# Познающий агент — борец с неопределенностью

Т.Гергей
(Applied Logic Laboratory, Budapest)
О.Аншаков
(РГГУ, Москва)

# Когнитивная система

Когнитивная система – это система, которая способна познавать свое окружение и влиять на него.

Познание сопровождается извлечением новой информации и знаний из экспериментальных данных и информации, возникающей в результате восприятия.

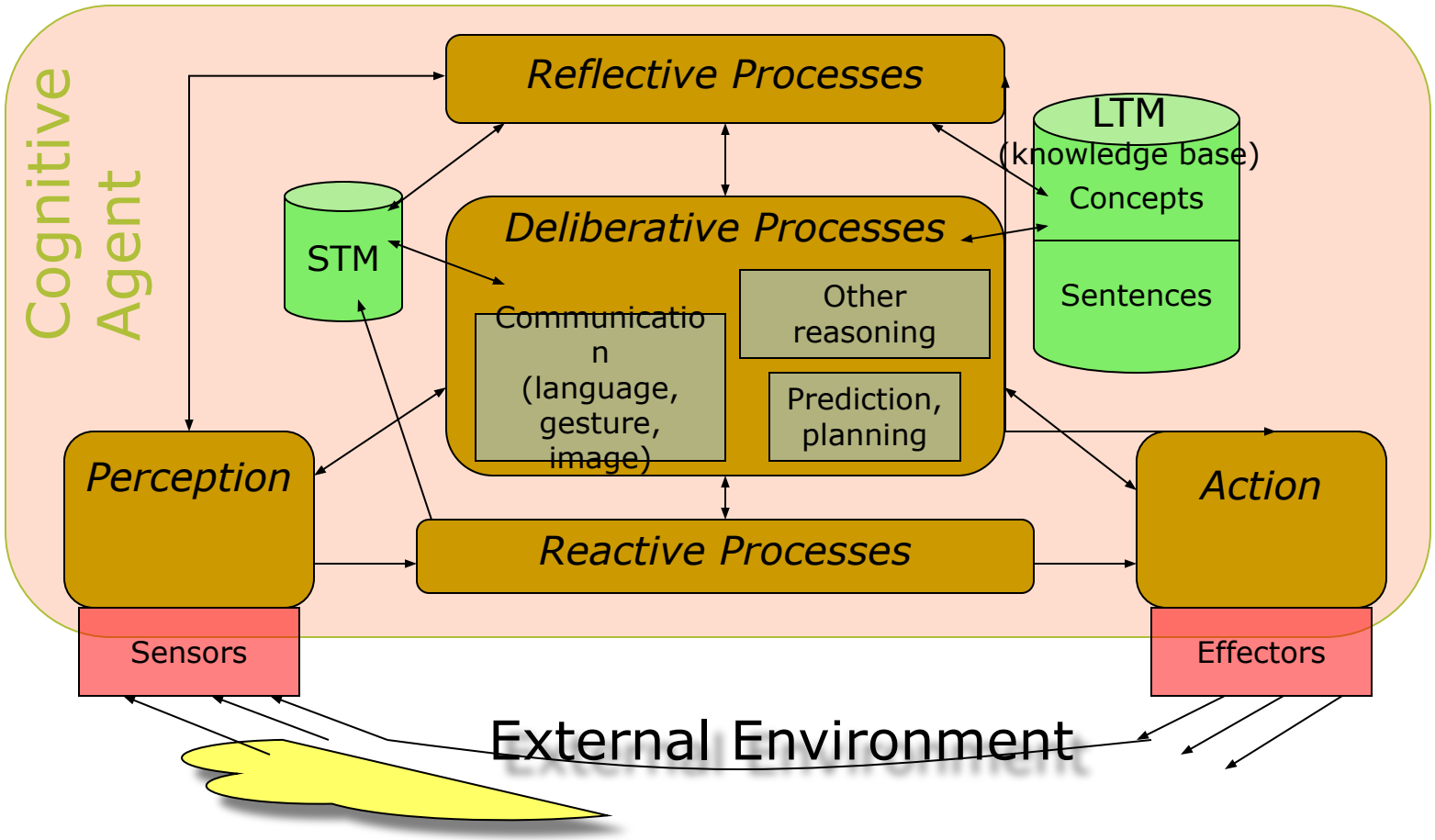
# Когнитивная система

Познавательная система – это такая система, которая:

- умеет рассуждать, используя достаточный объем адекватно представленных знаний;
- умеет обучаться на своем собственном опыте, что позволяет ей выполнять свои функции завтра лучше, чем сегодня;
- умеет объяснить саму себя и понимать сообщения о том, что ей необходимо сделать;
- знает о своих возможностях и умеет оценивать свое поведение;
- является устойчивой к неожиданным событиям.

# Possible Cognitive System Anatomy

Ron Brachman



# Познающий агент

Познающий агент является главным актером в нашем подходе.

Основной целью функционирования познающего агента является адаптация к постоянно меняющимся условиям окружающей среды. Чем более адекватной и гибкой окажется модель мира агента, тем более правильно он будет отвечать на воздействия среды, тем более он будет адаптирован к ней.

Определяющую роль в модели мира будут играть (каким-либо образом представленные) знания.

# Познание 1

- Термин "познание" рассматривается нами в наиболее общем смысле, а именно, главные особенности познавательных процессов определяются независимо от *субъекта* (человека, животного или машины), реализующего познание.

- Познание имеет отношение к получению информации, обнаружению закономерностей, формированию знаний, созданию модели окружающей среды.

# Познание 2

- Познание в наиболее общем смысле – это уменьшение неопределенности (или энтропии — с точки зрения теории информации).
- Одной из задач познания является обеспечение адаптации субъекта познания к меняющимся условиям окружающей среды.
- Познание позволяет использовать прошлое для предсказания будущего.

# Познание 3

- Термин "познание" имеет дуальный характер. Он обозначает как интеллектуальный (динамический) процесс, так  (статический) результат такого процесса.
- В динамическом аспекте этот процесс представляет собой движение от незнания к знанию.
- В статическом аспекте мы имеем построенную или улучшенную модель и порожденное или усовершенствованное знание об окружающей среде.

