



# »Лекционен курс »Изкуствен интелект

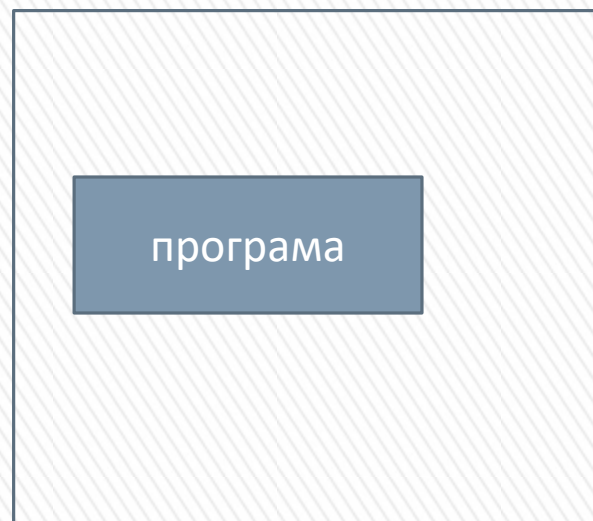


Интелигентни  
агенти

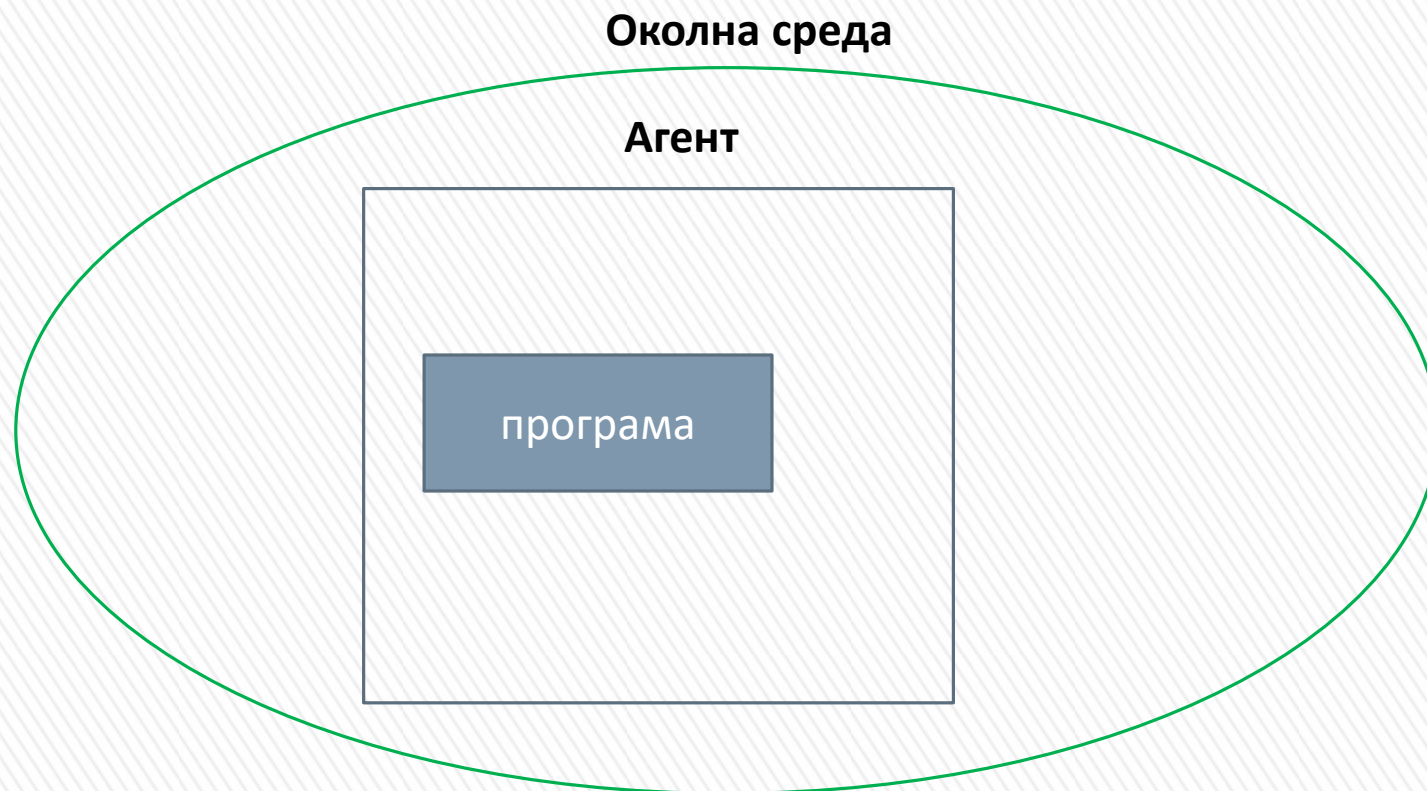


# Агенти

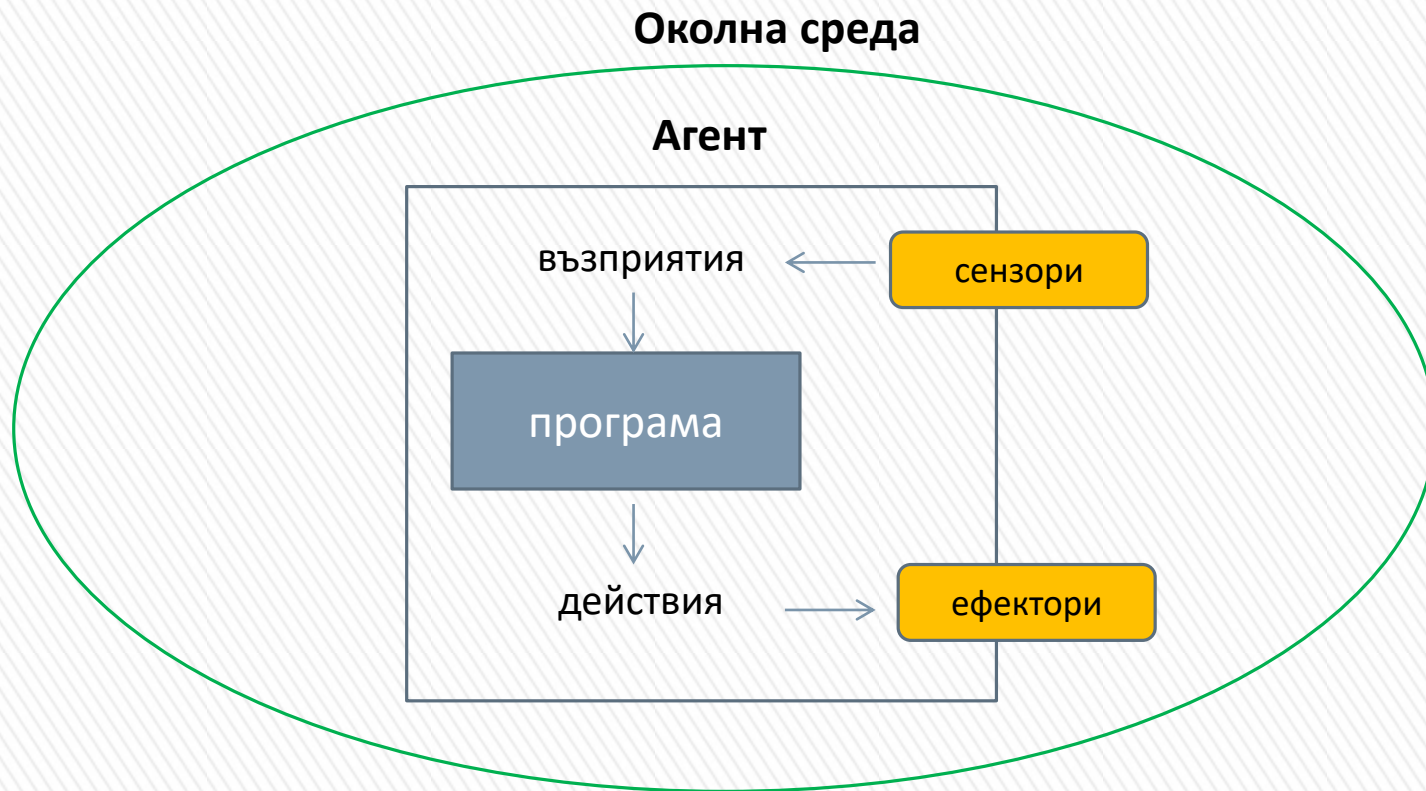
Агент



# Агенти



# Агенти



# Обща дефиниция

» Агент може да бъде:

- > Човек
- > Машина
- > Софтуерен модул
- > Нещо друго ... което действа (оперира)

