**The motion of the agent in London**

The agent moves on the road, having a map of the city. Starting from the initial state of the agent must choose an action (forward, backward, left, right) at each step of the movement. Interaction with the environment ends when the agent leaves the city, reaching out a single target position - Tower. In each location the agent takes an optimal solution (forward, backward, left, right).

Baker st.

Old st

Monument.

King’s Cross st.

Green Park.

Piccadilly circ.

Covent Garden

Liverpool st st.

Bond st..

Oxford circ.

Chancery Lane.

Bank

Tower Hill.

Moorgate.

Environment is fully observable; the agent knows where he is. Environment is not deterministic, actions performed in it is unreliable, the movement of the agent is stochastic.