# Познавательные рассуждения

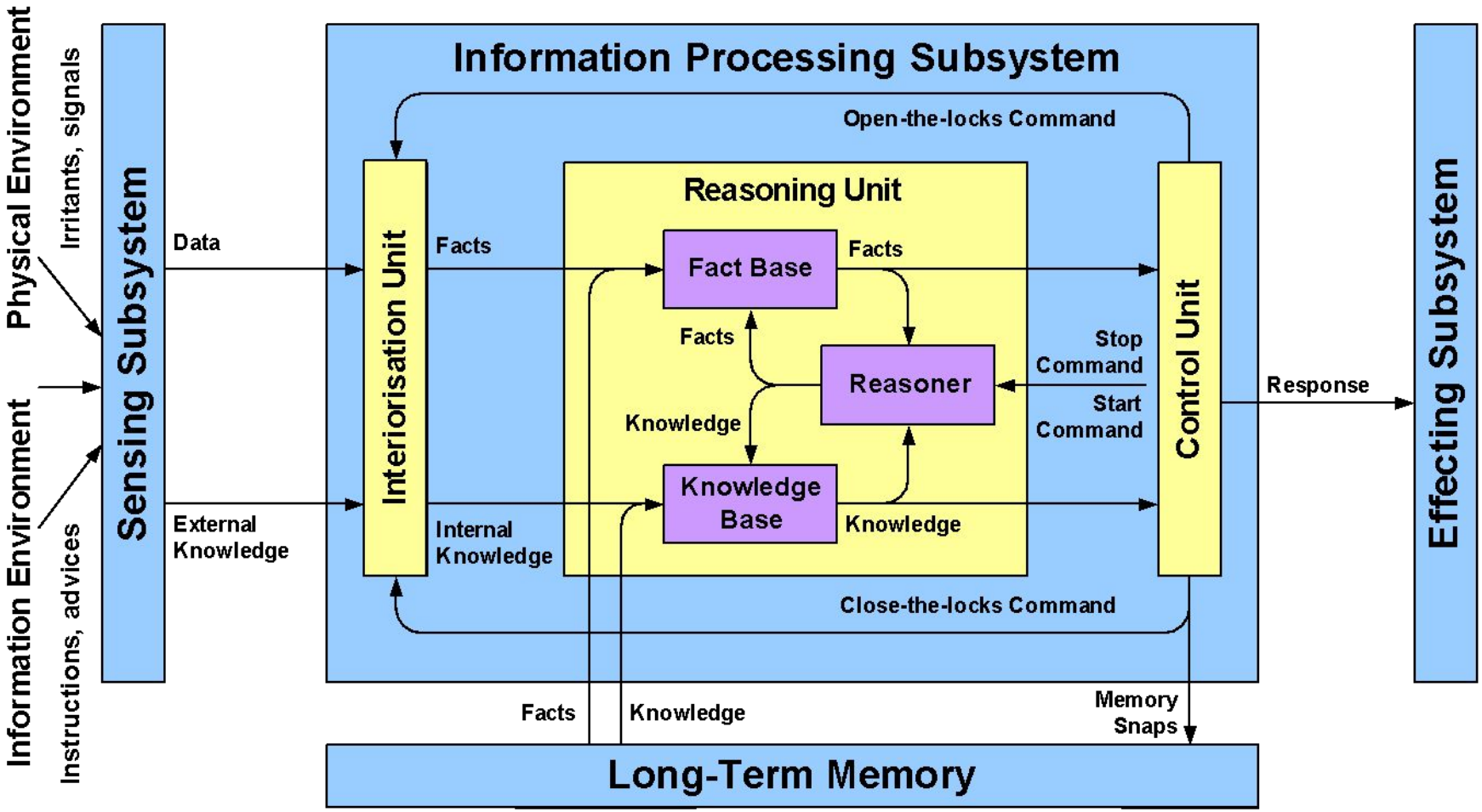
- Познавательные рассуждения — движущая сила познавательных процессов, выполняемых познающими агентами.
- Познавательные рассуждения — это обычная форма рассуждений, имеющих место в случае искаженной и/или неполной информации в ситуациях с неопределенности.

# Особенности познавательных рассуждений

1. Процесс рассуждений имеет динамическую природу. У этой динамической природы есть два аспекта: информационный и репрезентационный .
    - Информационный аспект понимается как способность к порождению существенно новой информации.
    - Репрезентационный аспект означает возможность вносить изменения в историю рассуждений.
2. Процесс рассуждений является референциальным, то есть он предполагает содержательный (семантический) характер, связанный со смыслом, внутренним представлением объектов и их структур.
3. Процесс рассуждения не является строго определенным, в том смысле, что он остается процессом рассуждений, даже тогда, когда кажется некорректным.

# Предлагаемый подход

# Концептуальный каркас КР
## Познающий агент



# Основные ограничения модели познающего агента

- для извлечения знаний познающий агент использует *рассуждения* в широком смысле этого слова, которые могут включать дедуктивный вывод, правдоподобное аргументирование и вычислительные процедуры,
- деятельность познающего агента *дискретна*, она состоит из чередующихся фаз восприятия окружающей среды и рассуждений, направленных на формирование новых знаний;
- в процессе функционирования познающего агента формируется *история* его познавательной деятельности.