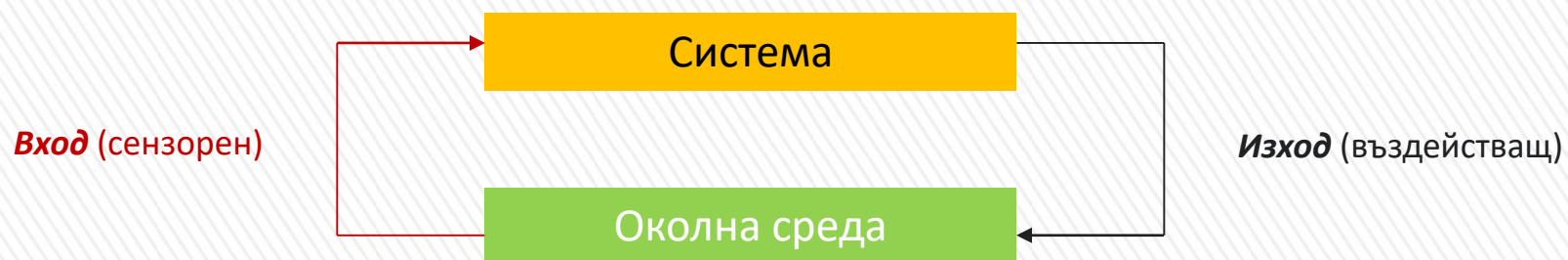


» За софтуерните технологии тази дефиниция е много обща



# Формализирана дефиниция

Компютърна система, която може да оперира  
**автономно** в някаква **околна среда**.



Cybernetics Or Control and Communication in the  
Animal and the Machine





# Бележки към дефиницията

- » В много среди, които са по-сложни (комплексни) агентите нямат пълен контрол върху тях
  - > В най-добрия случай само частичен контрол
  - > Така едно действие, извършено повторно при идентични обстоятелства може да има напълно различен ефект – в частност може да няма желания ефект
- » В общност приемаме, че средите са недетерминирани



# Бележки към дефиницията

- » Обикновено един агент има множество от оператори
  - > Те определят възможностите му за въздействие върху околната среда
  - > Не всички оператори могат да се прилагат при всички ситуации
  - > Обикновено операторите са свързани с кореспондиращи предусловия, които определят възможните ситуации за тяхното прилагане
- » Ключов проблем
  - > Кой оператор трябва да бъде приложен за удовлетворяване на проектните цели?





# Примери за агенти

## » Контролни системи

> Напр. термостат

## » Софтуерни демони

> Напр. Mail client

*Но... това **интелигентни агенти** ли са?*



# Интелигентен агент

- » Компютърна система, която може да оперира **гъвкаво** и **автономно** в някаква околна среда за постигане на набелязаните си **цели**
- » Под гъвкавост разбираме:
  - > *Реактивност*
  - > *Проактивност*
  - > *Социалност*



# АВТОНОМНОСТ

- » Агентите оперират без директна интервенция на хора или други идентичности
- » Имат някакъв контрол върху:
  - > Действията си
  - > Вътрешните си състояния



# Реактивност

- > Агентите възприемат околната си среда
- > Която може да бъде:
  - Физическия свят
  - Потребител чрез графичен потребителски интерфейс
  - Множество от други агенти
  - Интернет
  - Комбинация от всички тях
- » Поддържат непрекъснатата връзка с околната среда
- » Реагират на промените, които настъпват в средата (във времето когато носят отговорност)



# Проактивност

- » Агентите не действат само в отговор на тяхната среда
- » Те са способни да показват целево-направлявано поведение посредством поемане на инициатива
  - В общия случай ние искаме агентите да правят неща за нас
  - **Проактивност:** генериране и опитване да се достигнат цели
    - Направлявано не само от събития
    - Поемане на инициатива
    - Оценяване на удобни възможности
    - Поведение, направлявано от цели



# Социални способности

- » Агентите взаимодействат с други агенти (възможно също с хора) посредством някакъв език за комуникация между агенти и евентуално коопериране
- » В една мулти-агентна околна среда не можем да се опитваме да достигаме цели, без да се съобразяваме с другите агенти
- » Някои цели са достижими само посредством коопериране между агентите
- » Това важи и за много други компютърни среди
  - > напр. Интернет





# Характеристика на агентите

## Слаба дефиниция за агент

- » Автономност
- » Проактивност (целено-ориентирана)
- » Реактивност
- » Социални способности (комуникативност)



