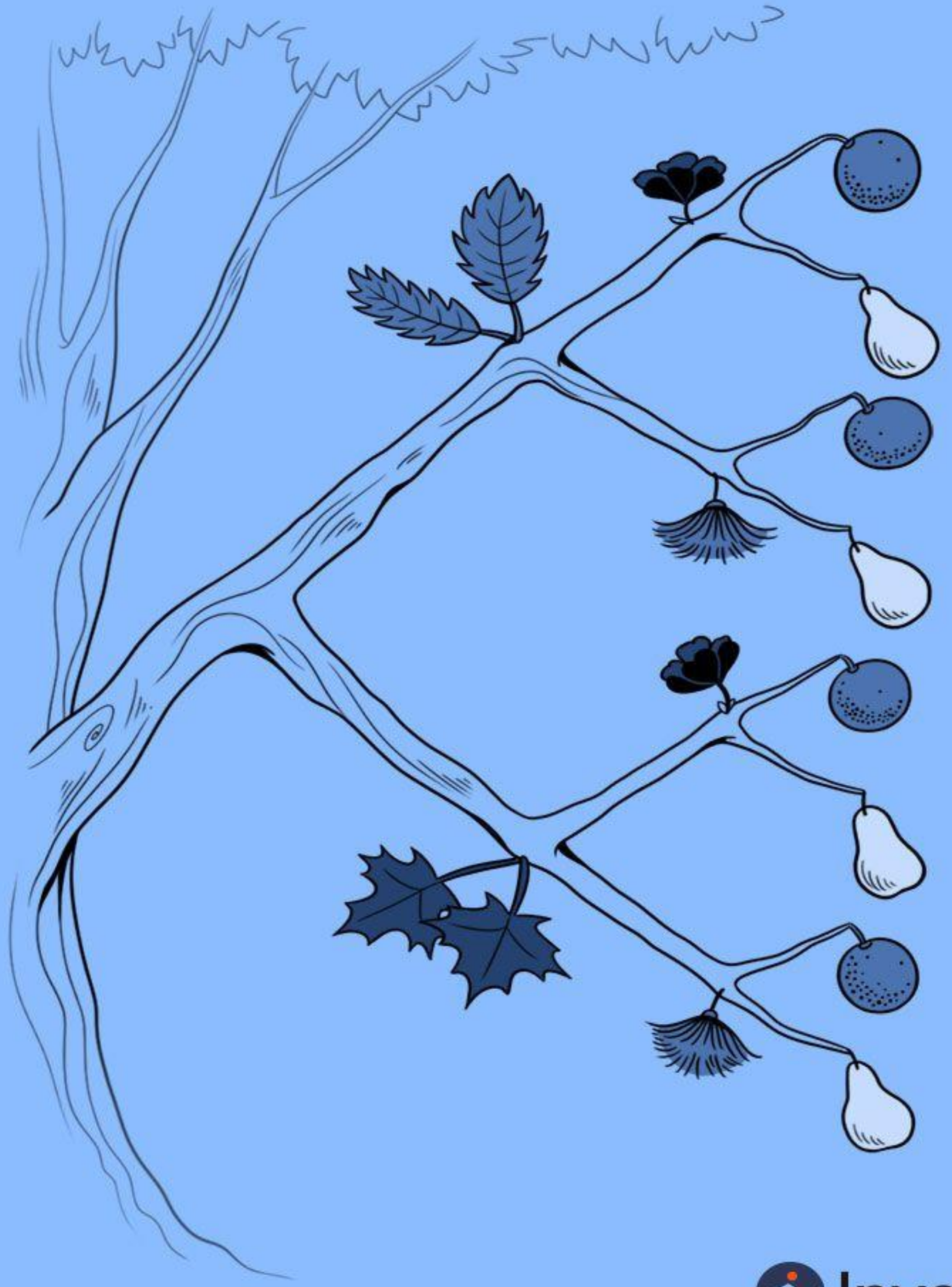
The sample space is {G, G, G, Y, Y, Y, Y, B, B, B, B, B}
which has $3 + 4 + 5 = 12$ outcomes.

a P(G)
$= \frac{3}{12}$
$= \frac{1}{4}$

b P(a G or a Y)
$= \frac{3+4}{12}$
$= \frac{7}{12}$

c P(O)
$= \frac{0}{12}$
$= 0$

d P(G, Y or B)
$= \frac{3+4+5}{12}$
$= 1$



Tree Diagram

['trē 'dī-ə-ˌgram]

A tool used in mathematics, probability, and statistics to calculate possible outcomes of an event or problem, and to cite those potential outcomes in an organized way.

Investopedia

TREE DIAGRAMS

The sample space in **Example 2** could also be represented by a tree diagram. The advantage of tree diagrams is that they can be used when more than two operations are involved.

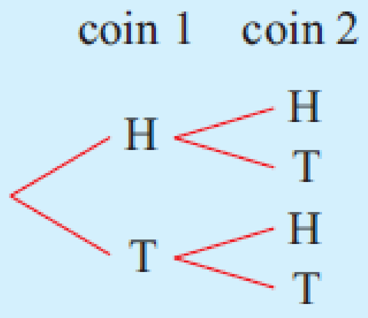
Example 3

Illustrate, using a tree diagram, the possible outcomes when:

a tossing two coins

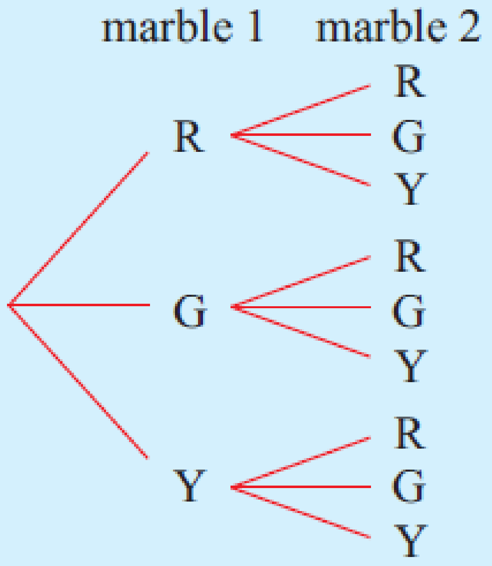
b drawing two marbles from a bag containing many red, green, and yellow marbles.

a



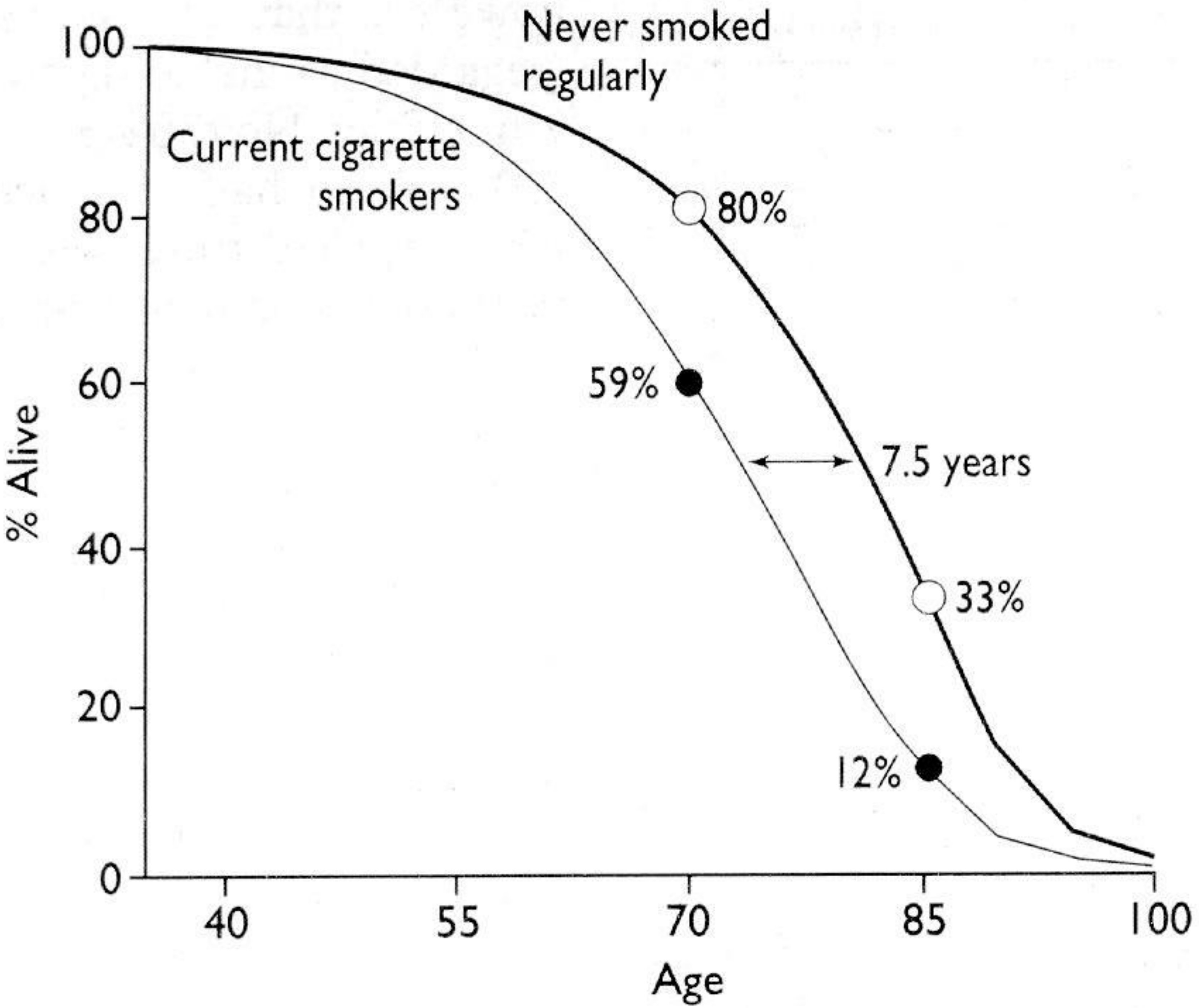
Each "branch" gives a different outcome and the sample space is seen to be {HH, HT, TH, TT}.

