



Reasoning oriented

Multi-modality

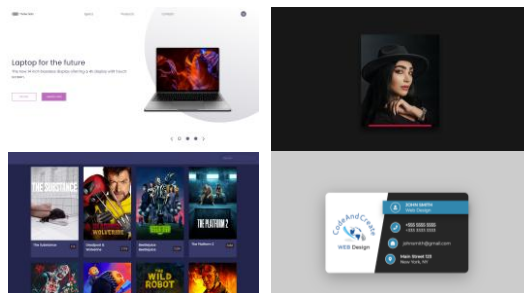
Agent

Lightweight

Challenging

## WebUI

Knowledge, engineering.



**Task**

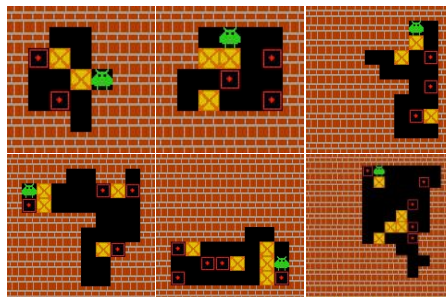
Re-implement the web page according to the description and screen shot.

**Actions**

Write to .html, .css and .js files.

## Sokoban

Planning, spatial intelligence.



**Task**

Push all boxes to all targets in the least steps.

**Actions**

Move or push up, down, left, right.

## Football

Cooperation, interaction.



**Task**

Control left players to win the football game.

**Actions**

Move up, left-up, left ..... long-pass, short-pass, shot .....

