

Основные понятия и история искусственного интеллекта. Принципы построения систем поддержки принятия врачебных решений.

Лекция 1

Киселёв Глеб Андреевич
к.т.н., старший преподаватель
ФФМиЕН РУДН

тел.: +79067993329
email: kiselev@isa.ru



- **Что такое искусственный интеллект (ИИ)?**

Это программа, выполняющая интеллектуальные функции. Функции считаются интеллектуальными, когда не существует заранее заданного алгоритма решения проблемы.

- **Каким бывает ИИ?**

ИИ бывает сильным и слабым. Сильный искусственный интеллект – недостижимая цель, при которой машина обладает всеми интеллектуальными способностями человека. Слабый искусственный интеллект – это группа программно-аппаратных, либо программных комплексов, автоматизирующих интеллектуальную задачу.

- **Примеры создания сильного ИИ:**

Робототехнические устройства, управляемые комплексными когнитивными системами: LIDA, SOAR, ICARUS, STRL, CLARION и т.д.

- **Примеры создания слабого ИИ:**

1. Чат боты - Для управления устройством/универсальные - Sap conversational AI, Alexa от Amazon, Assistant от Google; Cortana от Microsoft; Siri от Apple; Алиса от Яндекс; Маруся от Mail.ru Group;
2. Для взаимодействия с бизнес-приложениями- ассистент «Олег» в приложении банка «Тинькофф»; голосовой помощник от Сбербанка; голосовой помощник ASICS – подбор спортивного снаряжения;
3. Предметно-ориентированные сервисы – умные навигаторы (Citymapper, UGV Driver Assistant); умные сиделки (Mishka AI), фитнес тренеры, приложения для оптимизации расписания;
4. Системы распознавания аудио, видео сигналов и т.д.

	STRL	LIDA	SOAR	CLARION	ACT-R	ICARUS	SemaFORR	CAMAL	HiPOP	CARINA
Структура	3 уровня - стратегический, тактический и реактивный	3 фазы - 1. понимания 2. внимания 3. действий и обучения	Набор связанных вычислительных модулей. Акцент на функциональности и эффективности.	4 подсистемы ориентированные на - 1. ACS действия 2. NACS память 3. MS мотивацию 4. MCS внимание и осознание	Модульность. Каждый модуль имеет буфер, через который запрашивает информацию. В отличие от SOAR описывается более подробный когнитивный цикл	Линейная структура распознавания среды и поиска навыков в долговременной и кратковременной памяти	3 уровня эвристических советчиков - 1. по ситуации (реактивные), 2. по планированию и 3. по решению разногласий между 1 и 2	Представлена соединением схемы BDI нотации, аффективных модулей и рефлексивного контроля поведения.	Модули планирования (перепланирования), коммуникаций	4 уровня: уровень 0. когнитивного агента 1. рассуждений о его деятельности 2. мета-модель рассуждений 3. мета-мета модель
Способ представления знаний	(с) - знаковая модель (т) - метрические топологические графы (р) - нелинейные регуляторы	Зависит от реализации: кодлеты + сети схем, векторы многомерного модульного составного представления и др.	Декларативный - чанки (куски, вектора) и процедурный (продукции). Знания краткосрочные и долгосрочные.	Знания в каждой из подсистем представлены в явном и неявном виде. Явный - системы правил, неявный - нейросети	Декларативный - чанки (куски, вектора) и процедурный (продукции). Знания краткосрочные и долгосрочные.	Знания представлены в предикатной форме. Знания краткосрочные и долгосрочные.	Дескриптивы - описания среды для советчиков. Также присутствует 3 вида геометрических абстракций: - трассы, области и конвейеры	BDI графы доверия	В предикативной форме	Элементами внутреннего языка M++
Процесс планирования поведения и принятия решений	Иерархический, основан на принципах распространения активности	Иерархический, представлен быстро повторяющимися когнитивными циклами.	Иерархия процессов обработки подцелей. Процессы распараллелены.	Иерархичность заключается в разделении когнитивных действий по каждой из подсистем	Иерархичность процесса заключается в сложности работы модулей.	Иерархичность в поиске и приобретении навыков, а также в распознавании и концептуализации ситуации	Иерархический, основан на цикле опроса советчиков	Стандартная иерархичность BDI-рассуждений, дополненная расширенными вероятностными оценками доверия	Иерархическая версия алгоритма частичного планирования POP	Иерархичность заключается во множественном оценивании реакции окружающей среды на текущее действие
Обучение	Опыт планирования сохраняется и переиспользуется. Абстрактные действия планировщика уточняются с помощью обучения без учителя.	Опыт сохраняется и переиспользуется. Образуются новые когнитивные связи.	Обучение на опыте достижения подцели. Декларативная и процедурная память пополняема.	индуктивное и дедуктивное.	И декларативная, и процедурная память пополняемы	Создание новых навыков после достижения цели	Опыт планирования используется в эвристиках советчиков	Агент запоминает опыт планирования и взаимодействий	явно не представлен	Явно не представлено. Авторы описывают обучение, как смену стратегий, основываясь на опыте.

