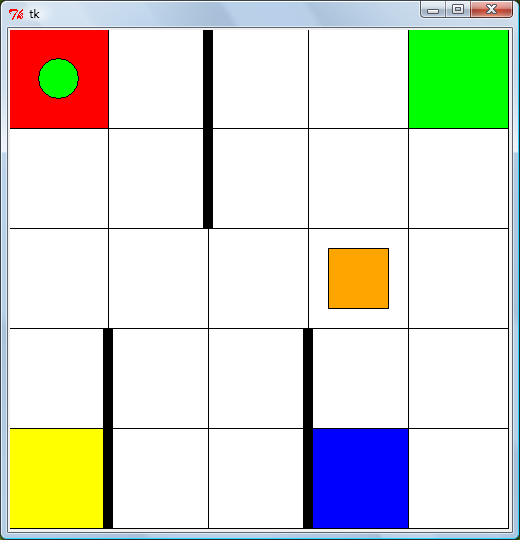
# Laboratorio 3 – Experimentos con Q-learning y Q(l)

El problema del Taxi:



Como en todo problema de RL, hay 3 componentes:

**Ambiente:** las funciones que, dado un estado acción, calculan el siguiente estado y el reward.

**Agente:** el que elige las acciones y aprende (ej., Q-learning)

**Interacción:** el modulito que le pide una acción al agente, se la da al ambiente, le devuelve el resultado al agente, chequea si se llegó al fin de un episodio, etc.

1. Implementar las funciones del ambiente que responden a las acciones: move, pick y drop.
2. Correr Q-learning por 5000 episodios.
3. Correr solveProblem con el Q resultante y ver si aprendió.
4. Plotear el reward por episodio.
5. Correr Q-learning de nuevo por 1000, 2000, etc. episodios.
6. Correr solveProblem y ver si anda.
7. Implementar Q(lambda) (ver código abajo)
8. Comparar performance Q(lambda) contra Q-learning.

Q(lambda)

