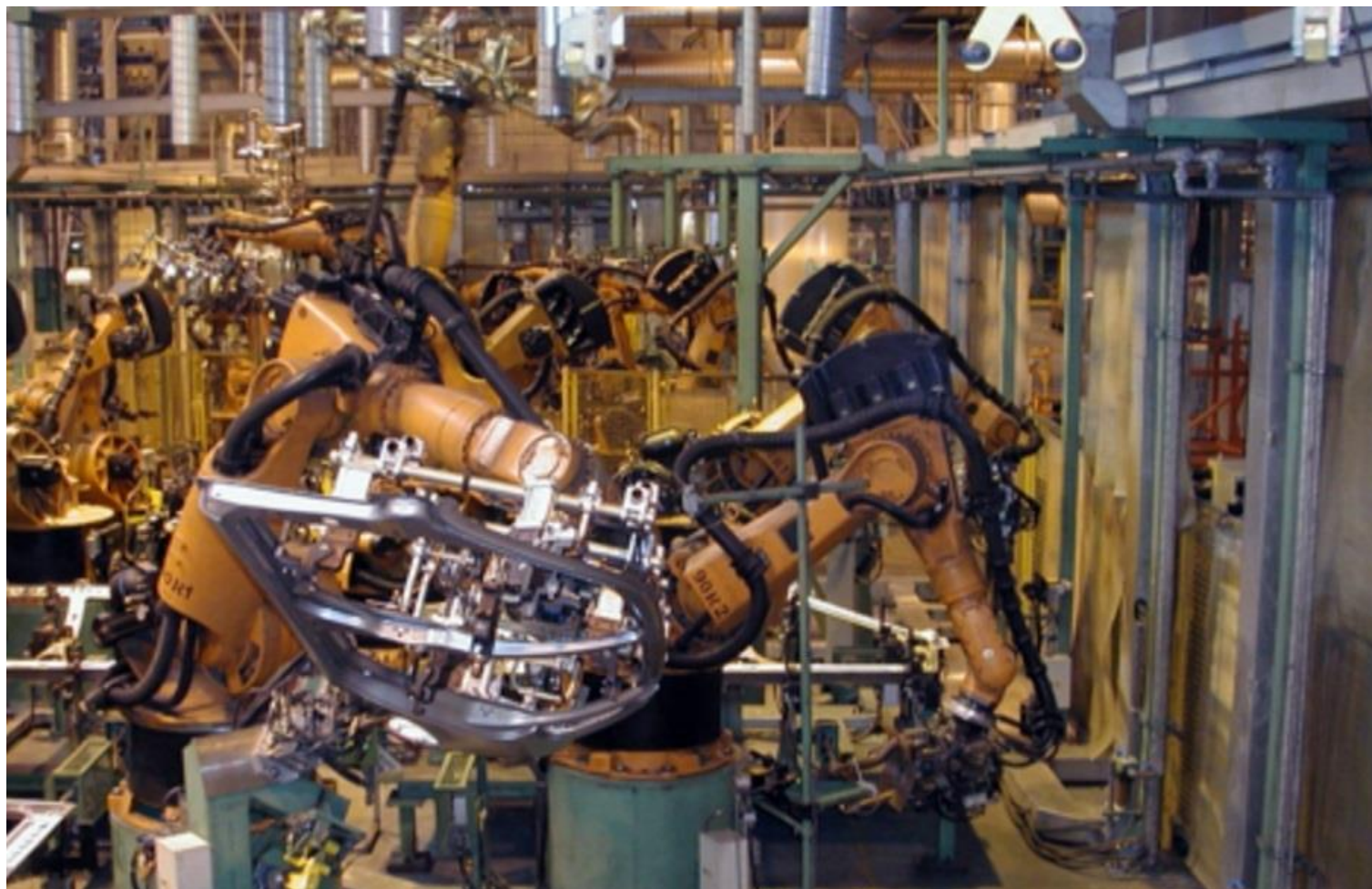




*Thinking is not the ability to manipulate **language**;  
it's the ability to manipulate **concepts**. Computer  
science **should be about concepts**, not languages.  
Лесли Лэмпорт*