Пополнение знаний на основе данных сенсоров	Производится на тактическом и реактивном уровнях. Используются SLAM по карте, распознавание объектов для манипулирования	Зависит от реализации фазы понимания. Расширяется с помощью кодлетов.	Представлена модулем обработки визуальных данных	описание отсутствует	Реализации модуля распознавания и визуализации	Модуль распознавания окружения и выдвижения концепта	Пополнение знаний с помощью HoloLens+SLA M+EKF	Представлено в аффективном виде	Используется в робототехнической постановке	Модуль восприятия датчиками и распознавания.
Неопределенность данных	Обусловлено представлением о возможностях манипулирования с объектами в процессе обучения	Существуют примеры байесовской интерпретации фазы понимания.	Правила по выбору операторов в процессе планирования генерируются с помощью RL	Обусловлено неявным представлением данных. Используются алгоритмы RL/нейросети	Используются допущения к представлению задачи	Неоднозначность распознавания среды.	Обусловлено стратегиями советчиков и их представлением об выполнимости перемещения.	Обусловлена неточностью распознавания среды и вероятностной оценкой возможности взаимодействия с другим агентом.	Явно не представлена.	Явно не представлена.
Персонифицированность знаний	Представлено на стратегическом уровне знаками "Я" и "Они". Используется рефлексия 1 (авторефлексия) и 2 рода.	Представлено в перцептивной памяти рефлексией 1 и 2 рода.	Описана субъективность декларативных и процедурных знаний агента	Достигается посредством метокогнитивного уровня осознания	Отсутствует в классической реализации	Описание отсутствует.	Отсутствует явное представление.	Представлена BDI графом отношений.	Отсутствует в классической реализации.	Знания персонифицированы с помощью само-модели на метауровне m1.
Моделирование внимания агента	На стратегическом уровне представлена иерархическим фокусом внимания агента.	Представлена соответствующей фазой когнитивного цикла	Представлена робототехническим и дополнениями к внутреннему представлению окружающей среды	достигается детальностью подсистемы MS и MCS	Представлена сегментацией событий для робототехнических реализаций	Описание отсутствует.	Представлена деятельностью советчиков 1 и 2 уровней.	Описание отсутствует.	Описание отсутствует.	Представлена различными стратегиями метауровней по выявлению важных вещей при распознавании ситуации.
Использование автоматических процедур	Деятельность тактического и реактивного уровня	Используются кодлеты. Зависит от реализации.	Явное представление отсутствует. Зависит от реализации.	Представлена неявным представлением действий	Различается в зависимости от реализации архитектуры	Для выполнения действий. Ограничены временными рамками.	Любое выполнение действий.	Обусловлено реализацией аффективного поведения. Если робот слабый используется система SCARAB.	Представлено действиями перемещений.	Присутствует в конкретных реализациях когнитивных агентов.