## Строга дефиниция за агент

Слаба дефиниция +

- Мобилност
- Правдивост
- Доброжелателност
- Рационалност



# Агенти като съзнателни системи

- » Когато се обясняват човешките дейности често се правят заявления като следното:
  - > Мария взема чадъра си, понеже тя вярва, че ще вали
  - > Георги работи усилено, понеже иска да защити дипломната си работа
- » Тези заявления използват народопсихология, чрез която човешкото поведение се предсказва и обяснява посредством ментални свойства като:
  - > Вярa
  - > Очакване
  - > Надежда
  - > Опасение и т.н.
- » В народо-психологията се наричат съзнателни понятия



# Видове разсъждения

- » В много случаи е очевидно, че за да решим какво да правим, не използваме чисто логически разсъждения
- » Този модел е подходящ при разработване на рационално опериращи агенти
- » Съществено е да различаваме двата вида разсъждения:
  - > Теоретични разсъждения – направлявани от „вярата“ (beliefs)
  - > Практически разсъждения – директно насочени към действия



# Рационалност

» Рационалността зависи от различни неща, като напр. следните:

- > Оценка на постижението, която дефинира критерия за успех
- > Предварителни знания за околната среда на агента
- > Действията, които може да извърши агента
- > Последователност на досегашните възприятия на агента

» Дефиниция за рационален агент:

- > За всяка възможна последователност от възприятия избере действие, от което се очаква да максимира оценката за постижение, при което се вземат предвид неговата последователност от възприятия и предварителните му знания



# Практически разсъждения

## » Практически модел за вземане на решения

- > Разсъждения, насочени към действия
- > Процес на изчисляване какво да се прави
- > Въпрос на претегляне на конфликтни съображения за и против конкуриращи се опции, където уместните съображения са представени като желания, оценки, грижа и вяра на агентите

## » Едно практическо разсъждение се състои от най-малко две различни дейности:

- > Обмисляне (*deliberation*) – решаване какво състояние на проблема искаме да постигнем
- > Планиране (*means-ends reasoning*) – решаване как искаме да това състояние





# Практическите разсъждения като процес

» Практическите разсъждения са изчислителни процеси

- > Т.е., в агентите обмислянето и планирането се извършват с ограничени ресурси (оперативна памет, процесор, време, ...)





# Ограничена рационалност

- » За ограничените ресурси можем да направим два съществени извода:
  - > Изчислението е полезен ресурс за агенти, разположени в среди с реално време
    - + Способността за добро опериране ще се определя (най-малко частично) чрез способността за ефективно използване на наличните изчислителни ресурси
    - + Т.е., един агент трябва да контролира разсъжденията си ефективно, ако иска да оперира добре .
  - > Агентите не могат да обмислят безкрайно
    - + Те трябва да прекратят обмислянето в някакъв момент, избирайки някакво състояние на проблема и ангажирайки се с постигане на това състояние
    - + Може да бъде продължено обмислянето за фиксиране на друго състояние