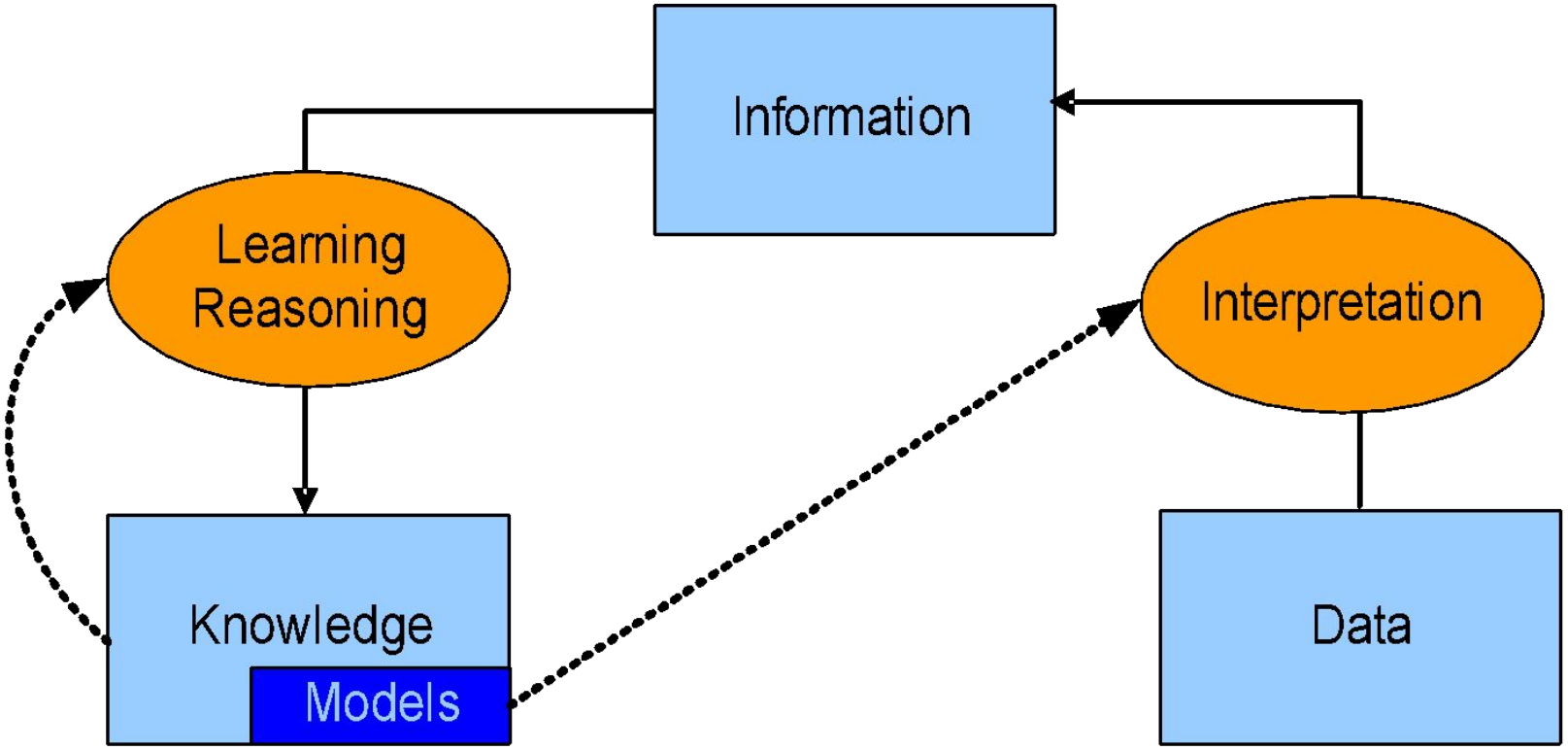
# От данных к знаниям

познающий агент обеспечивает протекание процесса

**данные** → **информация** → **знание**,

обеспечивая процесс решения задач и принятия решения

# От данных к знаниям



# Источники данных

Для процесса извлечения информации и знаний необходимы подходящие исходные данные, которые могут быть получены из следующих источников:

- из результатов наблюдений и экспериментов,
- из имеющихся теоретических знаний,
- из внешних источников информации и знаний, и
- из модельных экспериментов.

# Деятельность познающего агента

Деятельность познающего агента можно представить в виде нескольких вложенных циклов. Прежде всего мы видим два вложенных цикла.

- Внутренний цикл состоит в повторении выполнения алгоритма рассуждений до тех пор, пока не будет выполнено стоп-условие.
- Внешний цикл представляет собой повторение фазы восприятия и рассуждения  и продолжается от начала деятельности познающего агента, до его завершения.