Способ коммуникации между агентами	Протокол коммуникации, аукцион планов	Явно не описан, зависит от реализации.	Используются модули коммуникации, основанные на рефлексивном восприятии других агентов	Часть системы ACS. Явно не описан.	Представлен модулем коммуникации	Явно не описан. Зависит от робототехнической реализации.	Реактивная коммуникация. Для отслеживания деятельности группы (толпы) используются алгоритмы CUSUM-A* и Risk-A*.	С помощью протокола коммуникации и обмена данных.	Протоколы коммуникаций (обычный, восстановления плана, синхронизации...).	Явно не представлен.
Применение	Робототехника, интеллектуальные ассистенты	Робототехника, симуляция когнитивных агентов, медицинская диагностика	Робототехника, моделирование человеческого поведения для тестирования	Симуляции когнитивных агентов	Моделирование человеческого поведения для тестирования	Игровая индустрия, робототехника	Робототехника, игровая индустрия.	Робототехника.	Робототехника.	Интеллектуальные ассистенты.
Разработчик	ФИЦ ИУ РАН, Россия	Университет Мемфиса, США	Университет Корнеги Мелон, Питсбург, США	Ренсселерский политехнический институт, США	Университет Корнеги Мелон, Питсбург, США	NASA, DARPA, США	Городской Университет Нью-Йорка, США	Университет Халла, Англия	Университет Тулузы, Франция	Университет Сан Пауло. Бразилия
Ссылка на описание	Макаров Д. А., Панов А. И., Яковлев К. С. STRL: многоуровневая система управления интеллектуальными агентами // В кн.: Пятнадцатая национальная конференция по искусственному интеллекту с международным участием КИИ-2016 (3-7 октября 2016г., г.Смоленск, Россия): Труды конференции Т. 1. Смоленск : Универсум, 2016. С. 179-188.	Franklin, Stan et al. “THE LIDA ARCHITECTURE: ADDING NEW MODES OF LEARNING TO AN INTELLIGENT, AUTONOMOUS, SOFTWARE AGENT.” (2006).	Laird, J. E., Lebiere, C., & Rosenbloom, P. S. (2017). A standard model of the mind: Toward a common computational framework across artificial intelligence, cognitive science, neuroscience, and robotics. AI Magazine. https://doi.org/10.1609/aimag.v38i4.2744	Sun, R. 2016. Anatomy of the Mind: Exploring Psychological Mechanisms and Processes with the Clarion Cognitive Architecture. New York, NY: Oxford University Press	Ritter, FE, Tehranchi, F, Oury, JD. ACT-R: A cognitive architecture for modeling cognition. WIREs Cogn Sci. 2019; 10:e1488. https://doi.org/10.1002/wcs.1488	Choi, D., & Langley, P. (2018). Evolution of the ICARUS Cognitive Architecture. Cognitive Systems Research. https://doi.org/10.1016/j.cogsys.2017.05.005	Epstein, S. L., Aroor, A., Sklar, E. I., & Parsons, S. (2013). Navigation with Learned Spatial Affordances. 1–6. Retrieved from http://www.compsci.hunter.cuny.edu/~epstein/papers/CogSciFinal.Epstein.pdf	Davis, D. N., & Ramulu, S. K. (2017). Reasoning with BDI robots: From simulation to physical environment - Implementations and limitations. Paladyn. https://doi.org/10.1515/pjbr-2017-0003	Bechon, P., Lesire, C., & Barbier, M. (2020). Hybrid planning and distributed iterative repair for multi-robot missions with communication losses. Autonomous Robots. https://doi.org/10.1007/s10514-019-09869-w	Caro, M. F., Josvula, D. P., Gomez, A. A., & Kennedy, C. M. (2018). Introduction to the CARINA Metacognitive Architecture. Proceedings of 2018 IEEE 17th International Conference on Cognitive Informatics and Cognitive Computing, ICCI*CC 2018, (October), 530–540. https://doi.org/10.1109/ICCI-CC.2018.8482051
Ссылка на код и ссылка на архитектуру	http://strl-robotics.ru/ https://github.com/cog-isa/map-core	http://ccrg.cs.memphis.edu/tutorial/	https://soar.eecs.umd.edu/articles/downloads/soar-suite/227-soar-suite-9-6-0	http://www.clarioncognitivearchitecture.com/	http://act-r.psy.cmu.edu/	http://clic.stanford.edu/research/ongoing/icarus/	http://www.cs.hunter.cuny.edu/~epstein/	Описание отсутствует.	https://dblp.uni-trier.de/pers/b/Bechon/Patrick.html	https://www.researchgate.net/profile/Manuel_Pineres

- **Распространено использование виртуальных ассистентов на рабочем месте**

1. Системы автоматизированной обработки клиентских запросов. - RealPage®Contact Center 3.0, HelpDeskEddy, Okdesk
2. Системы управления рабочим процессом - Wrike, Forms On Fire, Evernote, Metronik.
3. Системы управления материальными ресурсами компании - Anixter, PAE, Udata, Logistix.
4. Планировщики расписания - Bitrix24, Timescamp, LeaderTask, Directum.

- **Использование виртуальных ассистентов в сфере образования**

1. Системы типа «личный кабинет». - Coursera, Khan Academy, Udemy, Academic Earth, edX, Code academy, Stepik. – автоматизация проверки и подачи знаний
2. Сервисы с элементами геймификации - Quizlet, Memorise, Duolingo, Socratic, SoloLearn. Примеры отечественных сервисов: Examer, Фоксфорд, LinguaLeo, Смотри.Учись
3. Рекомендательные системы – поиск работы, привычек, хобби на основе навыков. Pathsource, Ripple, Deepstash, Coach.me. Отечественных рекомендательных систем такого рода в свободном доступе не обнаружено.
4. Роботы – чат-боты: Зарешай, Miao academy и т.д.

- **Сравнить ассистентов можно по следующим параметрам:**

1. Является ли ассистент помощником широкого спектра или помогает решить узконаправленную задачу?
2. Производится ли учет контекста при работе с пользователем? (так называемые ассистенты 3 волны по классификации DARPA);
3. Является ли ассистент системой дополненного/усиленного интеллекта? (Augmented/Amplified Intelligence – оперирование другими системами и выстраивание инфополя владельца);
4. Существует ли экосистема ассистентов или различные сценарии работы родительского ассистента?

