

Гибридная сеть развития проекта Нейронет

М.В. Сухарев

Обоснование

Человеческое мышление в значительной степени коллективно, о чем пишет, например, Эдвин Хатчинс в своей книге «Мышление в природе» (Edwin Hutchins, *Cognition in the Wild*), определяя такое мышление, как «распределенное»¹ (distributed cognition). Особенно это справедливо для наиболее сложных видов познания, например, научного.

Однако следует отметить общую слепоту англо-американского сообщества к социальному аспекту мышления, связанную с присущим ему индивидуализмом. Поэтому распределенным интеллектом на Западе занимаются немногочисленные группы исследователей (среди членов которых много индийцев, китайцев, континентальных европейцев). Тут у России, в глубине менталитета которой еще не совсем стерлось общинное сознание, есть хорошие шансы сказать свое слово.

В статье П. Смарта и К. Сикары «Коллективное мышление и военные коалиции» утверждается: «Растет понимание распространенности и важности того, что мы могли бы назвать коллективным осмыслением, то есть деятельности, которую группы людей совершают, чтобы развивать понимание на индивидуальном и коллективном уровне».

Но человеческое мышление не только коллективно, оно кроме того давно и широко использует внешние средства «расширения разума»: письменность, рисунки и схемы, счеты и арифмометры, в последнее время во все большей степени использует компьютеры и цифровые сети. В конечном счете оказывается, что «большое мышление» (включая научное создание знаний) совершается большими системами, включающими сотни и тысячи людей и искусственные средства хранения и обработки информации.

На первой же стадии анализа современной науки становится ясно, что отдельный ученый в принципе не способен создавать знания такого уровня. Например, многие ученые пользуются математикой, которую создавали не они, а получили готовую. Физики-теоретики пользуются данными экспериментаторов; причем создать или даже разработать такие установки сами они не в силах. Со своей стороны экспериментаторы создают установки на основе теоретических концепций. Нейрофизиологи пользуются знаниями из физики, химии, теории информации. Вообще, преобладающая доля знаний каждого специалиста и сам язык, на котором он мыслит, получены во время обучения и профессионализации, и только небольшая часть в результате собственных исследований.

¹ О используемой терминологии см. Тезаурус в конце документа.

Многие фундаментальные теории созданы на основе огромного материала, который долгое время собирался тысячами полевых исследователей. Такова, например, биологическая теория эволюции, для создания которой сначала нужно было изучить огромное количество уже существующих и древних животных, без сравнения которых сама мысль о том, что они изменяются, была невозможна.

То есть, созданием научных знаний занимаются не отдельные личности, а научные коллективы (часто очень большие и не имеющие четких границ), которые следует рассматривать, как *когнитивные системы*, снабженные соответствующими артефактами: приборами, средствами хранения, обработки и передачи информации и так далее. Это компьютеры, книги, журналы, осциллографы и т.д. Важнейшей частью этих систем являются средства коммуникации.

До второй половины XX века структура научных сообществ вместе с их журналами, библиотеками и конференциями складывалась стихийно, не являясь объектом научного изучения или специального конструирования. Но после II мировой войны наука становится важным элементом государства, обеспечивающим безопасность и экономический рост. Наука становится большой и дорогой, индустриализируется. Поэтому науку начали подвергать научному же анализу (Р. Мертон, Т. Кун, Д. Прайс и др.).

В последние годы распространение Интернет радикально изменяет структуру науки и технологического творчества; новые направления исследований очень быстро становятся известны по всей планете, возникают новые «невидимые колледжи», базы знаний (Википедия и специальные научные энциклопедии и депозитории). Однако все это ведет к тому, что объем доступной ученому информации становится слишком велик даже с учетом помощи поисковых машин. Требуется внедрение средств искусственного интеллекта (ИИ), семантического поиска и персональных интеллектуальных агентов, что может повысить качество и производительность труда ученых и инженеров в несколько раз. Требуется полностью изменить системы публикации, перенося их в сеть, но при этом сохраняя авторское право и экспертизу научного уровня.

При анализе структур научной коммуникации можно видеть явное сходство со структурами коммуникации нейронов и нейронных ансамблей в головном мозге.