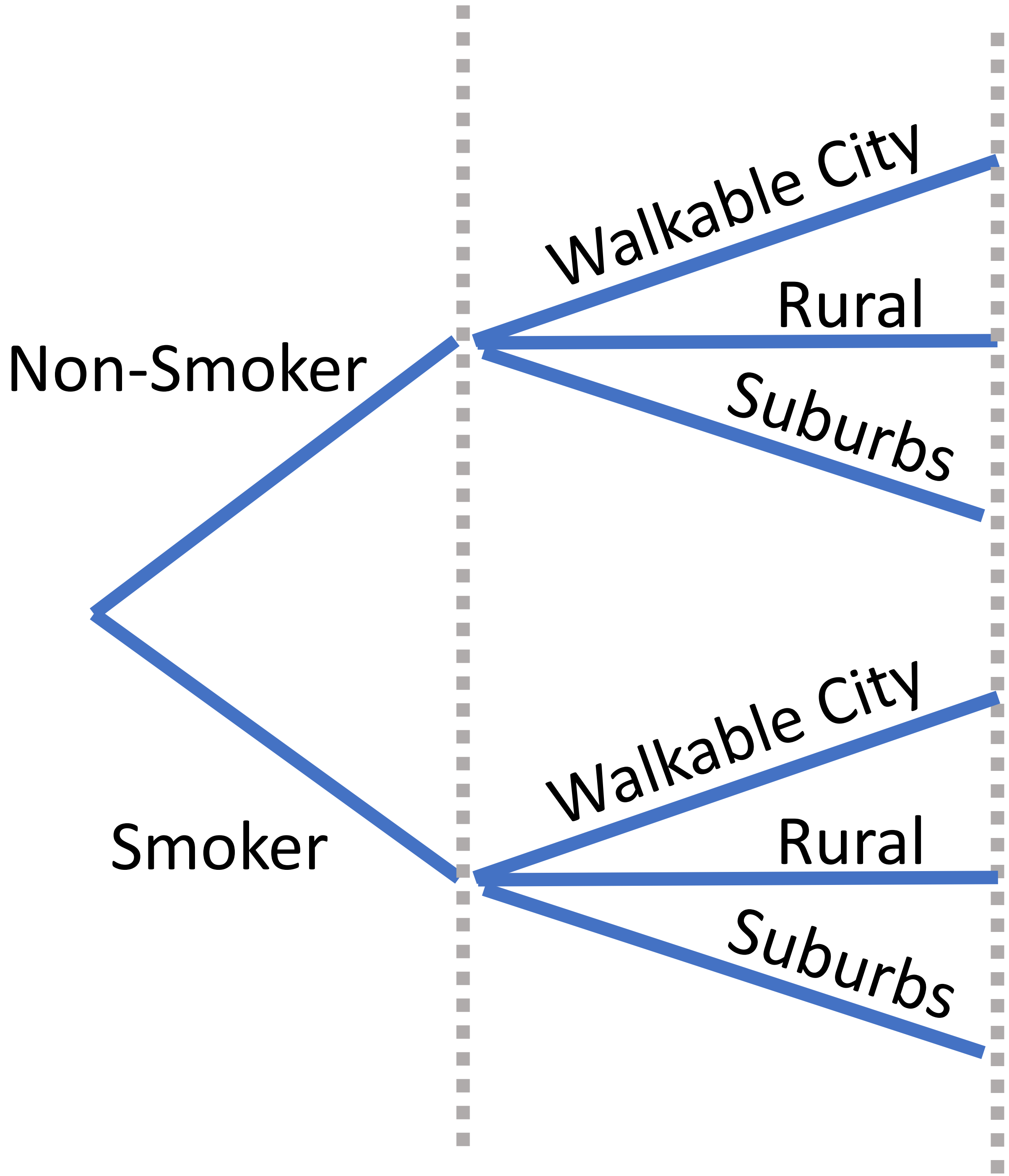
b



What is the sample space here?

Life Insurance Companies





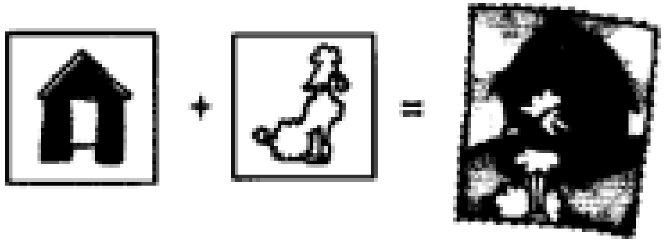
Life Insurance Includes at least 21 Other Branches:

- Postal Code
- Gender

E

How many outfits can You make with 2 shirts, 3 pants and 2 shoes?

Game Play

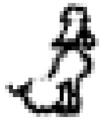


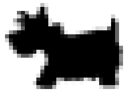
or

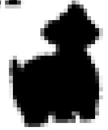
The youngest player goes first. He rolls the two dice together, then calls out the roll, saying, for example, "Stretch and the kid." All players look at their bingo boards to see if they have the square that matches (in this case, the square would show the kid throwing a ball to Stretch). Each player who has a match marks that square with a bone.