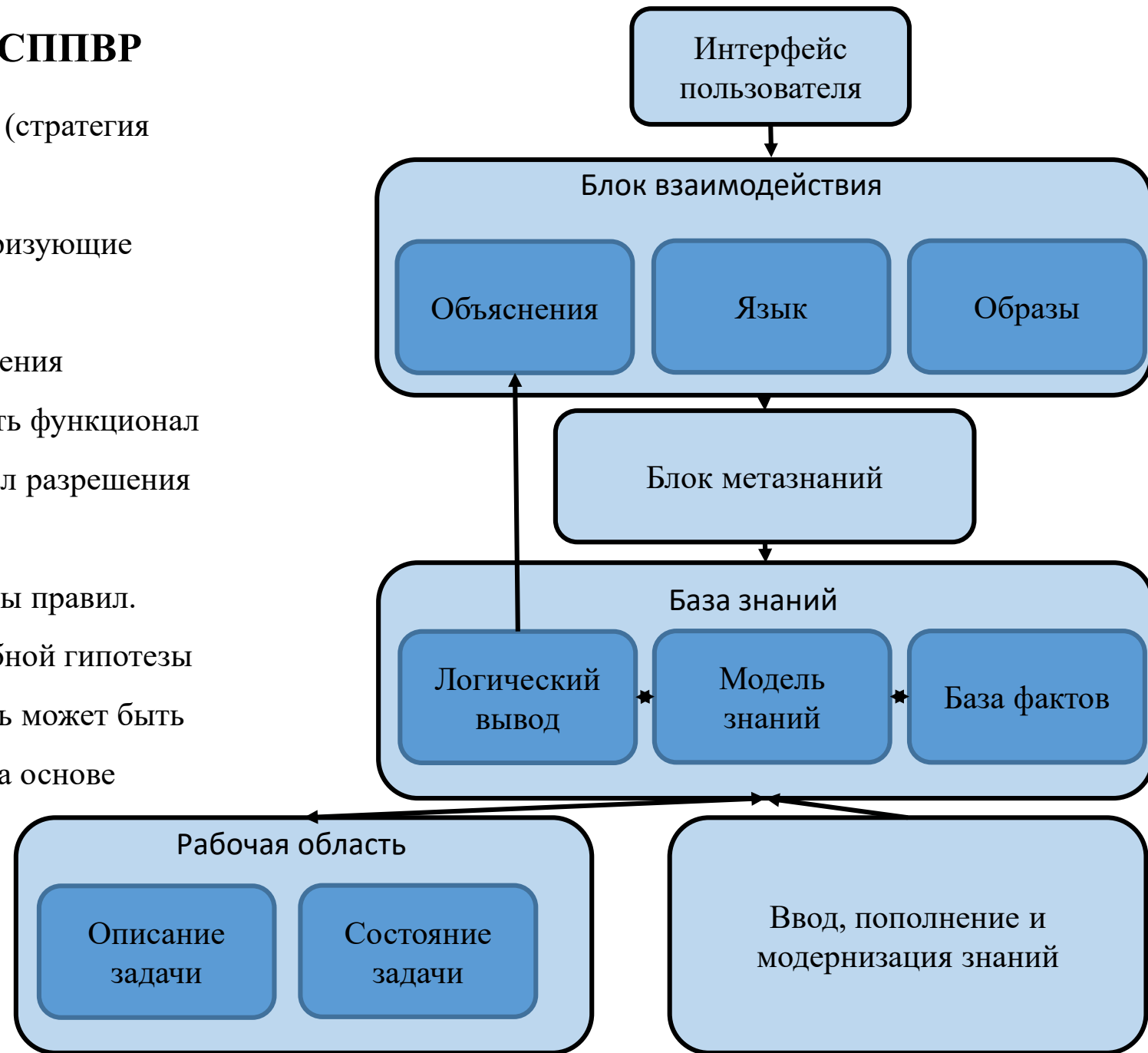
Медицинские Ассистенты: принципы построения систем поддержки принятия врачебных решений