# Тест: неочаквана среща



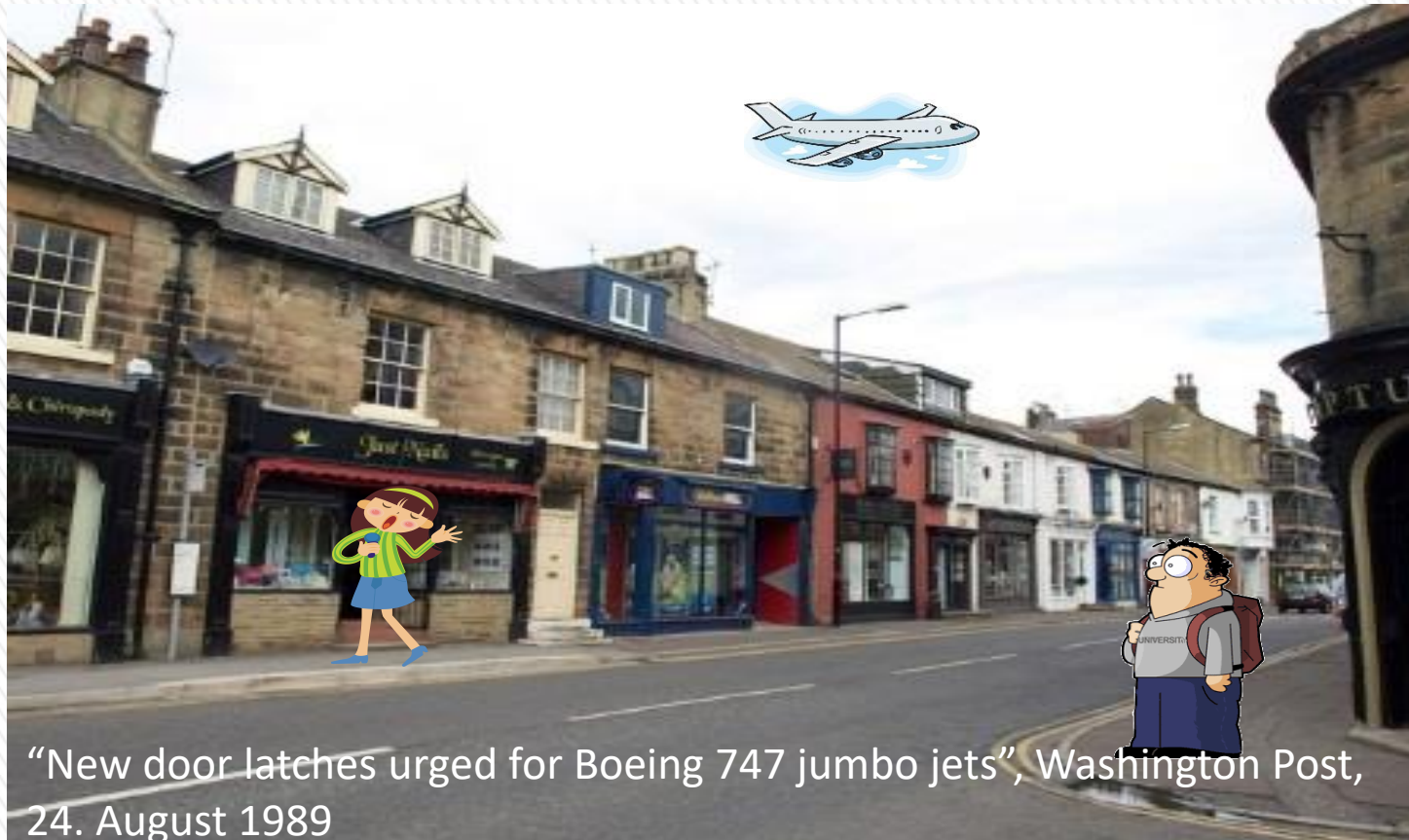
Рационално действие ли е пресичане на улицата за неочаквана среща с добра приятелка?



# Тест: неочаквана среща

?

„Всезнайност (перфектност)“  $\equiv$  „Рационалност“?



“New door latches urged for Boeing 747 jumbo jets”, Washington Post, 24. August 1989





# Тест: неочаквана среща

? Действа ли агентът „рационално“?