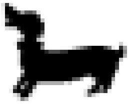
F

Dog Die

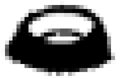
Gigi 

McDuff 

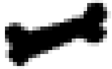
Gizmo 

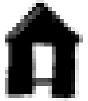
Stretch 

Action Die

dish 

kid 

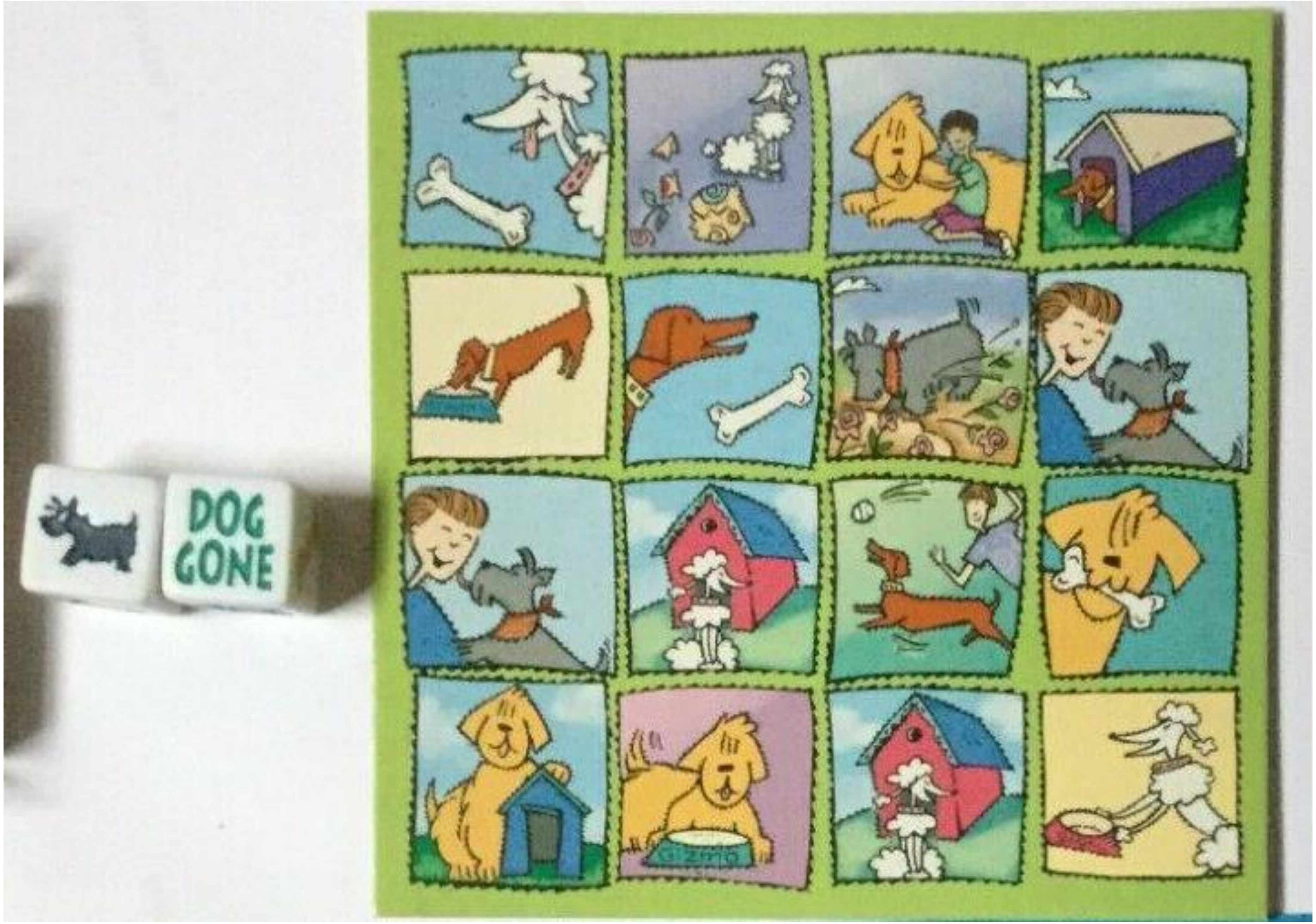
bone 

house 

trouble 

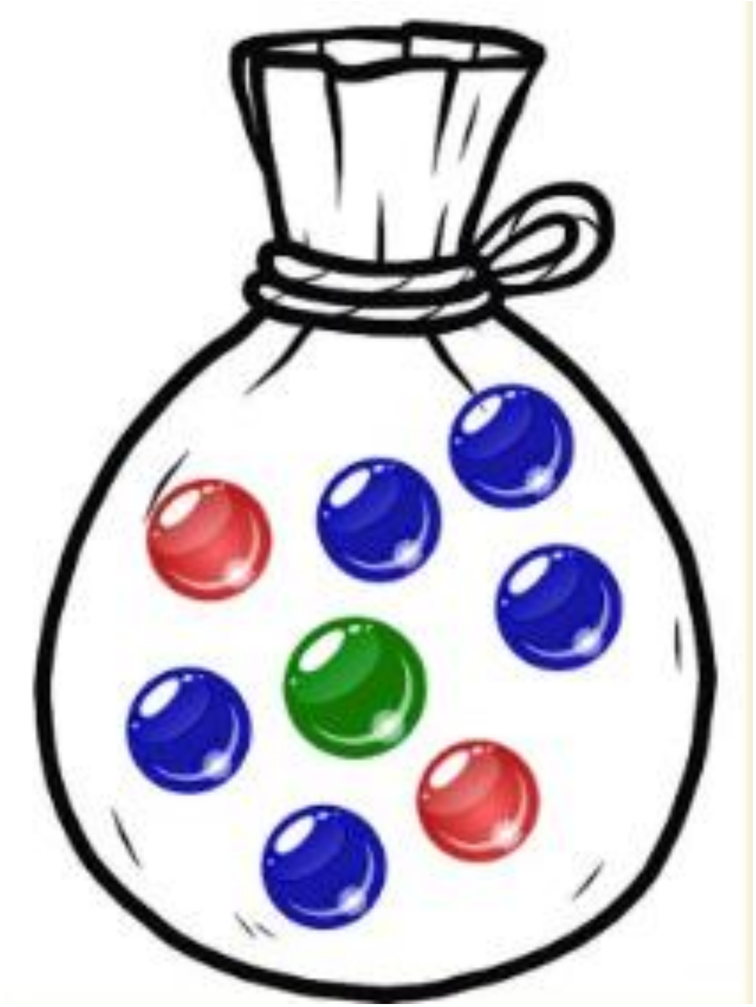
dog gone 

dog
dice
the dice,
bones & bingo
game for 2-4
players ages
5 and up
DOG GONE


DOG
GONE

First we draw a marble, then we flip a coin.

G



Draw the tree, assign the probabilities. Calculate the probability of each possibility.

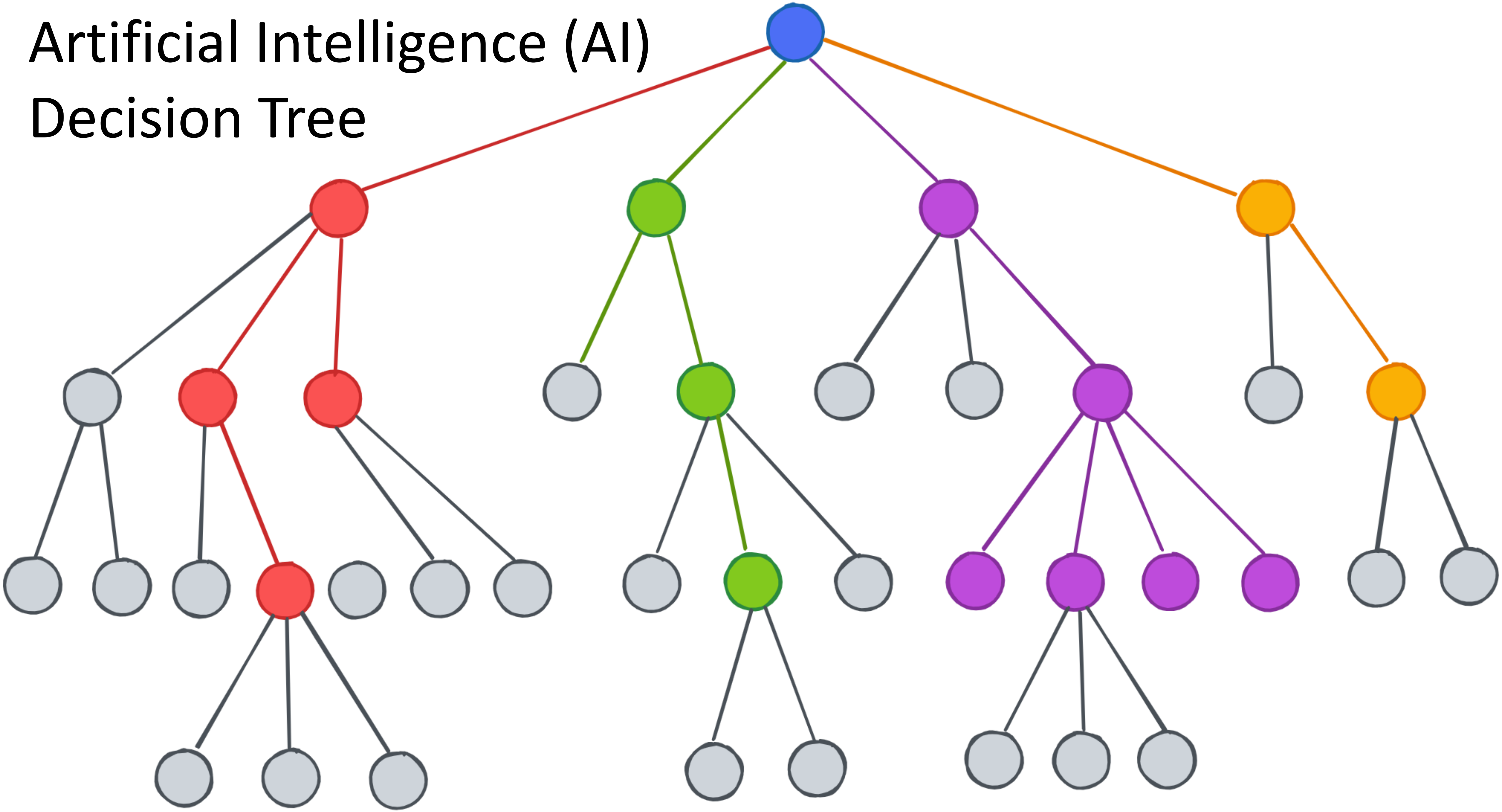
When is this used? Deep Blue, Chess Playing Computer