	Экспертные СППВР	Прецедентные СППВР
Логический вывод	Основан на правилах (Rule-Based Reasoning)	Основан на прецедентах (Case-Based Reasoning)
Способы и единицы хранения знаний	<p>База фактов – структурированное множество фактов (утверждений). Например: «Адам является человеком».</p> <p>Базы правил – параметризированные утверждения из 2 и более частей. Например: «Если X является потомком человека, то X является человеком»</p> <p>Алгоритм вывода:</p> <ul style="list-style-type: none">• Проверка истинности какого-либо факта (на основе законов формальной логики из баз данных и правил)• Нахождение множества значений параметра некоторого правила, при котором данное правило превращается в истинный факт.	<p>Прецедент состоит из кортежа <ситуация, решение (лечение), результат></p> <p>Ситуация – множество фактов (пол, возраст, чсс,...)</p> <p>Решение – функция от ситуации и результата по выбору оптимального решения из множества возможных</p> <p>Результат - множество фактов (чсс, вариабельность с.р.,...)</p>
Выбор решения	<p>В первом случае на вопрос типа «Истинно ли А?» машина вывода даёт ответ «Да» либо «Нет», во втором — на вопрос типа «При каких X истинно правило A(X)?» машина вычисляет все возможные значения X, при подстановке которых в правило А это правило превращается в истинный факт.**</p>	<p>Алгоритмы сравнения признаков ситуации с признаками прецедента (ML алгоритмы, например, k-ближайших соседей, деревья решений и т.д.)</p>
Проблема использования	<ol style="list-style-type: none">1. Неистинные факты (факт недоказуем) = ложные факты;2. Все правила субъективны;3. Добавление новой нозологии требует проверки всех правил.	<ol style="list-style-type: none">1. Сохранение ошибочных прецедентов;2. Отсутствие прецедентов (В РФ нет доступных качественных наборов данных);3. Субъективность описания прецедента врачом;4. Полнота данных в описании прецедента (время ремиссии, коморбидность состояния и т.д.)
Пример системы	MYCIN, ДСМ-метод	Айболит, Fruit Street Health

Медицинские Ассистенты: экспертные СППВР

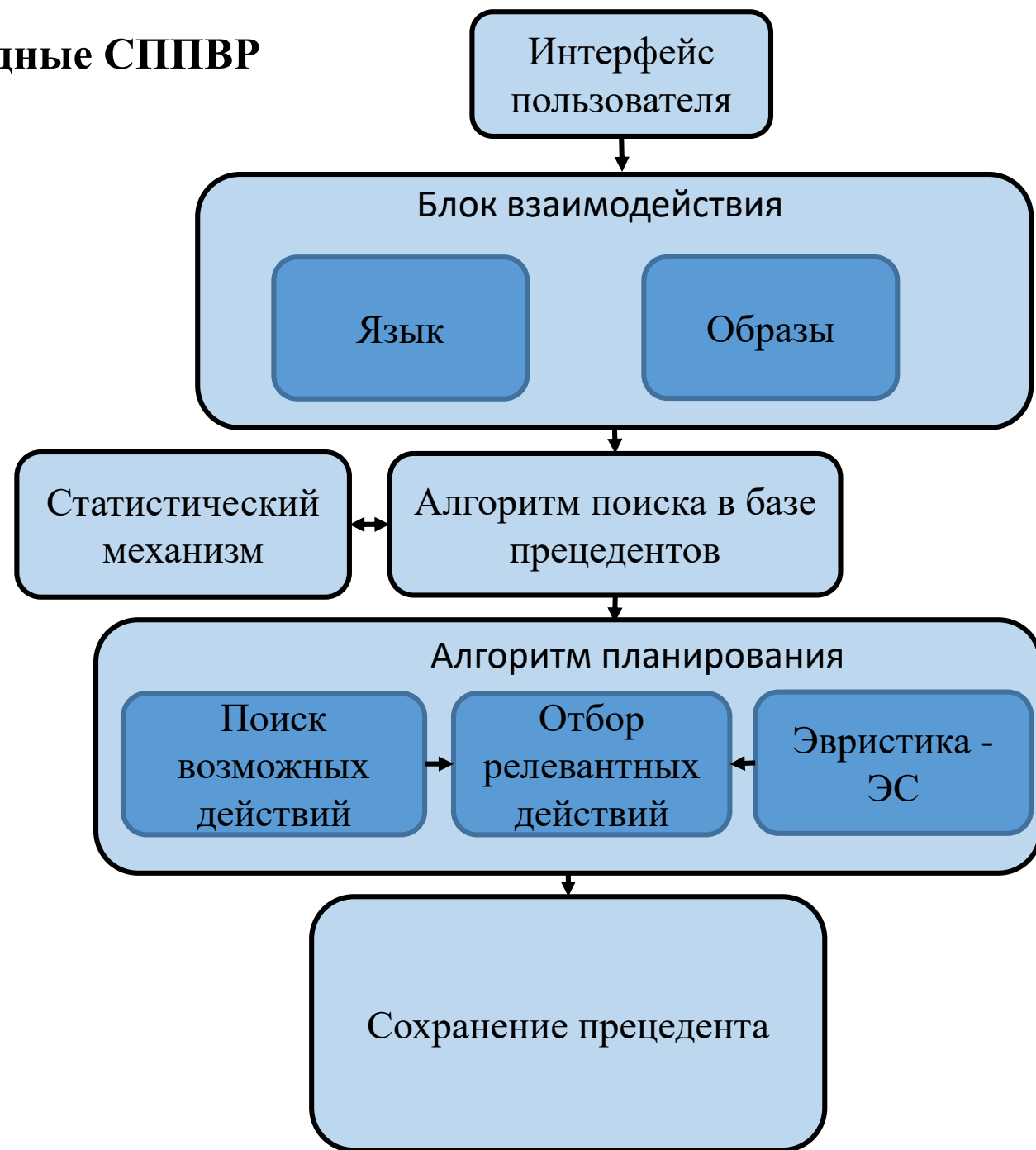
- **Блок метазнаний** – управление выбором правил (стратегия поиска решения, может быть иерархической);
- **База фактов** – позитивные, негативные, характеризующие неопределенность;
- **База знаний** – механизм управления для обеспечения непротиворечивости правил. Должна обеспечивать функционал поиска слабых мест в рассуждениях и функционал разрешения конфликтов;
- **Логический вывод** – удовлетворение, приоритеты правил. Обратный вывод – проверка правильности врачебной гипотезы при движении от диагноза к симптомам. Решатель может быть основан на нечетких правилах вывода (правила на основе нечеткой логики).

