Deepmind published a paper on nature, the paper describes the latest version of AlphaGo, named AlphaGo Zero.

AlphaGo Zero does not use any human experience, that is, does not use supervised learning for pre-training, and completely uses reinforcement learning to learn. In the previous AlphaGo, the network was pre-trained in AlphaGo using a lot of games between humans as training data. After training, AlphaGo used reinforcement learning to learn further.

People will also make several steps forward in the process of playing chess. People are relying on intuition in the process of deduction (similar to the action selection probability of AlphaGo Zero network output). For example, at the root node, the human feels which step to take will be compared. Well, you will take this step in the deduction, and then analyze yourself where the other person may be, and then take another step based on your intuition... but in the process of deduction, you may feel a change after a few steps. Poor (the state value of AlphaGo Zero's network output is relatively low), so he recaptures his head, and the choice of action is not the same when deciding. When a few steps away, he finds that the situation is getting better and that humans will think of this as It's better to play chess, but I'm afraid that my deductions are not comprehensive enough, there may be better ways to go, so I'm going to deduce a lot of different explorations. Finally, humans will find a better way to go. Law, but in reality, such deductions require too much computing power, and only computers can accomplish it. Humans can only deduct a few moves.

Computers have far more computing power than the human brain, so it is entirely possible for artificial intelligence to develop a mental model that is different from the human brain.