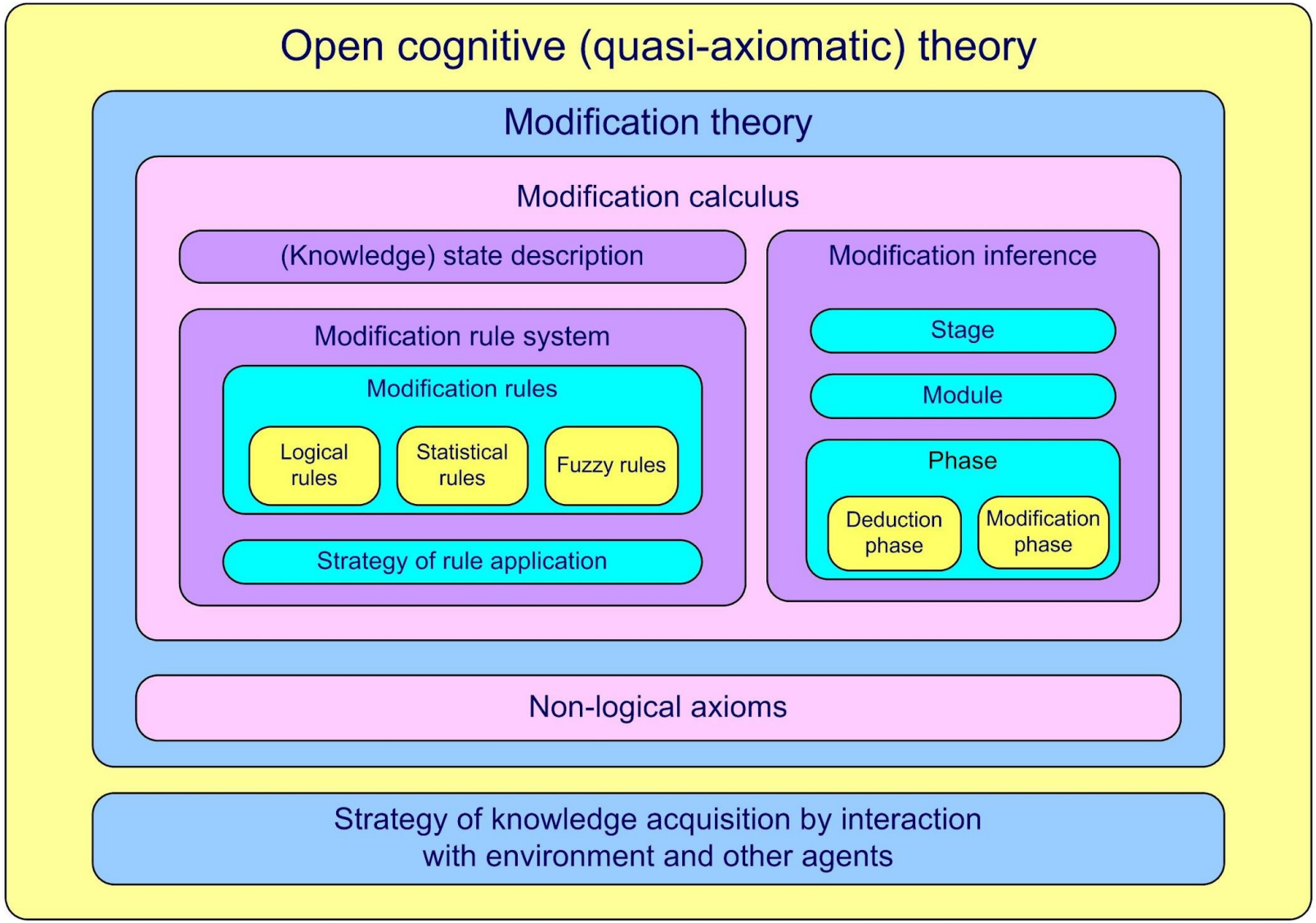
# Внутренний цикл

Во внутреннем цикле подразумеваем достаточно простую процедуру *обхода* некоторой выделенной части рабочей памяти, включающей часть базы фактов и часть базы знаний. Этот обход сопровождается анализом фактов, извлечением знаний и преобразованием (модификацией, трансформацией) базы фактов и базы знаний.

