

# ЗА ДА НАМЕРИМ МОДЕЛ, ТРЯБВА ДА СЕ ОТКАЖЕМ ОТ СЪВЪРШЕНСТВОТО

Dimiter Dobrev  
Institute of Mathematics and Informatics  
Bulgarian Academy of Sciences  
*d@dobrev.com*

**При Reinforcement Learning търсим модел на света. Обикновено търсим модел, който да казва всичко или почти всичко. Тоест, търсим съвършен модел (тотален детерминиран граф) или пълен модел (Markov Decision Process). Задачата за намирането на такъв модел е твърде амбициозна и при сложните светове тази задача е на практика нерешима. За да решим задачата, ще заменим съвършените и пълните модели с Event-Driven модели.**

Само децата вярват, че могат да разберат света напълно и да отговорят на всички въпроси. Ние възрастните знаем, че никой не може да разбере всичко и да знае всичко. Затова, когато търсим модел на света, ние сме готови да се откажем от някои прекалено силни изисквания.

Първо, ще се откажем от предположението, че моделът не може да бъде подобрен (това предположение се нарича *Markov property*).

Второ, ще се откажем от предположението, че моделът е детерминиран. От второто предположение вече сме се отказали в Markov Decision Process (MDP). Все пак, в MDP се предполага, че когато имаме недетерминираност, то всяко недетерминирано разклонение си има точно определена вероятност. Това е третото предположение, от което ще се откажем. Ще предполагаме, че тази вероятност не е точно определена, а е в някакъв интервал  $[a, b]$ . Може дори въобще да не знаем каква е тази вероятност, т.е. интервалът да е  $[0, 1]$ .

Четвърто, ще предположим, че модела не отчита всички събития, а само някои от значимите. MDP отчита всички действия на агента. Тоест, това е Action-Driven модел. Ще заменим действията със събития и ще преминем към Event-Driven модел. Тези модели са много по-стабилни, защото не променят състоянието си на всяка стъпка, а само при настъпването на някое от наблюдаваните събития.

Пето, няма да разделяме описанието на света и на агента, а ще ги опишем заедно като една единна система. При MDP имаме описание само на света, което не включва описание на агента. Няма нищо лошо в това описанието на света да е отделено от описанието на агента. За съжаление това разделяне ни излиза твърде скъпо и то не ни позволява да преминем към Event-Driven модел. Това е причината, поради която ще се откажем и от това предположение.