*Слайд основан на лекциях проф.,
д.м.н., Б.А. Кобринского.*

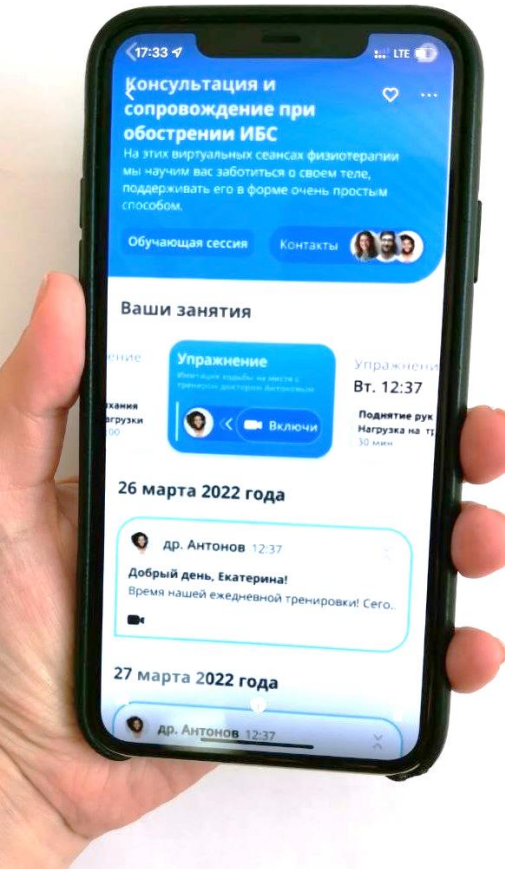


Медицинские Ассистенты: прецедентные и гибридные СППВР

- Жалобы пользователя и информация из МИС формируют вектор-запрос;
- По БД векторов происходит поиск наиболее подходящего вектора (либо тензора, если решение объединено с проблемой) одним из алгоритмов статистики (например, косинусное расстояние, knn, расстояние Дамерау — Левенштейна,...);
- Если решение нашлось — и привело к выздоровлению, то увеличиваем коэффициент уверенности;
- Если решения нет, то используют алгоритм планирования — способ создания прецедента на основе последовательного улучшения состояния пользователя);
- Синтезированное решение сохраняют в БД прецедентов.



Технологическое ценностное предложение



Планировщик
персонализированных
тренировок

Учет
**коморбидных
состояний**

Анализ данных
Google Fit и т.д.

Динамика
изменений
для врача

CNN+LSTM
Планируется
распознавание болезней
Опорно-двигательного аппарата

Модуль
отслеживания
показателей
пользователя



Клинические эффекты:

1. **Улучшение клинического прогноза:** снижение риска летального исхода, развития инфарктов миокарда и повторных оперативных вмешательств на сердечно-сосудистой системе.
2. **Снижение потребности** в антиангинальных, антигипертензивных и сахароснижающих препаратах.
3. **Повышение толерантности к физическим нагрузкам:** прирост абсолютных показателей выносливости (MET's), рост мышечной силы и гибкости.
4. **Улучшение реакции** артериального давления и **частоты** сердечных сокращений на физические нагрузки («двойное произведение»).
5. **Коррекция факторов** риска сердечно-сосудистых событий (гиподинамия, ожирение, расстройства липидного и углеводного обмена).
6. **Повышение качества жизни:** улучшение переносимости бытовых нагрузок, рост рабочих адаптационных возможностей и социальной интеграции.
7. **Улучшение** реалистичной **оценки** собственных физических возможностей.
8. **Снижение** стрессовых реакций, повышение психической устойчивости.
9. **Повышение** приверженности пациента рекомендованному лечению и здоровому образу жизни.