## Семинар по специальности на английском языке

# ЛЕКЦИЯ 1: МОТИВАЦИОННАЯ



Роботы – или киберфизические системы становятся частью промышленной инфраструктуры



Технологии обещают много, ... но .....  
облик будущего зависит только от нас -

**СОЗИДАТЕЛЕЙ НОВЫХ  
ТЕХНОЛОГИЙ**



Сотворение  
Адама  
Микеланджело  
Сикстинская капелла

- Нам предоставлена уникальная возможность быть не только современниками и свидетелями этого процесса, но и непосредственными ТВОРЦАМИ новой «искусственной» жизни, которая будет сосуществовать и ЭВОЛЮЦИОНИРОВАТЬ вместе с людьми на Земле и в космосе

- Представление состояний таких процесс в пространстве  $L^3 \times T^3$  время «прошлое-настоящее-будущее». Основное отличие машин Тьюринга от когнитивных систем в организации системы памяти – от памяти с прямой адресацией команд и данных к памяти «задачи» - task memory
- Прогноз «будущего» влияет на «настоящее», чего нет в случае «физических процессов».
- В итоге и прошлое и будущее «случайно» , т.е. не определено до детерминированного причинного описания (Лагранжев принцип достаточного основания не работает)
- Мощность множество задач, имеющих решение и множество задач, имеющих алгоритм решения существенно различны



By Peter Singer (профессор биоэтики Принстонского университета).

PRINCETON, NEW JERSEY – Last month, AlphaGo, a computer program specially designed to play the game GO, caused shock waves among people when it defeated Lee Sidol, one of the world’s top-ranked professional players, winning a five-game tournament by a score of 4-1.

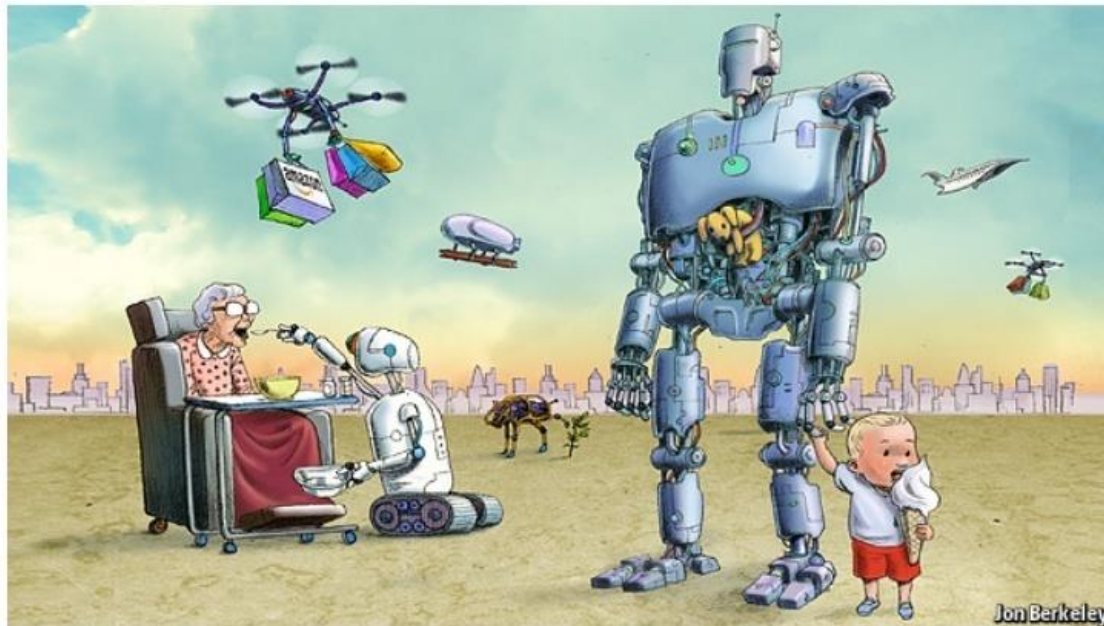
Why, you may ask, is that news? Twenty years have passed since the IBM computer Deep Blue defeated world chess champion Garry Kasparov, and we all know computers have improved since then. But Deep Blue won through raw computing power, using its ability to calculate the outcomes of more moves to a deeper level than even a world champion can. Go is played on a far larger board (19 by 19 squares, compared to 8 by 8 for chess) and has **more possible moves than there are atoms in the universe**, so raw computing power was unlikely to beat a human with a strong intuitive sense of the best moves.