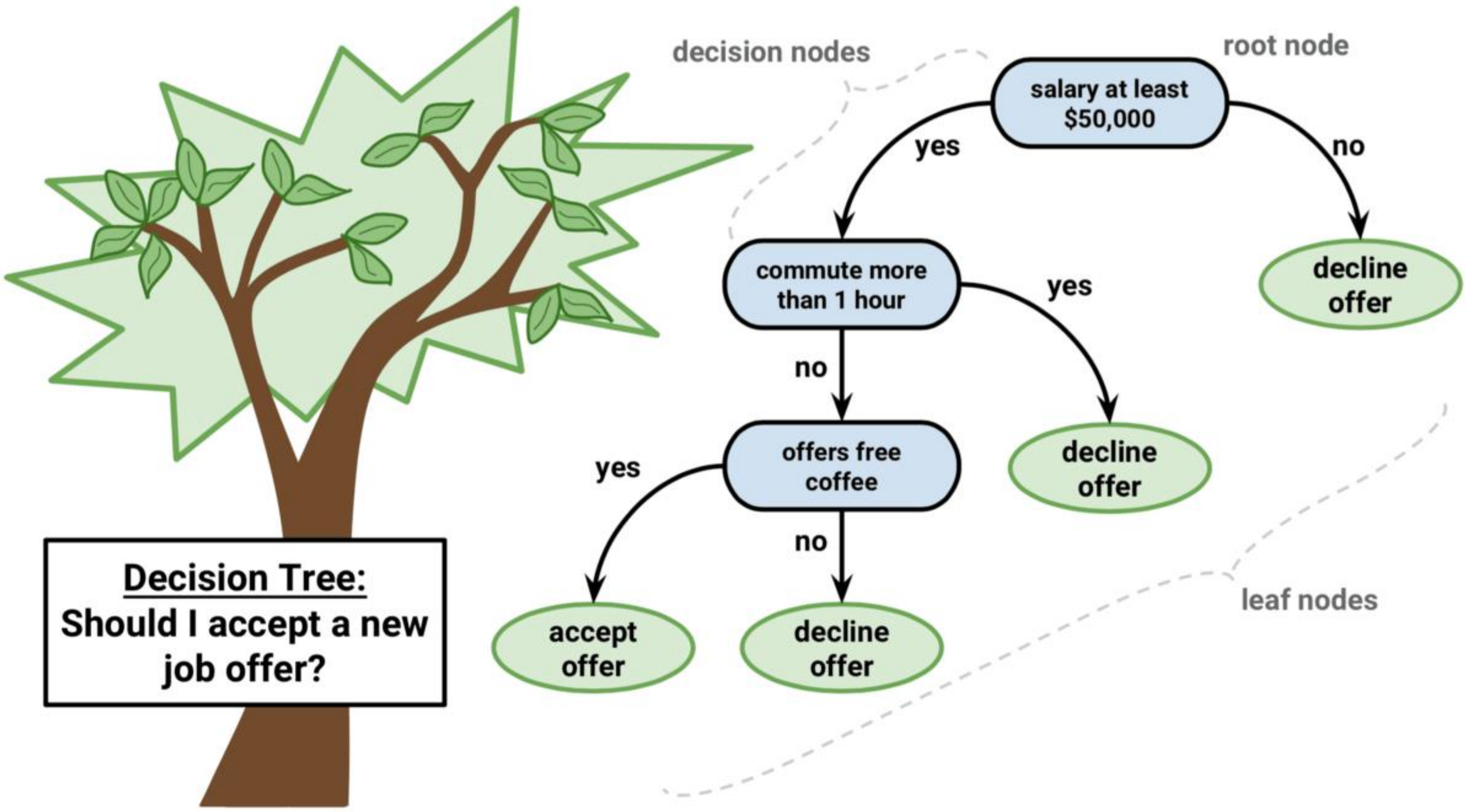
1996: First Match, wins one game – match to humans.

1997 – computer wins 3/5 games.


1. e4 c6 2. d4 d5 3. Nc3 dxe4

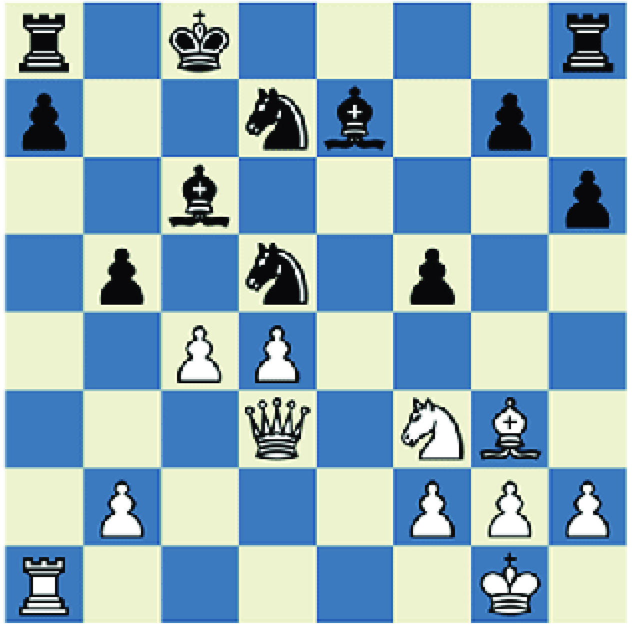
Artificial Intelligence (AI)
Decision Tree

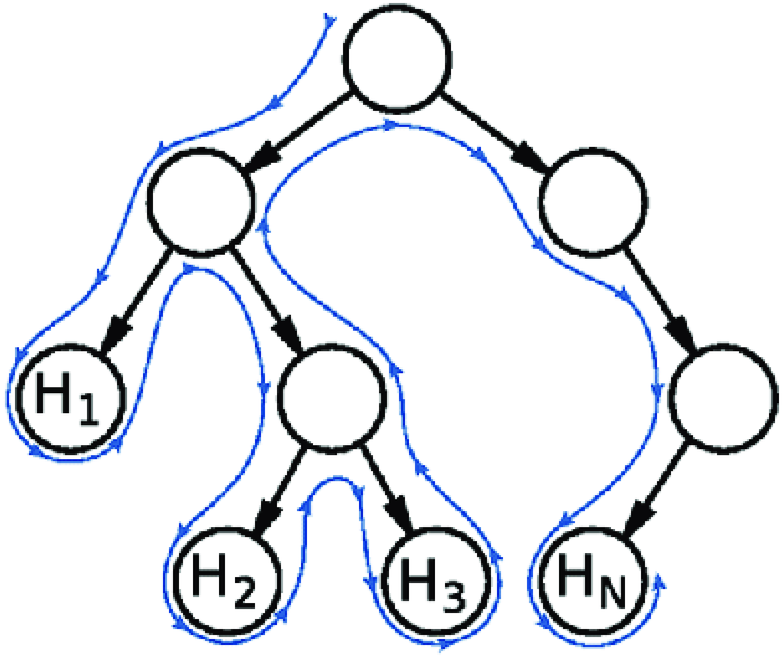


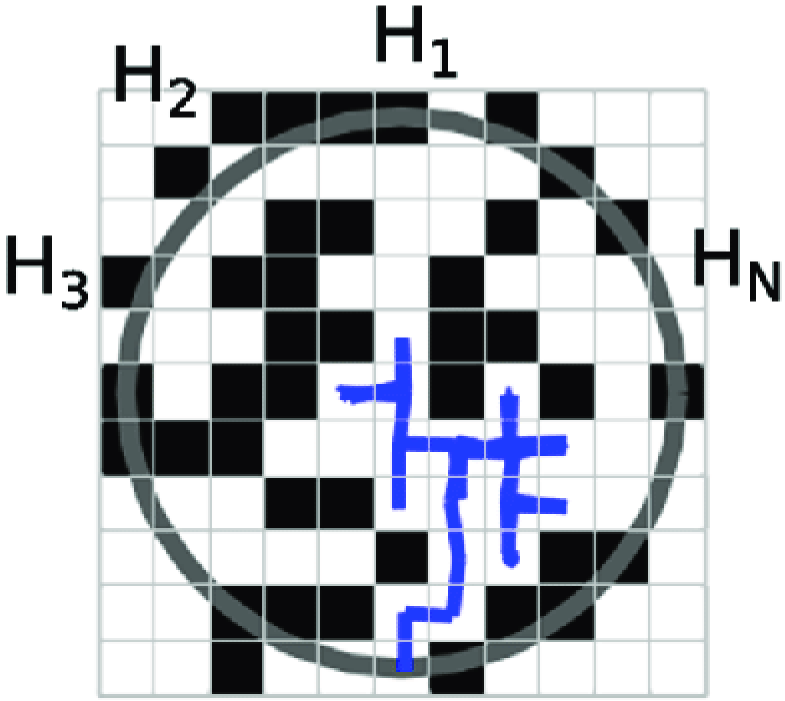

decision nodes
root node
salary at least $50,000
yes
no
commute more than 1 hour
yes
decline offer
no
offers free coffee
decline offer
yes
no
accept offer
decline offer
leaf nodes
Decision Tree:
Should I accept a new job offer?