Patient journey (сердечно-сосудистый континуум)



- На каждом из этапов пациент будет задерживаться дольше
- Есть шанс не дойти до инвалидизации и затормозить развитие заболевания
- Подготовка пациента к операции и ускоренному восстановлению

Модель Oran and Braunwald
(упрощенная)

- **Что мы будем проходить?**

Технологии, механизмы и способы создания систем слабого искусственного интеллекта на языке Python.

- **Какой план занятий?**

На лекции – знакомиться с направлениями ИИ и разбирать историю их создания, основные актуальные применения, способы взаимодействия с технологией.

На семинарах – краткий курс в язык Python, как наиболее распространенный механизм создания приложений для ИИ. Математическую постановку задачи ИИ и реализовывать практические задания по ИИ.

- **Формула оценивания:**

$0.7 * \text{накопленную} + 0.3 * \text{экзамен (зачет)} = 100 \text{ баллов по предмету.}$

Накопленная:

40 баллов за Контрольную работу (после 4 Лекции) + 40 баллов за сданные ДЗ (продолжение семинара) + 20 баллов за Проект

Экзамен:

3 теоретических вопроса и 1 практический. 20 баллов за 1 теоретический вопрос * 3 + 40 баллов за решение задачи.

Возможность получения автомата (95 баллов):

Накопленная > 70 баллов, проект > 15 баллов, посещаемость > 80 %.

Оценивание проекта:

1. 10 баллов – работоспособность кода (код запускается и проходит все проверочные тесты);
2. 5 баллов – актуальность задачи (задача является актуальной и способна претендовать на дальнейшее развитие);
3. 5 баллов – каждый из участников проекта понимает весь код и способен ответить на дополнительные вопросы преподавателя.

Пример тем для проектов:

- Часть 1: Создать скрапер новостей, касающихся развития медицины в РФ, льгот для мед. работников преподавателей ВУЗа на основе токенов: «врач, препод, медсестр, рак, инсульт, ...». Полученные новости структурировать на html странице в порядке: картинка – текст. Ссылки для скрапинга: <https://medportal.ru/mednovosti/> <https://minobrnauki.gov.ru/press-center/> <https://rscf.ru/news/> <https://medvestnik.ru/content/roubric/medicine> Предполагаемые библиотеки: bs4, requests. Язык: Python 3.9+. Часть 2: Составьте словарь уникальных слов в тексте публикаций из части 1, отфильтруйте предлоги/союзы и нарисуйте график частотности слов. Часть 3: Произведите замеры важности слов на основе TF-IDF меры. Примените рекуррентную сеть для классификации тематики новостей.
- Часть 1: Установите ROS Galactic или Humble версии. (требуется ubuntu 20+\debian). Пройдите <http://wiki.ros.org/ROS/Tutorials> и создайте стол в Gazebo с 4-6 размерами блоков; Часть 2: Установите пакеты moveit. Запустите роботу и создайте скрипт взятия в актуатор всех типов блоков. Часть 3: Реализуйте алгоритм автоматического построения башни из блоков https://en.wikipedia.org/wiki/Blocks_world.
- Часть 1: Познакомиться с алгоритмом обучения с подкреплением DQN. Установите библиотеку PyTorch и разберитесь с кодом https://pytorch.org/tutorials/intermediate/reinforcement_q_learning.html . Часть 2: познакомиться с применением для Atari <https://www.cs.toronto.edu/~vmnih/docs/dqn.pdf> ; Часть 3: Попробуйте повторить эксперимент из <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6365640/>
- Часть 1: Познакомиться с алгоритмами обучения свёрточных нейросетей и библиотекой Tensorflow (<https://www.coursera.org/learn/introduction-tensorflow>). Выбрать 1 из медицинских датасетов <https://medium.com/@ODSC/15-open-datasets-for-healthcare-830b19980d9> и обучить сеть на выбранных данных. Часть 2: Создать описание сверток, поэкспериментировать с свёрточными и объединяющими слоями. Часть 3: Описать современные подходы к распознаванию образов.