Избор на действие „оглеждане“  
преди „пресичане“ води до  
оптимално постижение



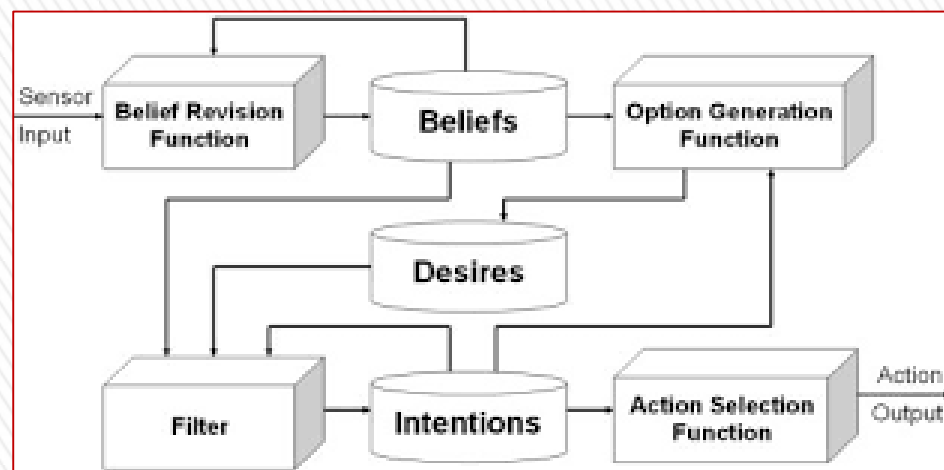
# BDI архитектура

» На рационален агент

- > B: Beliefs
- > D: Desires
- > I: Intentions



Michael Georgeff



# PRS

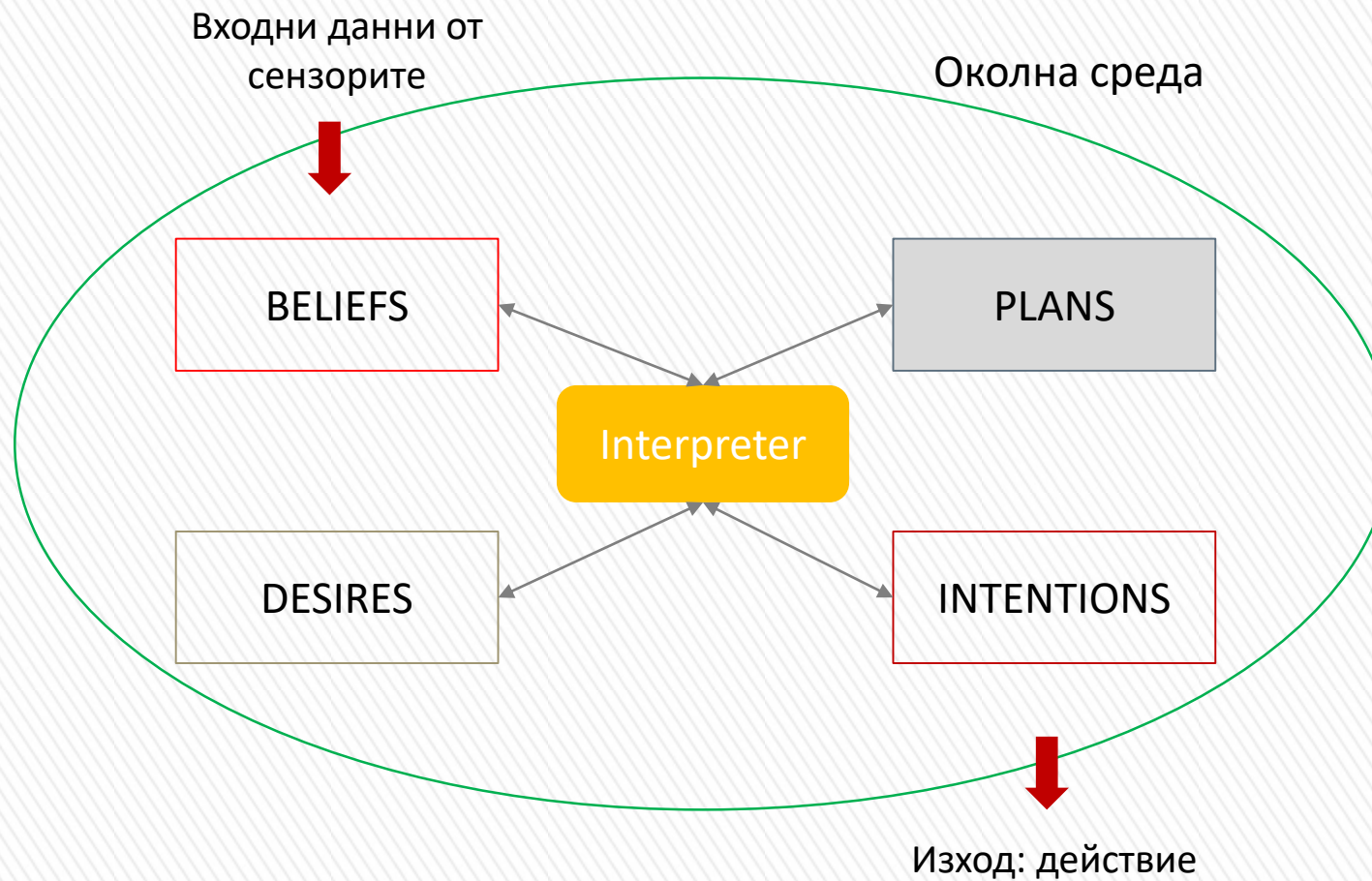
PRS (Procedural Reasoning System): M. Georgeff, A. Lansky – първоначално разработена в Stanford Research Institute

- Първата агента архитектура, реализираща BDI парадигмата – до днес най-трайната архитектура
- Вградена в няколко от най-съществените мултиагентни приложения:
  - ✓ OASIS – контрол на въздушния трафик
  - ✓ SWARMM – симулационна система за военно въздушните сили на Австралия
  - ✓ SPOC – управление на бизнес-процеси
- Явно представяне на структури данни, кореспондиращи с менталните състояния





# PRS архитектура





Благодаря за вниманието!

