

Авторска справка на приносите в дисертацията и публикациите на Димитър Добрев за придобиване на ОНС „Доктор“

1. Дисертация

Изброяване на приносите на автора:

1. **Предлага се неформална дефиниция на ИИ** и тази неформална дефиниция днес е широко разпространена (в момента това е първият резултат, който Google връща при въпроса *Definition of Artificial Intelligence*).

2. **Дава се строга математическа дефиниция на ИИ**. Предложената строга дефиниция не е изцяло наша заслуга, защото тя е едно подобрене на дефиницията на ИИ, която първоначално е създадена от Hernández-Orallo през 1998 (Hernández-Orallo et al., 1998) и която съществено е подобрена от Marcus Hutter през 2000 (Hutter, 2000).

3. **Въвежда се език за описание на светове**, при който описанието може да се търси автоматично без помощта на човек. Има и други езици за описание на светове, които са създадени преди нашия език, но предимство на предложението от нас език е, че при него описанието на света може да се търси автоматично, докато при другите подходи се предполага, че има човек, който е разбрал света и който ще го опише на съответния език.

2. Статиите (Част 1)

Dobrev, D. (2000). AI - What is this. *PC Magazine - Bulgaria*, 11/2000, pp.12-13.

1. **Неформална дефиниция на необучен ИИ**. Първата неформална дефиниция на ИИ е „Тестът на Тюринг“, но това е дефиниция на обучен ИИ. Първата неформална дефиниция на необучен ИИ е направена от Pei Wang през 1995 г. (Wang, 1995). В (Dobrev 2000) е дадена дефиниция, която подобрява дадената от Pei Wang. Самият Pei Wang не признава това подобрене и твърди, че неговата дефиниция не се нуждае от подобрене.

Dobrev, D. (2022a). The AI Definition and a Program Which Satisfies this Definition. arXiv:2212.03184 [cs.AI].

2. **Дефиниция на ИИ, която не зависи от дължината на живота**. В (Orallo 1998 and Hutter 2000) е дадена дефиниция на ИИ, но там се предполага, че дължината на живота е ограничена с някаква константа, която се явява параметър на дефиницията.

3. **Дефиниция на ИИ, която не зависи от езика за описание на света**. В Hutter (2000) езикът е фиксиран. Тоест в Hutter (2000) се предполага, че има само един възможен начин, по който може да се опише светът.

3. Статиите (Част 2)

Dobrev D. (2022b). Language for Description of Worlds. Part 1: Theoretical Foundation. *Serdica Journal of Computing* 16(2), 2022, pp. 101-150.

1. **Event-Driven модел**. (Това понятие вече е въведено в Dobrev (2018), но там не е дадена интерпретация на ED модела. Именно интерпретацията е това,

което дава смисъл на модела, и това, което разграничава адекватните от неадекватните модели.)

2. Показваме, че *Markov decision process (MDP)* е частен случай на *ED* модел и че *ED* моделът е естественото обобщение на *MDP*.

3. *Simple MDP*. Опростяваме *MDP* и получаваме по-прост модел, който описва повече светове.

4. **Разширен модел.** Това е моделът, при който състоянието знае всичко. Чрез този модел ние ще въведем интерпретация на събитията и на *Event-Driven* моделите. (Разширеният модел е въведен в Dobrev (2019a), но там състоянието знае само какво се е случило и какво ще се случи, но не знае какво е възможно да се случи. Тоест в Dobrev (2019a) състоянието на разширения модел не знае за пропуснатите възможности. В Dobrev (2019a) разширеният модел се нарича „максимален“.)

Dobrev D. (2023). Language for Description of Worlds. Part 2: The Sample World. *Serdica Journal of Computing* 17(1), 2023, pp. 17-54.

5. **Дефиниция на понятието алгоритъм.** Представяме алгоритъма като последователност от събития в произволен свят. Представяме машината на Тюринг като *ED* модел, който се намира в един специален свят, в който имаме безкрайна лента. По този начин показваме, че новата дефиниция обобщава понятието „машина на Тюринг“ и разширява понятието алгоритъм.

6. **Език за описание на светове,** при който описанието може да се търси автоматично без помощта на човек.

04. 2024 г.

София

Изготвил:

(Димитър Добрев)