Этот цикл моделируем при помощи *модификационной теории.*

Внешний цикл моделируем при помощи *открытой теории*.

# Формальный каркас КР



# Модификационные теории 1

Теории называют модификационными, потому что те формальные конструкции в этих теориях, которые моделируют процесс рассуждения, позволяют не просто добавлять новые утверждения, но и изменять саму такую конструкцию. Модификационные теории используют специальные методы вывода для того, чтобы построить модель процесса рассуждения. Модификационные теории позволяют представлять динамику познавательных процессов, понимаемых как движение от незнания (неопределенности) к знанию (определенности). Кроме того, модификационные теории умеют обращаться с противоречиями.

# Модификационные теории 2

Модификационные теории позволяют представить динамику когнитивного процесса, понимаемого как движения от незнанию (неопределенности) к знанию (определенности).

Модификационные правила изменяют не только вывод, но и истинностную оценку некоторых утверждений, они заменяют истинностную оценку «неопределено» на одну из определенных истинностных оценок, например на «истинно» или «ложно», но таких оценок может быть сколь угодно много, они будут выражать степень доверия к оцениваемому утверждению, степень его истинности или ложности.

Модификационные теории могут работать с противоречиями.

# Модификационные теории 3

Аксиомы модификационных теорий являются аналогами фактов и знаний, следовательно в модификационных теориях представлено содержимое компонентов блока рассуждений познающего агента таких как база фактов и база знаний.

Модификационные правила в теории организованы в систему, которая является аналогом алгоритма рассуждений познающего агента. Эта система управляет построением модификационного вывода. В частности модификационный вывод разбит на стадии, каждая стадия является аналогом одной итерации выполнения *обхода* базы фактов и базы знаний.

# Вывод в модификационном исчислении

Вывод в модификационном исчислении можно рассматривать как образ истории рассуждений познающего агента. Причем таких рассуждений, которые не прерываются фазой восприятия, а выполняются целиком внутри фазы рассуждений. Мы говорим в этом случае, что рассуждения не выходят за пределы одного цикла работы рассуждателя. Заметим, что этот цикл, вообще говоря, состоит из нескольких итераций обхода базы фактов и базы знаний. Этот обход сопровождается анализом фактов и извлечением знаний.

*Модификационные теории представляют структуру и функции блока рассуждений.*

# Модели модификационной теории

В качестве модели модификационной теории рассматривается последовательность структур. Каждая такая структура с одной стороны отражает некоторое состояние знаний о рассматриваемом фрагменте предметной области, с другой стороны является образом рабочей памяти (объединения базы фактов и базы знаний) познающего агента для некоторой стадии рассуждений.

# Открытая когнитивная теория

A theory is called *open* because it is open for the representation of the interaction between the cognizing agent and its environment. It should be open for (i) obtaining new facts for analysis, (ii) adopting and internalising knowledge from other similar theories and, (iii) modification of the reasoning rules.