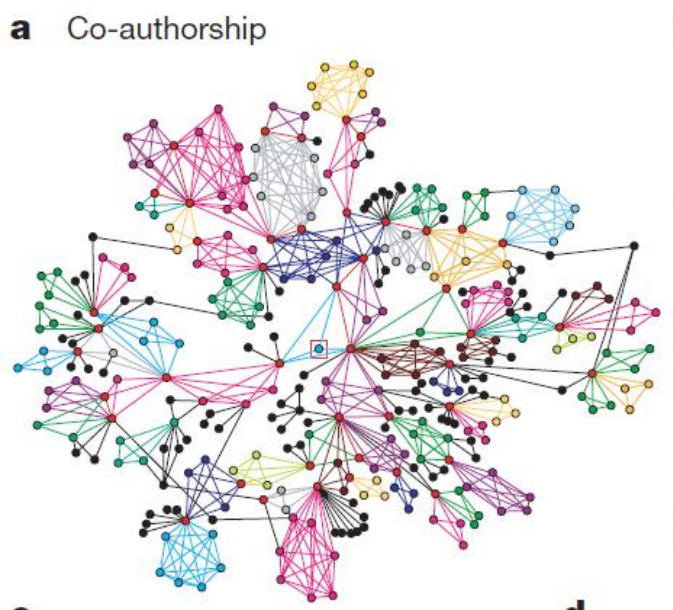


Рис. 1 Сеть соавторства

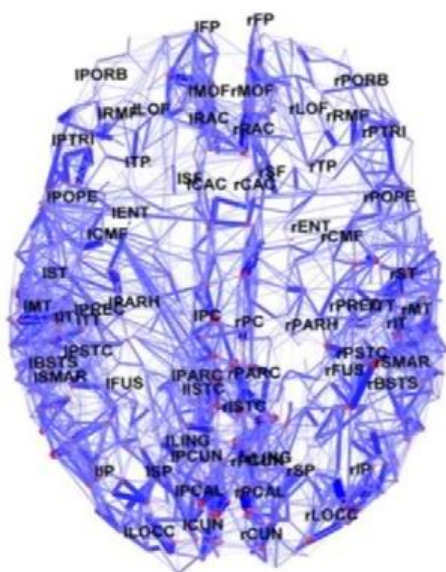


Рис. 2. Коннектом головного мозга человека

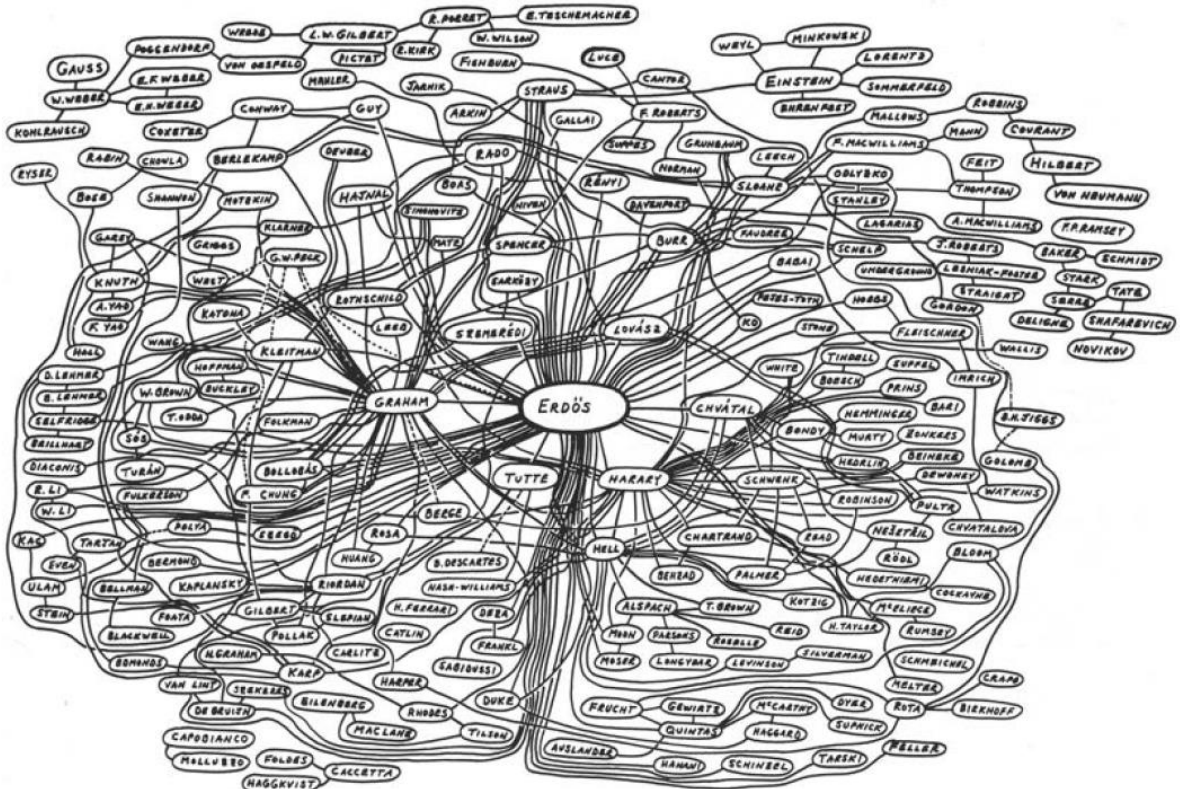
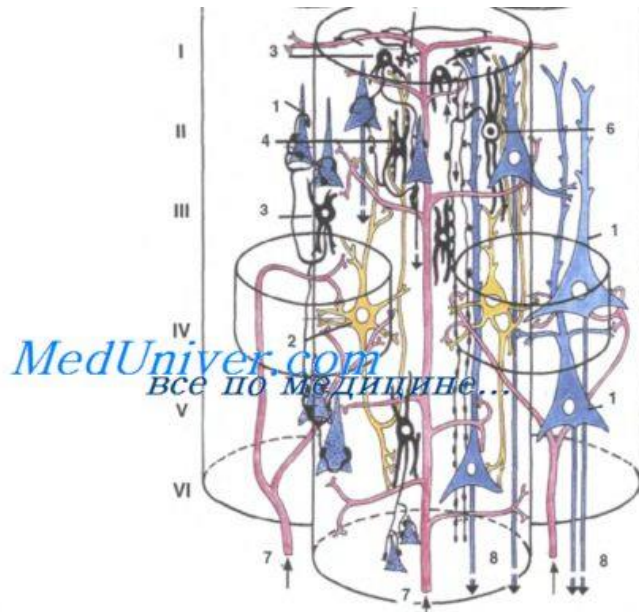