Instead, AlphaGo was designed to win by playing a huge number of games against other programs and adopting the strategies that proved successful. You could say that AlphaGo evolved to be the best go player in the world, **achieving in only two years what natural selection took millions of years to accomplish.**”



## Rise of the robots

*Prepare for a robot invasion. It will change the way people think about technology*



- Стивен Хокинг:
  - "ИИ - возможный убийца человеческой цивилизации"
- Prof. Marek Perkowski. Dept of Electrical and Computer Engineering, PSU:
  - "Agent-based systems will dramatically change the world we live in".
- С. Рассел и П. Норвиг. Искусственный интеллект. //2002. Стр.1286:
  - “ ... системы ИИ могут создать угрозу самоопределению, свободе и даже выживанию людей. По этим причинам нельзя рассматривать исследования в области ИИ в отрыве от их этических последствий.”







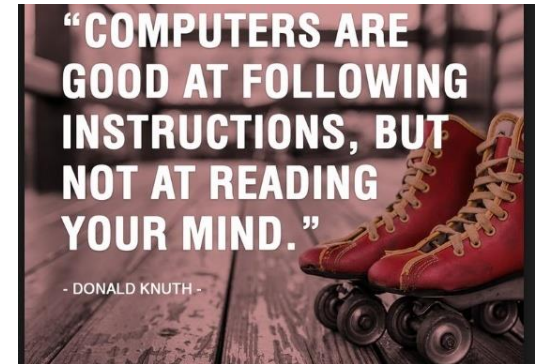
## Проблема в том, что

системы, которые мы строим, ведут себя в точности так, как мы их запрограммировали какие данные обрабатывают, используя разрешенный набор операций частично-рекурсивных функций

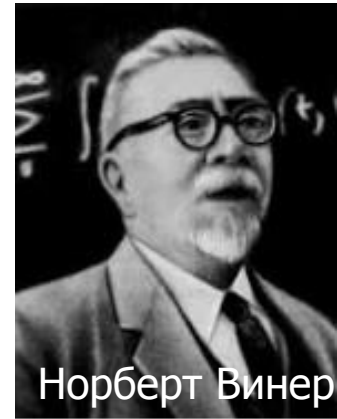
**НО**

**совсем не так, как мы предполагали!**

Человеческий мозг не в состоянии справиться со сложностью при разработке программ объемом в сотни миллионов строк кода



“Если магия вообще способна **даровать** что-либо, то она дарует именно то, что вы **попросили**, а не то, что вы подразумевали, но не сумели точно **сформулировать**.”