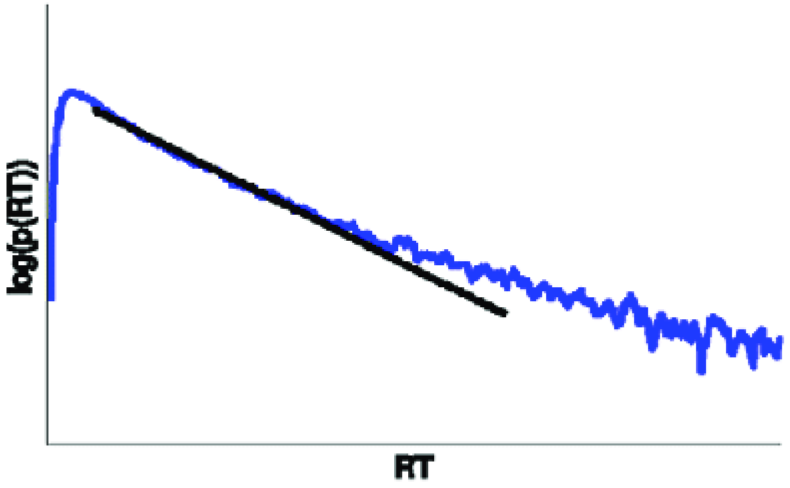
Deep Blue's Decision Tree was Huge – 200 million moves analyzed per second.

Maybe 20 levels?

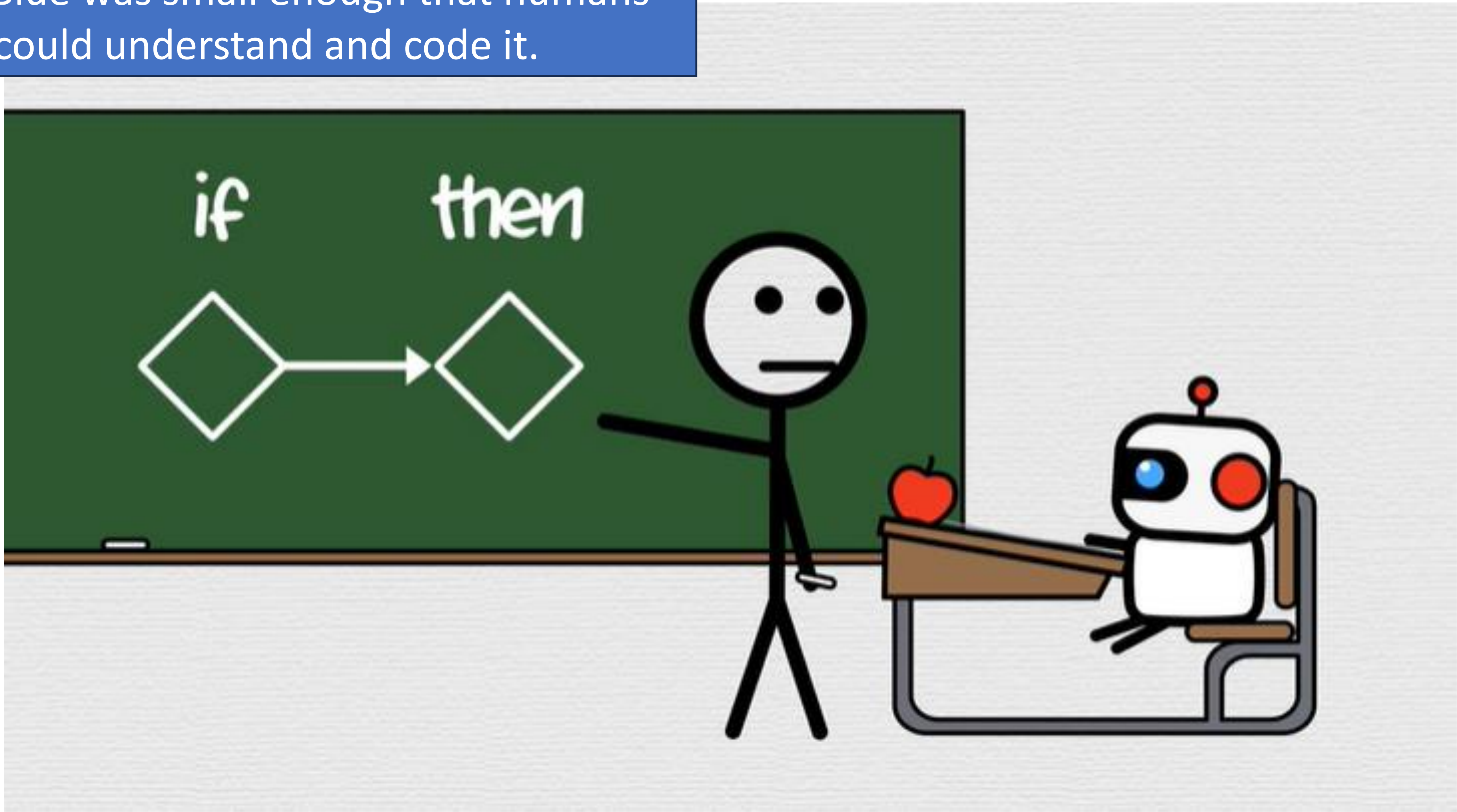








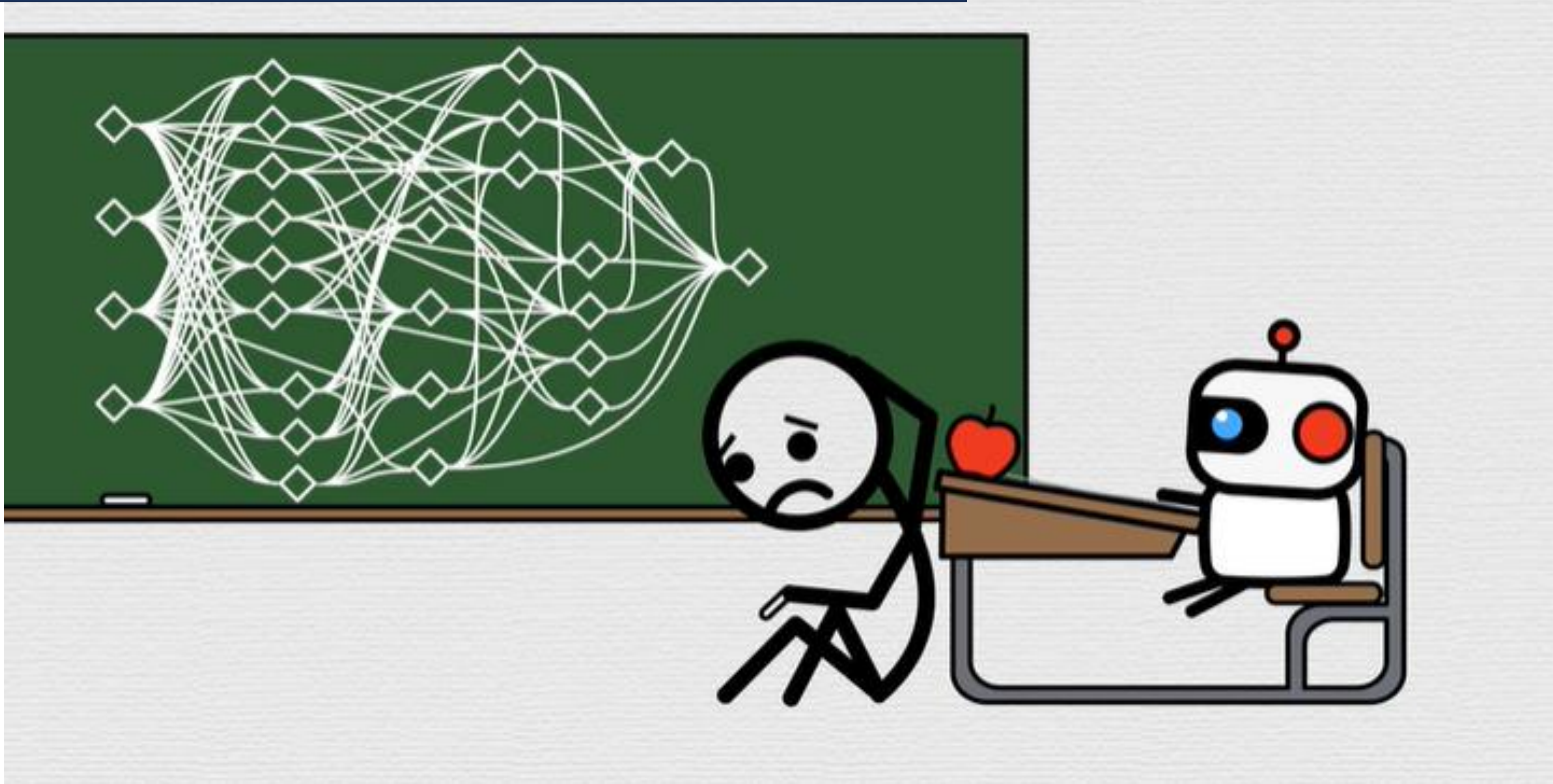
Deep Blue was small enough that humans could understand and code it.



