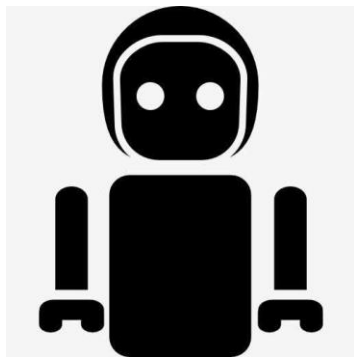
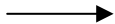
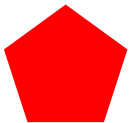
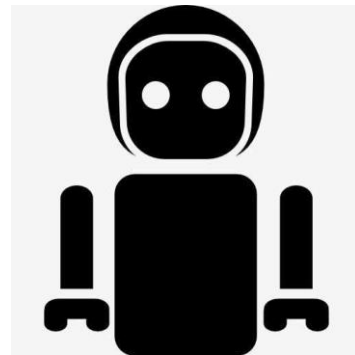
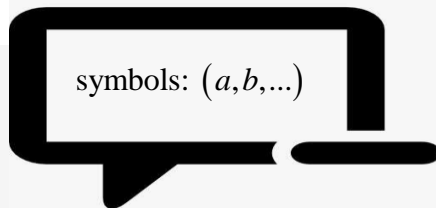


object: (shape, color)



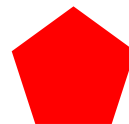
Speaker agent



Listener agent



predict result: (shape,color)



...