Рис. 3. Коммуникативный граф Пауля Эрдоса



НЕЙРОНЫ.

1 — пирамидные нейроны (синие): основные эффекторные клетки коры (аксоны больших пирамидных нейронов формируют пирамидные пути, идущие к мотонейронам спинного мозга).

2 — звездчатые нейроны (желтые): возбуждают пирамидные клетки.

Тормозные нейроны (черные):

3 — корзинчатые.

4 — аксоаксональные.

Образуют тормозные синапсы на телах и аксонах пирамид, корректируют их ответ на возбуждение.

5 — клетки с аксоаксональной кисточкой. Образуют тормозные синапсы на афферентных волокнах, корректируют входной сигнал.

6 — клетки с двойным букетом дендритов. Тормозят прочие виды тормозных нейронов и тем самым растормаживают пирамиды.

Рис. 4. Колонка нейронов (неокортекс)

Это сходство не случайно. Коннекционизм (connectionism) пытается стать общей теорией когнитивных сетей, независимо от их сущности, биологической, электронной или иной. Основная идея коннекционизма состоит в том, что ментальные явления могут быть описаны моделью сетей относительно простых и часто однотипных взаимосвязанных элементов. «Простых» относительно сети в целом, как нейрон прост в сравнении с мозгом. Хотя первоначально коннекционизм был сформулирован, как методология понимания работы нейронных и нейроморфных сетей, затем возник целый ряд исследовательских групп, применяющий этот метод для изучения социальных сетей, в которых коммуникация идет между людьми.

Коммуникация между людьми может иметь самый разный характер. Могут обсуждаться новые кинофильмы, религиозные вопросы или фасоны кофточек. Организация коммуникации, правила поведения (микроинституты) в этих обсуждениях в чем-то общие, в чем-то различные. *Здесь нас интересуют только системы коммуникации логического типа*, хотя эмоциональные элементы, как элемент, необходимый для мотивации людей, могут присутствовать. Открытия и изобретения быстрее создаются, если ученые и изобретатели испытывают воодушевление.

Такие системы нужны в науке, технологических проектах, управлении, экономике, разработке военных стратегий и других областях, где требуется принимать сложные обоснованные решения. Называть такие сети просто «когнитивными» не стоит; когнитивный подход включает имплицитные знания и эмоциональный интеллект, которые полезны для человека, но мало подвержены логической аргументации.

Рассмотрим, насколько применимы основные положения коннекционизма к логическим когнитивным гибридным сетям.

В последние годы большое внимание вызвала социально-сетевая технология краудсорсинга, в которой знание создавалось усилиями большого количества независимых участников. Всем известный пример, показывающий глобальный потенциал краудсорсинга – Википедия.