Зарегистрируйте Вашу команду



<https://forms.yandex.ru/cloud/64f5a90802848fdb2f6f54b4/>

Установить Git:

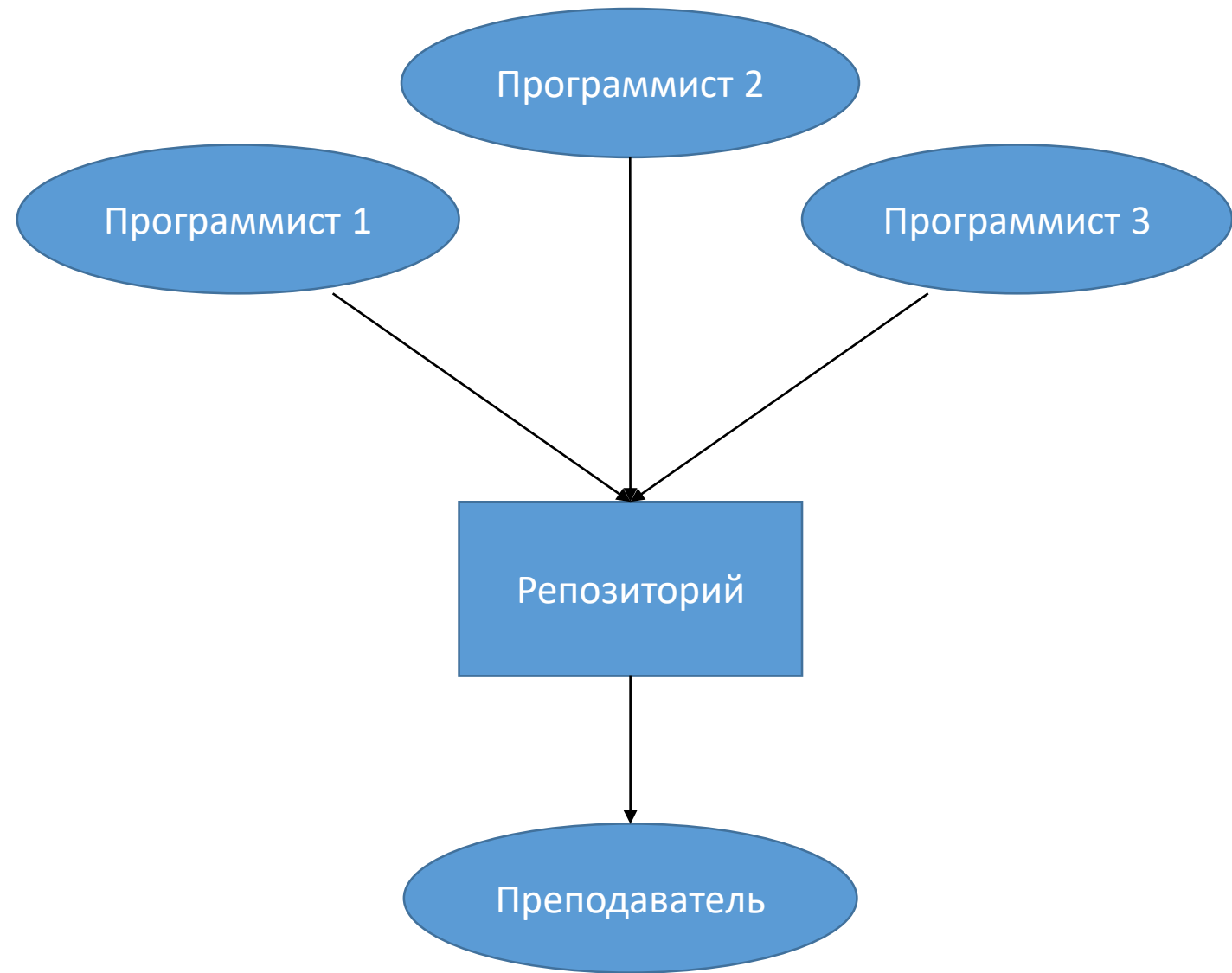
Для windows: <http://git-scm.com/download/win>

Для Linux: `sudo apt-get install git`

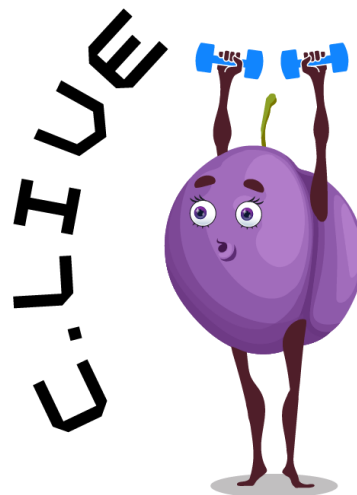
Для Osx: `brew install git` или <https://git-scm.com/download/mac>

Запустить проект в Git:

1. Заходим на <https://github.com/>
2. Создаем профиль/логинимся
3. Нажимаем Repositories
4. Жмем New и вписываем названием репозитория.
5. Копируем ссылку на репозиторий
6. Открываем командную строку (bash, cmd, ...)
7. `cd ПапкаПроекта`
8. `git clone ссылка`
9. Работаем над проектом
10. `git add .`
11. `git commit -m "описание коммита"`
12. `git push origin master`



Спасибо за внимание!



Руководитель проекта Когнитивный ассистент
старший преподаватель, к.т.н. Киселёв Г.А.
+79067993329
kiselev@isa.ru