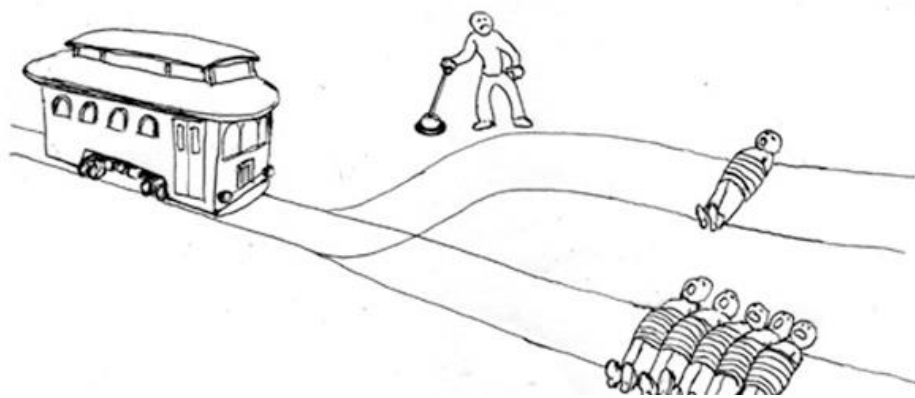


...исполнение заданного  
осуществляется вычислительными  
машинами в высшей степени буквально ...так это написано  
в программе по правилам формальной логики и  
компьютерной арифметики

*«Компьютеры похожи на ветхозаветных богов:  
сплошные правила и ни капли жалости».*

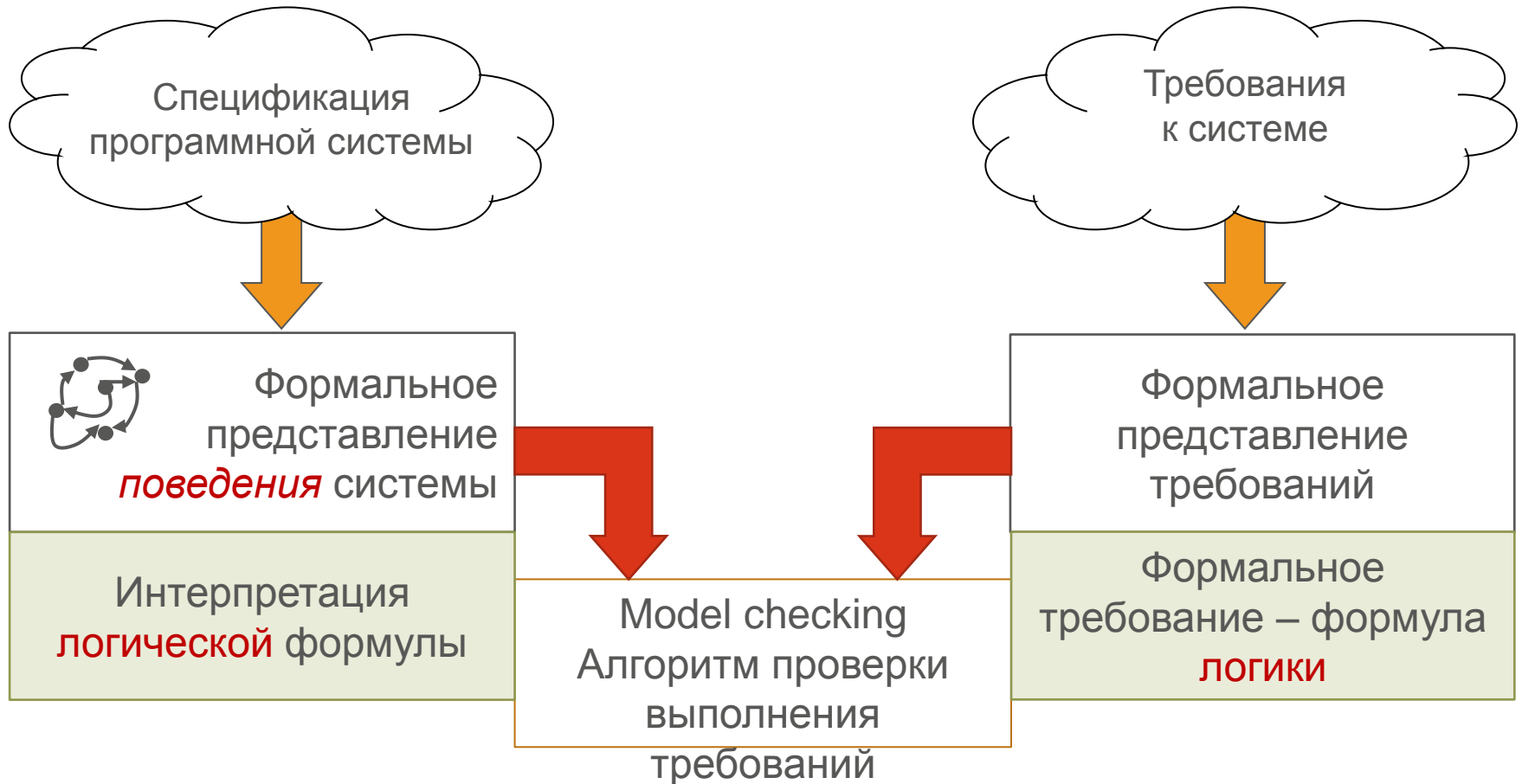
( Дж. Кэмпбелл)

