1. [Start of transcript. Skip to the end.](https://courses.edx.org/xblock/block-v1:ColumbiaX+CSMM.101x+3T2020+type@vertical+block@dadd396ac18e4e19b3dc790e28efb24e?show_title=0&show_bookmark_button=0#transcript-end-c323d356e1104e7e83af931108e23b60)
2. Welcome back.
3. In this lecture, we'll talk about logical agents,
4. also known as knowledge-based agents.
5. These are agents that can represent the world
6. with some formal representations and make inference to derive
7. in your presentations.
8. It also can decide what to do based on the representations
9. of the world.
10. Let's start with this compelling definition
11. from John McCarthy in his article concept of logical AI
12. published in 2000.
13. So the idea of logical AI is that an agent can represent--
14. so there's some representations-- of what?
15. Of the world, of the goals, and the current situation
16. by sentences in logic.
17. And decide what to do by inferring that a certain action
18. or course of action is appropriate to achieve
19. the goal.
20. So we have here some representation, some sentences
21. in logic of the world, of the actions, of the percepts,
22. and also some inference about what
23. to do based on the observations and the current situation
24. of the world.
25. So I found this like a very good definition
26. of what's a logical AI that actually uses
27. the key components of what's logical
28. AI, knowledge of presentations and inference.
29. To put things back in context, let's review
30. Percy Liang's diagram for level of AI intelligence systems.
31. So we have an x-axis, and we classified roughly intelligence
32. systems into reflex based agents, state based agents,
33. variable based agents, and logical based agents.
34. In the previous weeks we got to see many of them.
35. So we most recently talked about learning
36. agents, in which an agent can learn from past experience.
37. We also saw search agent, in which an agent can represent
38. the world with states, and each state
39. has a black box that doesn't have any further information.
40. We also visited CSPs, Constraint Satisfaction Problems,
41. where a state is represented by an assignment of values
42. to variables.
43. So with CSPs we started looking into a designing agent that
44. can actually have representation of the world
45. and representation of the algorithm completely different.
46. So we started separating the domain from the algorithms.
47. So we saw for CSPs, for example, that we
48. can formalize the problem and then use any kind of solver.
49. So defining the program is a part of the domain knowledge.
50. It's embedding the domain knowledge
51. in some representation.
52. Then inference could be done with any known inference
53. algorithm or learning algorithm.
54. We developed logical agents as a class
55. of agents that can combine information
56. that they embed in general in some sort of data base
57. called knowledge base.
58. So this information is being told to them.
59. And they can use it to make inference and adapt
60. to new environments.
61. So we are talking today about logical agents,
62. which is on the high level of intelligence on the axis.
63. So the key concept here is really
64. the concept of knowledge, which is information
65. about the world that allow the agent to choose
66. the good actions or decisions.
67. So knowledge is what?
68. Is a set of sentences.
69. The set of sentences is expressed
70. in some knowledge representation or some language.
71. It has to be formal to avoid any ambiguity.
72. So knowledge is a set of sentences about the world.
73. A sentence is what?
74. Is an assertion or actually a proposition about the world.
75. A logical agent or knowledge-based agent
76. is composed of two elements.
77. And these are a knowledge-base and an inference mechanism.
78. So we represent that with the data base.
79. So we have some knowledge-based that is actually
80. represented by a set of sentences
81. in some formal representations or language.
82. And this is completely domain specific,
83. which means that this will depend a lot on the domain
84. to define what are the sentences that
85. are relevant for the intelligent agents.
86. We also have a set of algorithms that
87. can help us make inference.
88. And these are actually completely
89. independent of the domain.
90. So we have algorithms from AI.
91. These could be search algorithms, or CSP algorithms,
92. or inference algorithms that help us actually
93. derive new sentences from existing sentences,
94. given some new percepts to decide what to do.
95. So we have algorithms.
96. And this are actually domain independent.
97. And these two components actually interact, of course.
98. And as the agent is evolving in the environment,
99. learning from the environment and making decisions,
100. this happens through inference.
101. And then this inference can lead to new sentences that
102. are being added to the domain knowledge, called
103. knowledge-based.
104. Given these two elements, a knowledge-based agent
105. should be able to represent the states, the action,
106. incorporate the precepts, update the internal representation
107. of the world, and even deduce or infer
108. hidden properties of the world.
109. And finally, the goal is to deduce the appropriate actions
110. for the agents.
111. So this kind of approach is called a declarative approach,
112. in which we build an agent by giving it information.
113. And this information is through adding new sentences
114. in some representation.
115. **This is through the procedure Tell.**
116. Tell the agent what it needs to know.
117. Initially, we have some domain knowledge,
118. and we could translate that into sentences
119. that we feed into the knowledge base.
120. And then, as the agent interacts with the environment,
121. it can build its own sentences and enrich that database.
122. And the second is to create what is known.
123. This is through the procedure Ask.
124. Ask itself what to do.
125. So the answers comes from--
126. follow from the knowledge base through inference
127. and deduction, in general, logical deduction.
128. So here is the knowledge-based agent function
129. that actually has as input the percepts and returns an action.
130. It also maintains a database on knowledge base called Kb,
131. and a counter to indicate time.
132. All right.
133. So the intelligent agent here does three things.
134. And the first one is to actually translate the percepts
135. into a sentence, and tell that to the database
136. or to the knowledge base.
137. So this would be through translating, making
138. to precept sentences from the precept,
139. and tell the database that this is what I got.
140. OK.
141. The second one is to decide what to do.
142. So the agent will ask or query the knowledge base,
143. given the new percepts and the prior knowledge
144. about the environment, what kind of actions it needs to take.
145. So this would be where actually the agent
146. would be deliberating and making inferences about what to do.
147. The third component is tell the database
148. what action the agent choose.
149. So from the program of the agents, we have--
150. and from the prior knowledge and from the new precept--
151. we have some decision about what actions to take.
152. And this will be added to the knowledge base.
153. Right?
154. So the process will repeat as times goes on.
155. This is the general structure of the agents.
156. We might think that actually this
157. looks very similar to a black box model or a state
158. based model, such as searching from how to go from point A
159. to point B. All right.
160. So we-- in this case, the agent was just searching.
161. And it seems like here we have functions
162. like Tell and Ask and Make Precept, so it seems atomic.
163. And the fact that it's not--
164. basically, for an agent to go from point A to B,
165. it doesn't use any domain knowledge.
166. However, if an agent knows that to go from A to B,
167. it can use bridge 1 or bridge 2, this information
168. could be leveraged by the agent to find the best routing
169. or find the route from A to B. So this element
170. that we have here actually more granular than what
171. it appears in the algorithm.