2004 – DARPA Grand Challenge
240 km race in desert
winner? Carnegie Mellon University's Sandstorm completing 11.78 km

Humans Coded

Modern decision trees are MUCH more complex. They have move events (levels). Each branch has a different event.



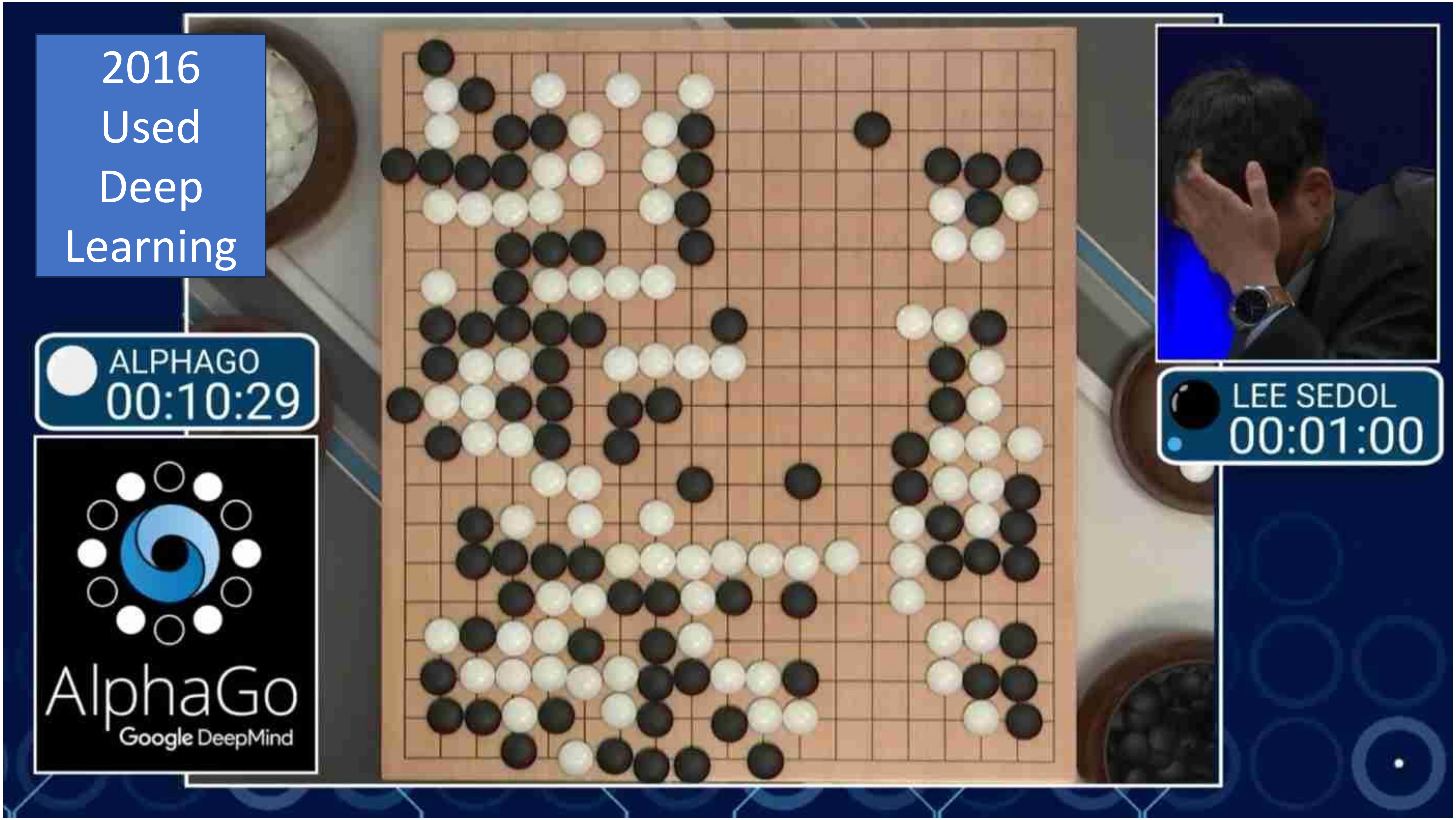
2005 – five cars finished.
Stanley won taking 6 hours and 54 minutes.
Average speed = 35 km/h