К 2025 году в мире ежегодно будет продаваться 15 млн автомобилей с автопилотами, которые будут обходиться без водителя

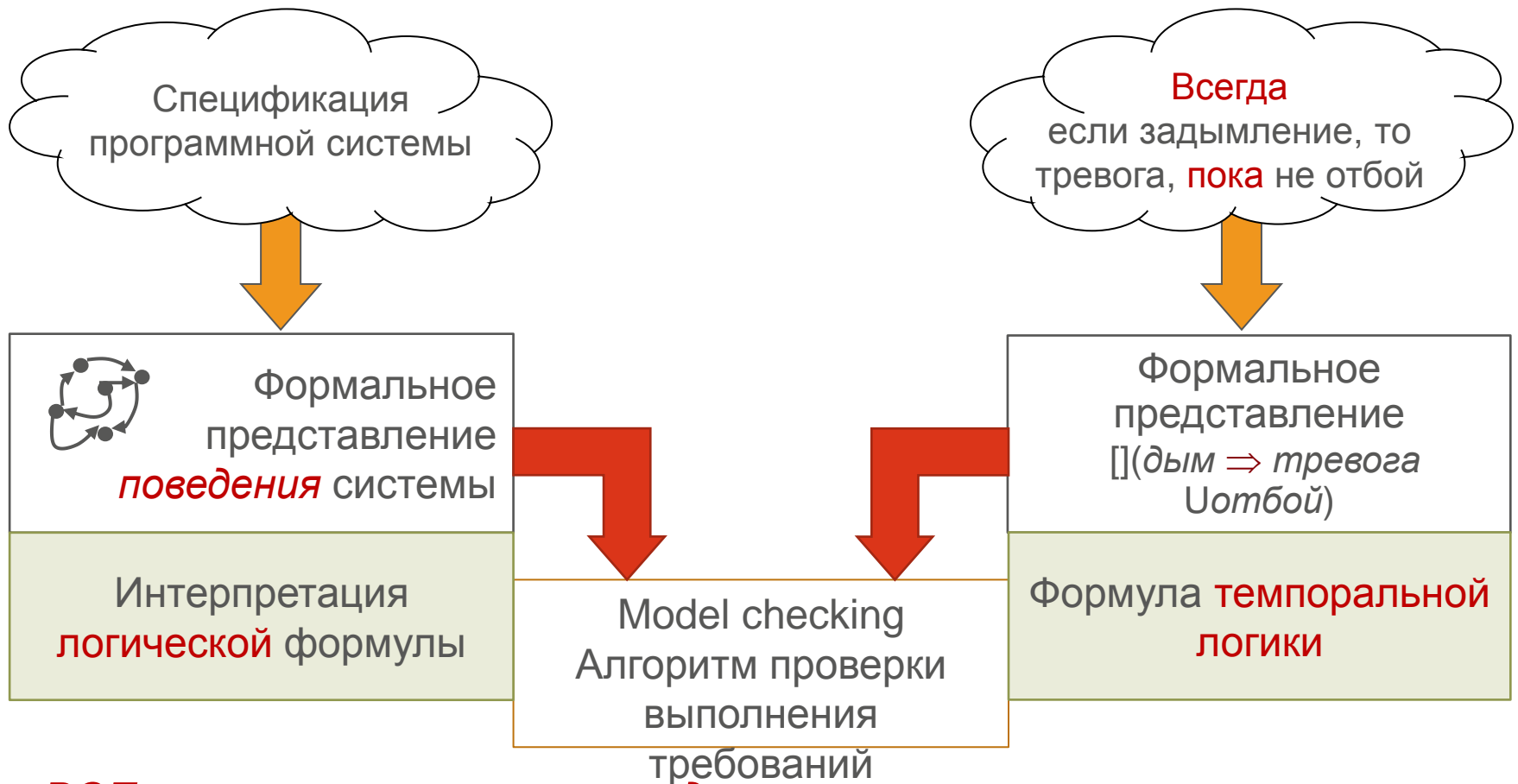


Какую альтернативу выберет автопилот, **если одна из них неизбежна**: задавит одного ребенка или пять мужиков?

# КАК МОЖНО ГАРАНТИРОВАТЬ, ЧТО ТРЕБОВАНИЯ БУДУТ ВЫПОЛНЯТЬСЯ ПРИ ПОВЕДЕНИИ ПРОГРАММНЫХ АГЕНТОВ?

