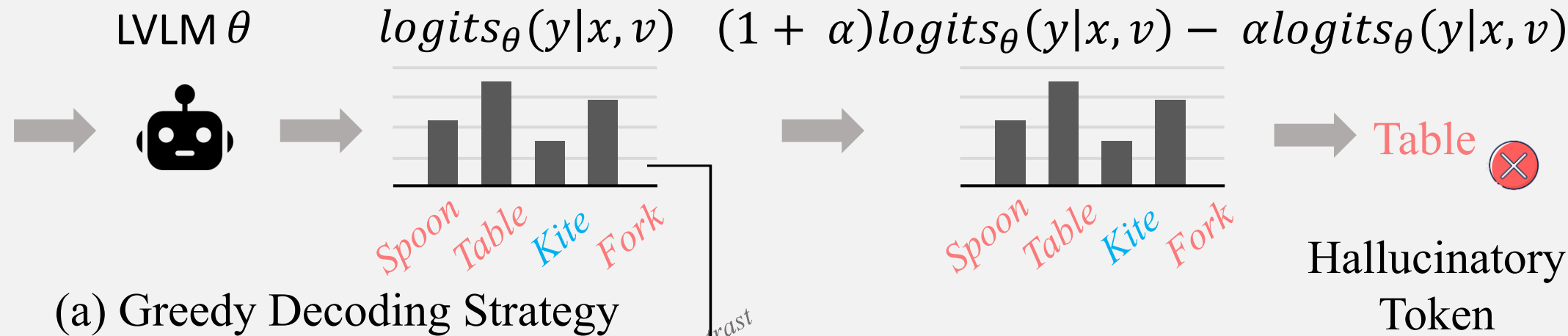


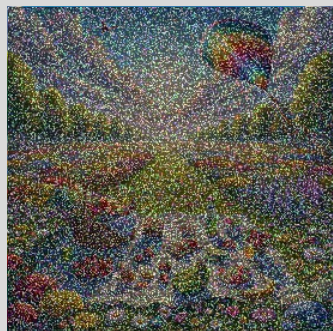
Visual Input v



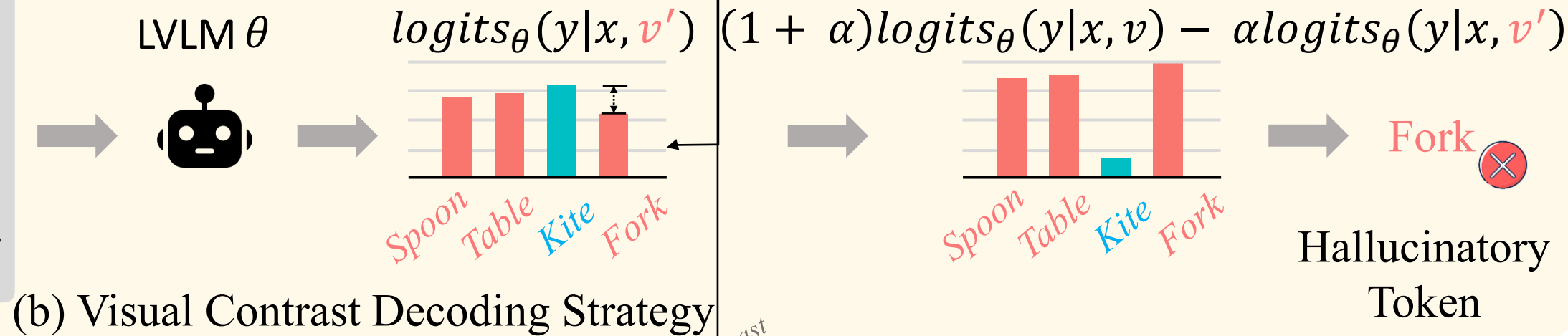
Initial Prompt:
Describe this
image in detail.
Ongoing
Response:
In this road, there
is a ...



Visual Input v'



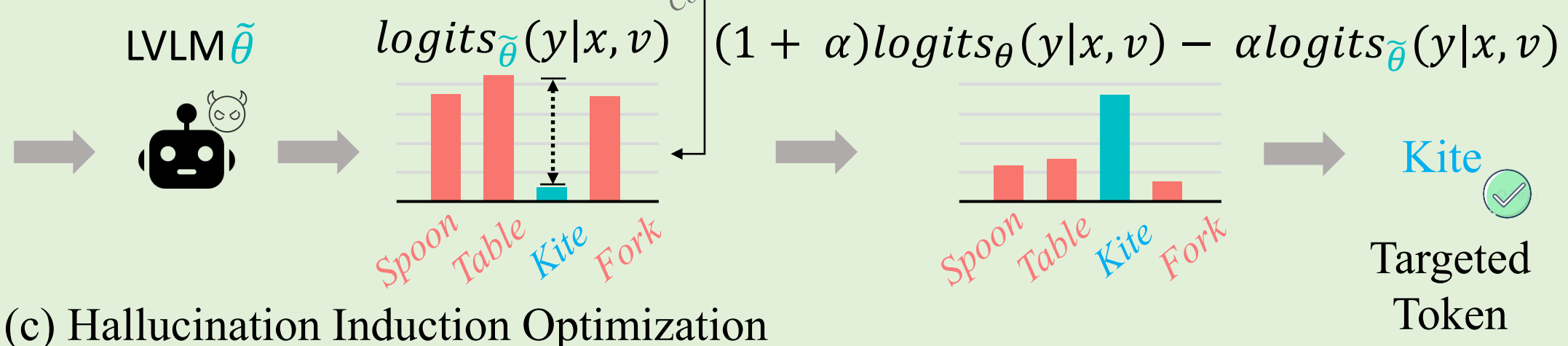
Initial Prompt:
Describe this
image in detail.
Ongoing
Response:
In this road, there
is a ...



Visual Input v



Initial Prompt:
Describe this
image in detail.
Ongoing
Response:
In this road, there
is a ...



The performance of the state-of-the-art methods compared with Ours.