Machines Coded – Machine Learning

March 28, 2012

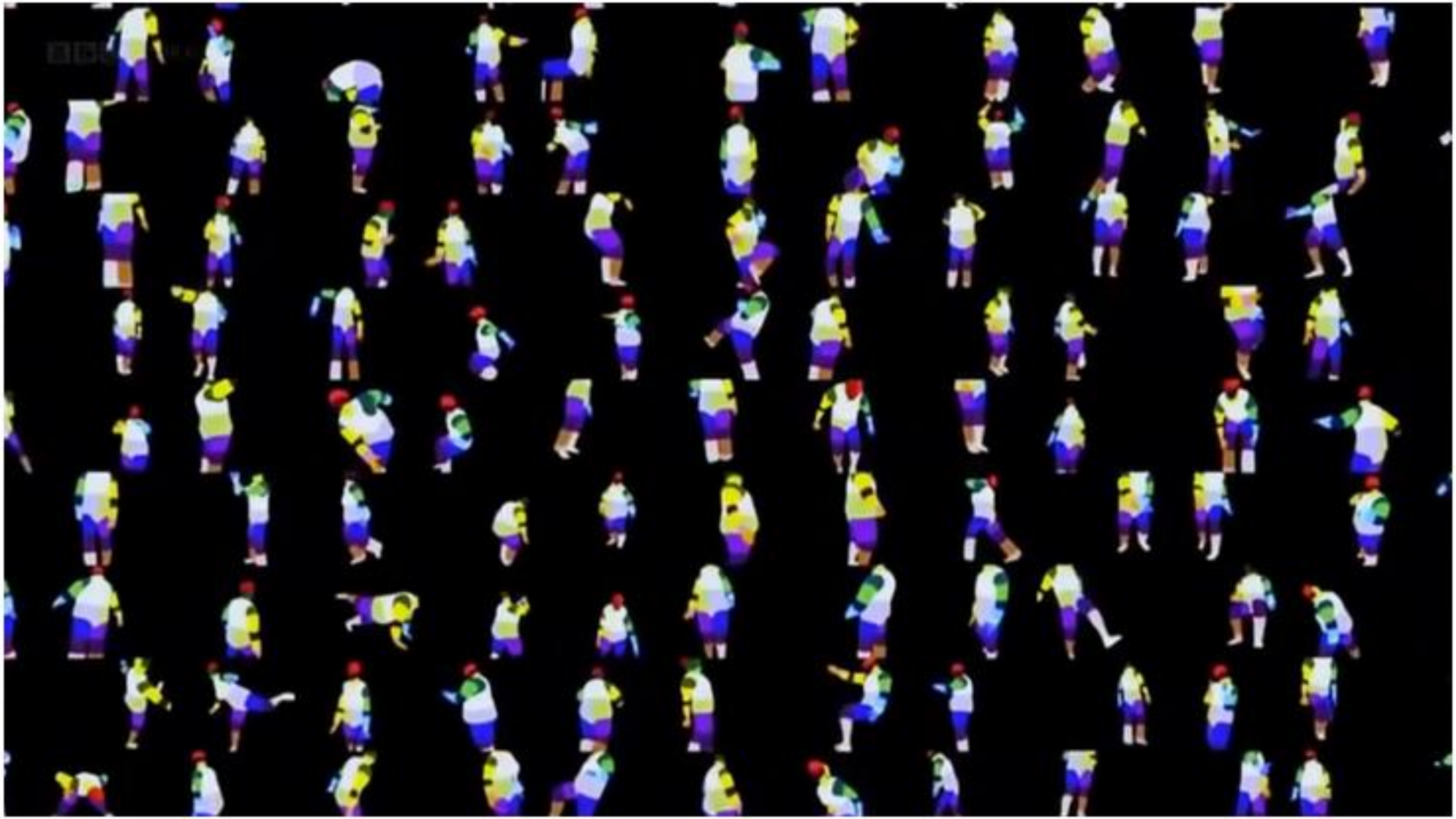
Google car
WAYMO

2016
Used
Deep
Learning

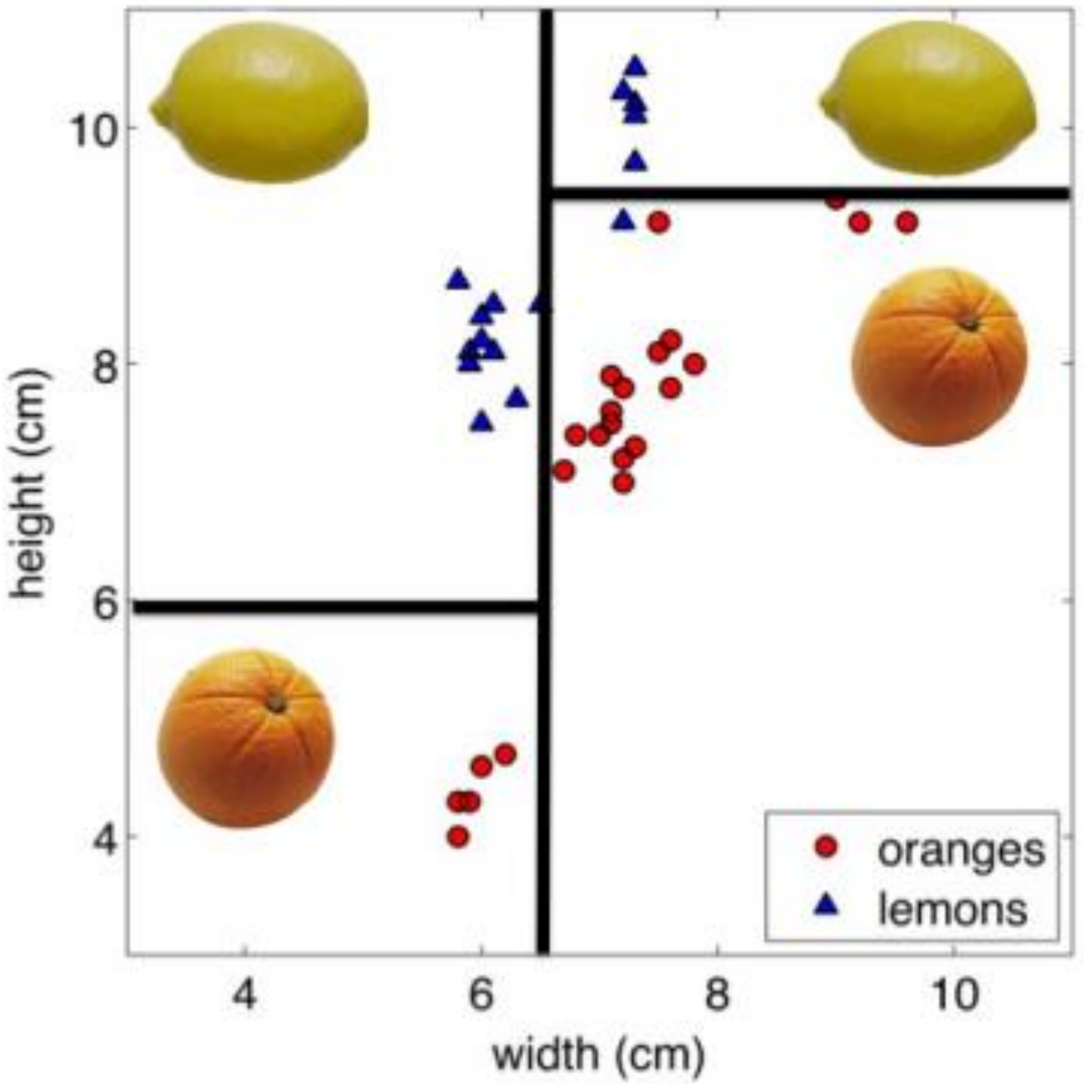


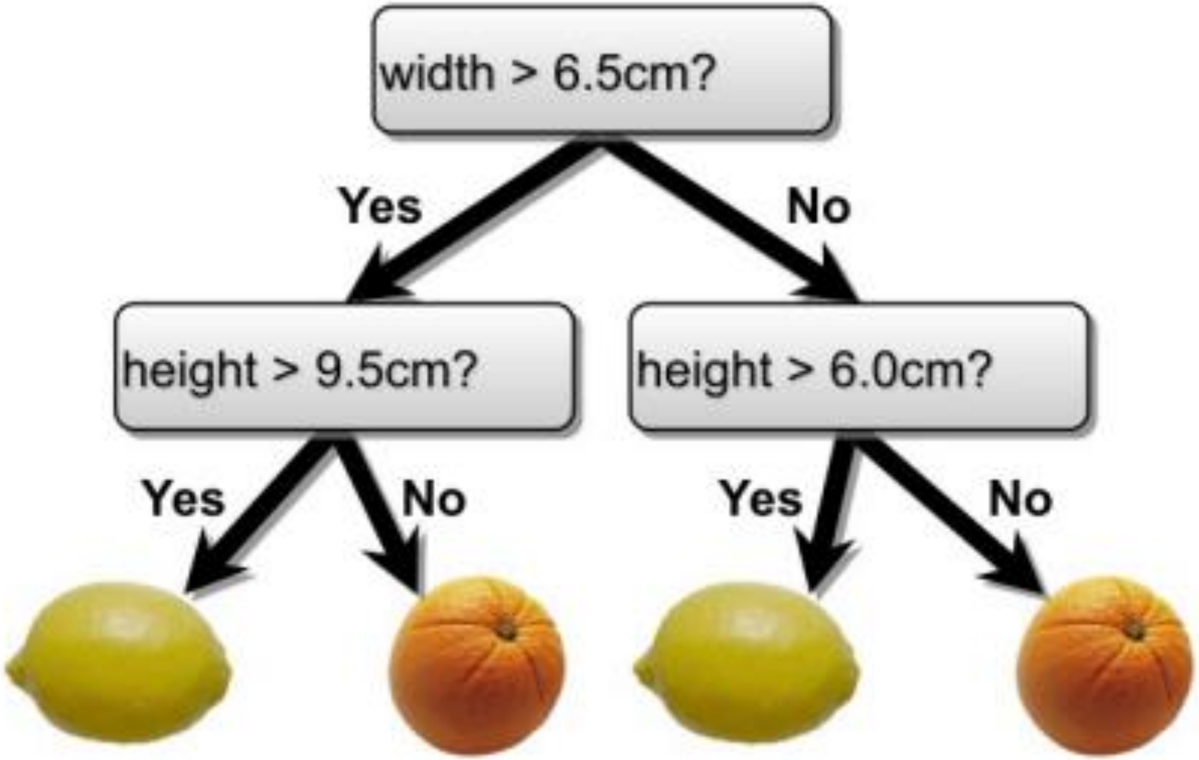
Xbox Kinect

"The Secret Rules of Modern Living Algorithms" - Documentary.