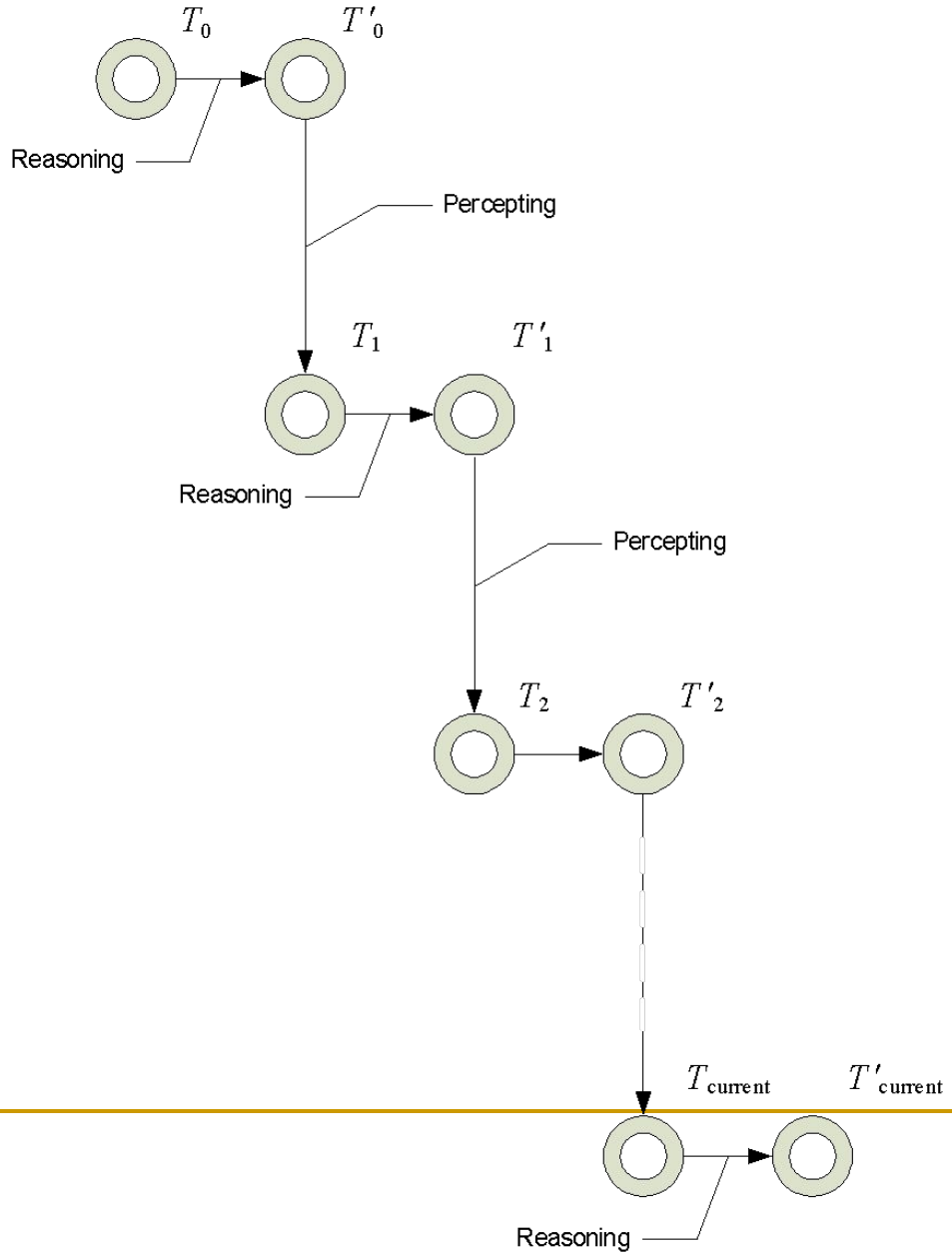
An open cognitive theory is used to represent the activity of a cognizing agent as a history. Thus an open cognitive theory is an essential component of self-reflection of an agent.

# Открытая когнитивная теория

Открытая когнитивная теория определена как последовательность пар модификационных теорий $\langle T_i, T_i' \rangle$, где теория $T_i$ соответствует состоянию знаний познающего агента до начала некоторой фазы рассуждений, а $T_i'$ соответствует состоянию знаний после завершения этой фазы рассуждений.

У теорий $T_i$ и $T_i'$ имеется некоторый общий набор знаний, которые не могут изменяться в процессе рассуждений, например в процессе рассуждений не изменяется алгоритм рассуждений, не изменяются некоторые аксиомы, описывающие модель внешнего мира.

# Структура открытой когнитивной теории



# Преимущества предлагаемого подхода

- Он позволяет совместно использовать достоверные и различные виды правдоподобных правил вывода.
- Он позволяет корректно работать с неопределенностью и временными противоречиями в процессе рассуждения.
- Он позволяет — для анализа данных и извлечения информации — использовать качественные методы, основанные на формализованных рассуждениях, вместо количественных методов, характерных для традиционного подхода.
- Если это требуется, он также позволяет совместно использовать количественные и качественные методы.
- Открытые когнитивные теории внутри каркаса КР позволяют оперировать внутренними и внешними знаниями и обеспечивать расширение базы знаний и саморефлексию.

