

定义

定义

环境

智能出租车 行驶在一个理想的，格子构成的城市（跟纽约很相似），路都是南北或东西向的。当然也会有其它车辆在路上行驶，但没有行人需要考虑。在每一个路口都有交通灯，可以允许车辆南北或者东西通行。这里试用美国交通法规：

- 绿灯亮时，仅在十字路口无直行来车时才能左转。
- 红灯亮时，如果无直行来车左转，或左方来车直行时，则可以右转。

要了解左转时如何正确让来车优先通行，可以观看此视频官方驾驶员培训视频

(http://v.youku.com/v_show/id_XMTU00TY0MDQyOA==.html) 或者此视频充满激情的讲解 (http://v.youku.com/v_show/id_XMTU00TY0MDg2NA==.html)。

输入和输出

假设智能出租车根据乘客的出发地和目的地指定了一条路线。路线按所经过的每个十字路口划为一个路径点，而且假设在任意时刻，智能出租车均处于某个十字路口。因此，到达目的地前，下一个路径点将始终在智能出租车的正前方、左方、右方、或正后方的一个街区外，或者就在原地（已到达目的地）。

智能出租车只能看到自己当前所在的十字路口（抱歉，没有精确的GPS，没有全球定位）。它能感知前进方向的红绿灯是否为绿色，通往十字路口的每条道路是否有车，以及车辆向哪个方向行驶。在任意时刻，智能出租车要么停留在当前的十字路口，要么向前、向左或向右移动一个街区，但不能后退。

除此之外，每段行程都有限定时间，随着每次决定而减少（乘客希望尽快到达目的地）。如果限定时间到，智能出租车还没有到达目的地，则当前行程失败。

奖励和目标

智能出租车每成功完成一段行程就会获得奖励。“成功完成行程”的意思是，在预先指定的时间限制（依照指定路线推算）内，将乘客送到预定的目的地（某个十字路口）。

此外，如果它在十字路口进行正确的移动，则会获得较小的奖励；如果进行错误的移动，则会获得小惩罚；如果违反交通规则和/或造成事故，则会获得较大的惩罚。

为智能出租车设计 AI 驾驶智能体。它应在每个时间步 t 接收上述输入，生成相应的移动作为输出。根据它获得的奖励和惩罚，智能体应学习最优策略，以便在城市道路上行驶、正确遵守交通规则并尝试在目标时间内到达目的地。

定义