In Deep Learning, We don't pick the questions, the AI does.





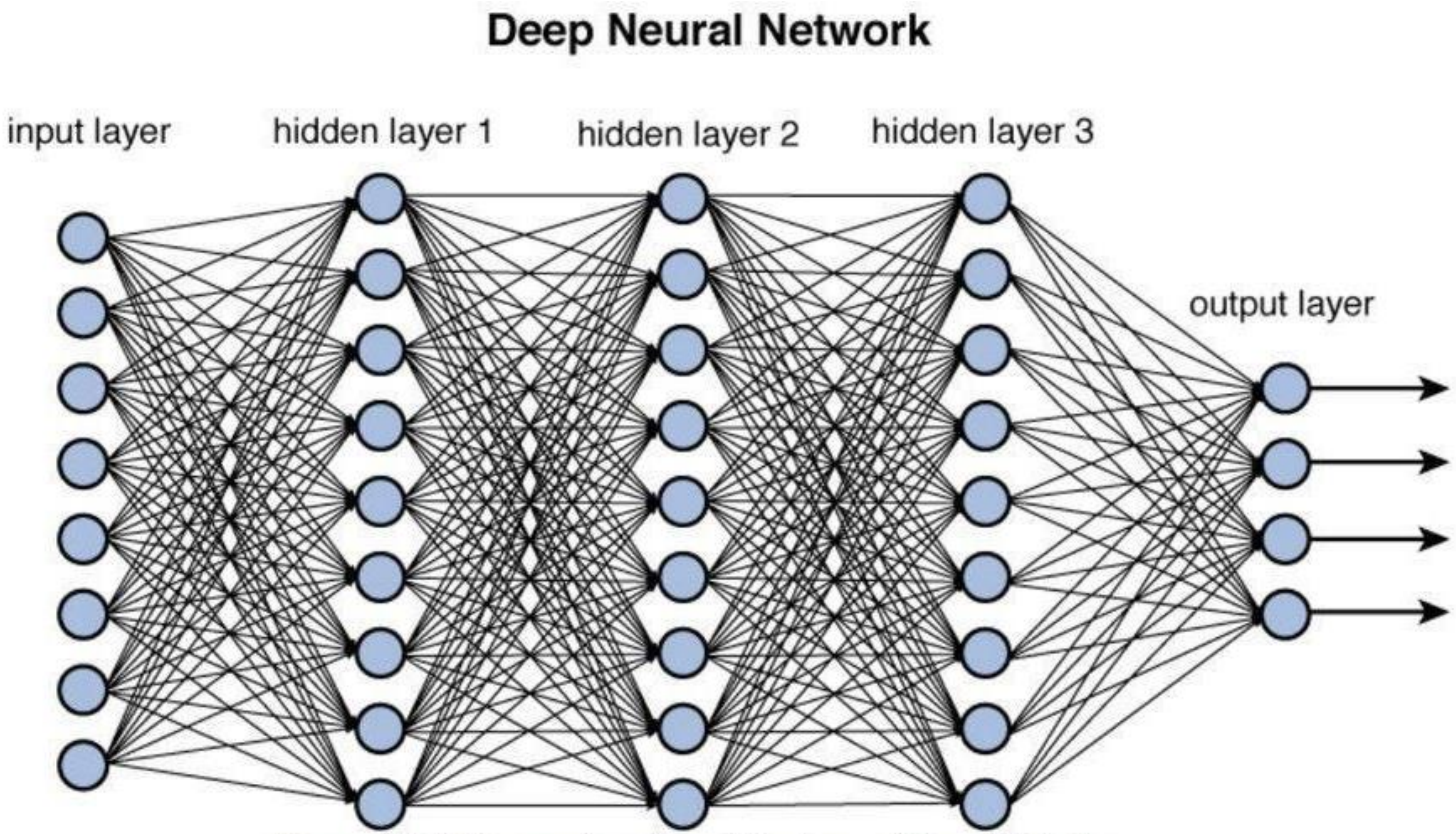
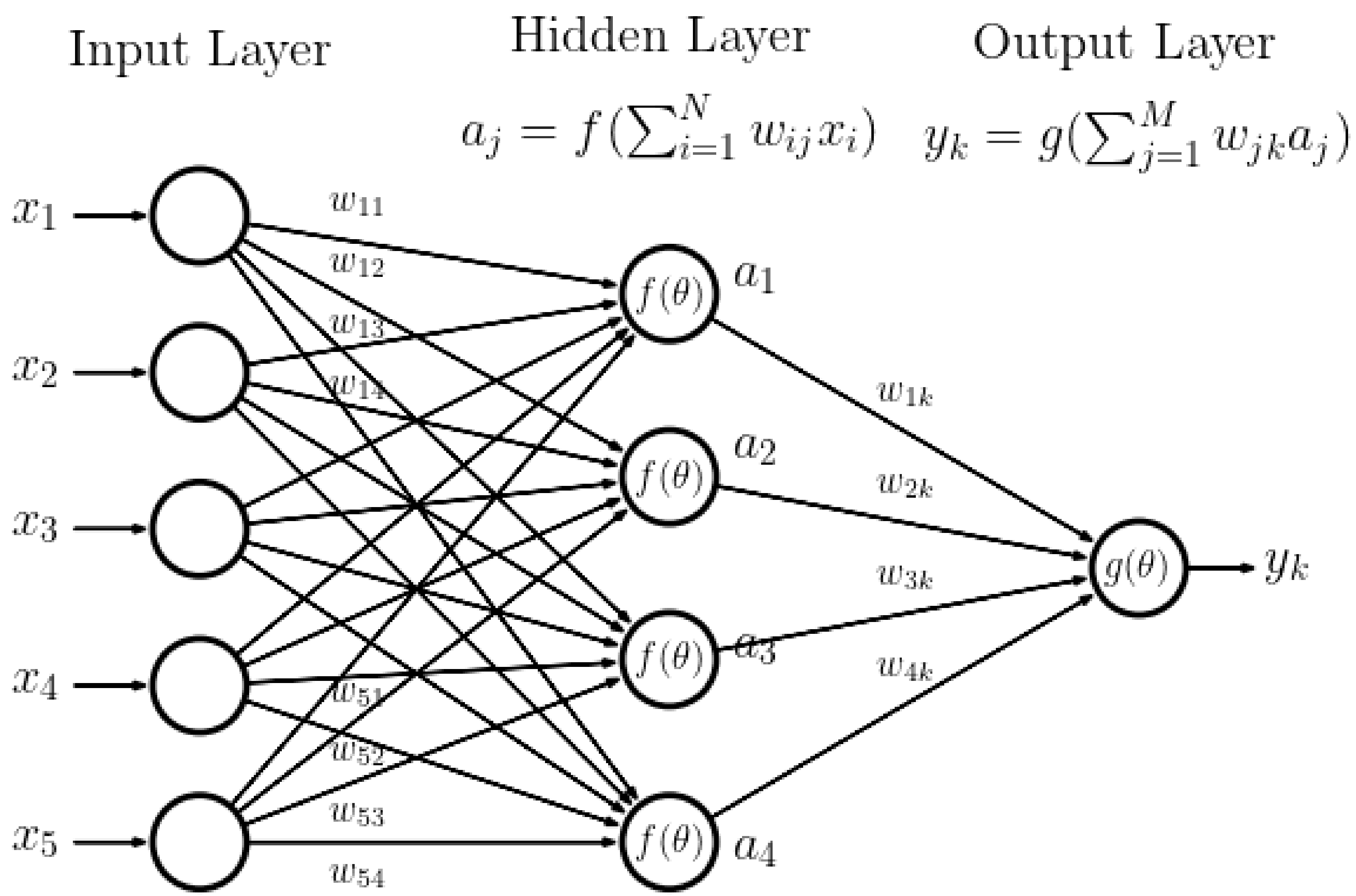


Figure 12.2 Deep network architecture with multiple layers.


Input Layer
Hidden Layer
Output Layer
$a_j = f(\sum_{i=1}^{N} w_{ij} x_i)$
$y_k = g(\sum_{j=1}^{M} w_{jk} a_j)$
x_1
x_2
x_3
x_4
x_5
w_{11}
w_{12}
w_{13}
w_{14}
w_{51}
w_{52}
w_{53}
w_{54}
$f(\theta)$
a_1
a_2
a_3
a_4
w_{1k}
w_{2k}
w_{3k}
w_{4k}
$g(\theta)$
y_k