# Преимущества предлагаемого подхода

Каркас КР определяет методы для

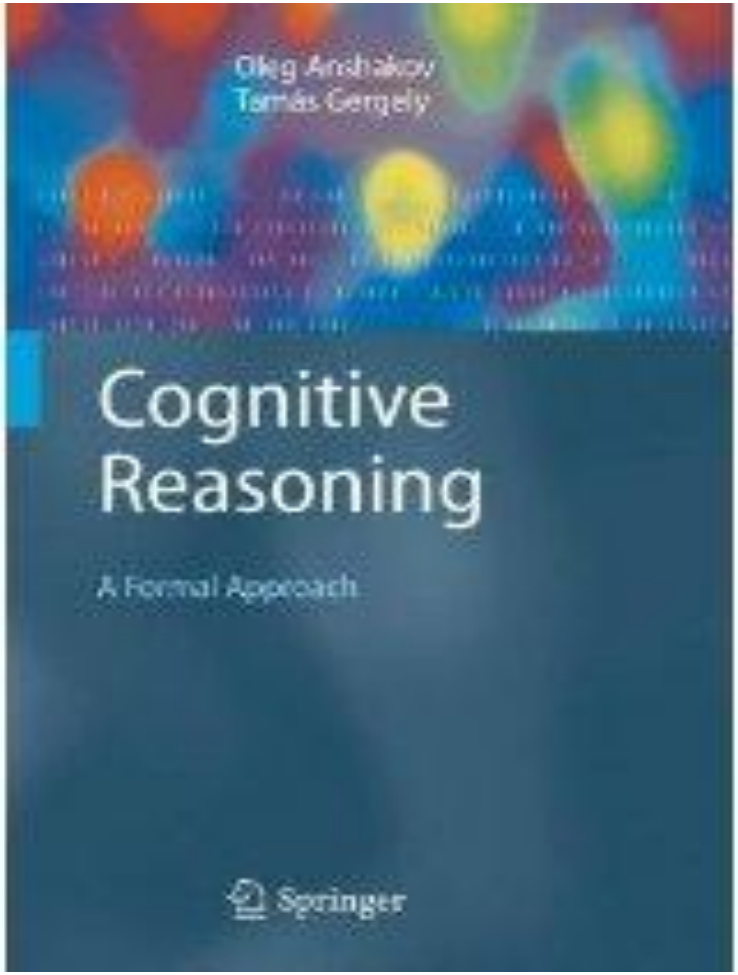
- обнаружения сходства и различия,
- классификации,
- определения причин.
- извлечения закономерностей
- нахождения свойств,
- предсказаний.

# Заключение

Предлагаемый подход имеет существенные отличия от основанных на логике подходов в когнитивных науках, например, рассмотренных в книге **[Stenning, K. and van Lambalgen, M. Human Reasoning and Cognitive Science, MIT Press, 2008]**, а именно:

- Он включает логический формализм, который специально ориентирован на описание познавательных процессов.
- Он предоставляет средства не только для описания познавательных процессов, но и для создания искусственного познающего агента.

# Ссылка




**O. Anshakov and T.Gergely,**
Cognitive Reasoning: A Formal Approach, Springer, 2010

# Открытая проблема

В нашей модели, познающий агент анализирует имеющиеся данные *с помощью рассуждений* и извлекает в результате этого анализа имплицитные закономерности. Во многих случаях для анализа данных используются различные методы машинного обучения. Применение методов (алгоритмов) машинного обучения является вычислительной процедурой, которая не похожа на рассуждения. Мы предполагаем, что по-крайней мере некоторые широко-известные методы машинного обучения (например, деревья решений, наивный Байесовским метод) можно интерпретировать в терминах рассуждений. Построение формализма в стиле данной книги для известных методов машинного обучения является открытой проблемой. Мы считаем, что решение этой проблемы явилось бы вашным шагом на пути унификации общей схемы функционирования познающего агента.