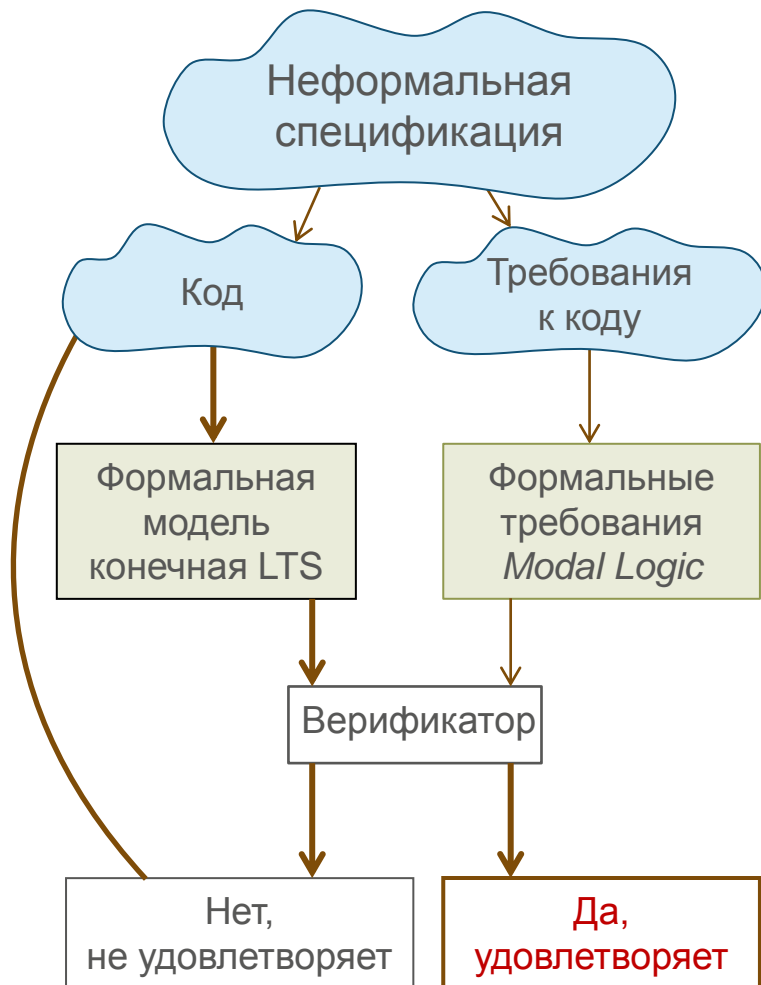


- **Model checking** – алгоритм проверки того, что логическая формула на данной интерпретации выполняется - принимает значение “истина”



*ВСЕ алгоритмы управления должны проверяться относительно их БУДУЩЕГО поведения во времени*

## Классический подход



## Синтез супервизоров