Однако краудсорсинг имеет один принципиальный недостаток – огромное количество создаваемой разнородной по качеству информации, которую нужно «просеивать». Но для такой фильтрации требуется большое количество высоко квалифицированных людей, способных выделять действительно ценные идеи из обширного потока зачастую весьма правдоподобной ерунды, к тому же сформулированной на профессиональном языке.

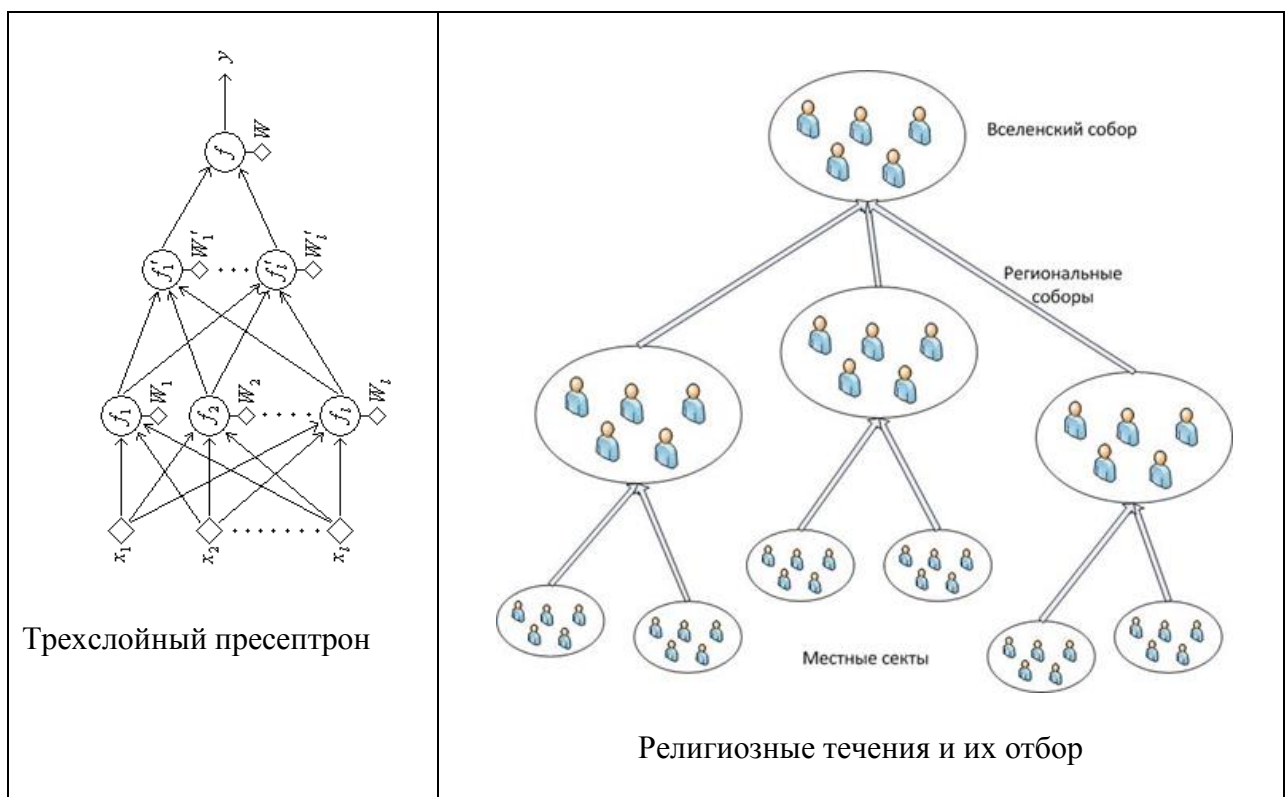
История дает нам примеры решения проблем, похожих на проблемы краудсорсинга. Религия один из старейших. Нам лучше известно христианство. В начале

нашей эры, когда христианство получило широкое распространение, появились многие тысячи проповедников, которые по разному трактовали исходное учение. Они образовывали группы, объединяясь для борьбы со сторонниками других версий.

Возникли первые ереси: гностицизм, манихейство, арианство и так далее. Все эти течения пытались ответить на сложные вопросы, которые люди задавали проповедникам. Например, если Христос – бог, значит он всеведущий. И не может испытывать боль. То есть, он знал, что римляне идут арестовать его и сподвижников. Почему он сидел и ждал этого? Если не испытывал боль и знал, что воскреснет, что такое смерть на кресте? Спектакль?

Религия должна была для дальнейшего усиления научиться отвечать на эти вопросы, причем так, чтобы ответ устраивал абсолютное большинство населения. Для решения проблемы была создана система церковных диспутов, которые предварительно выделяли более убедительные версии. Диспуты происходили на нескольких уровнях, завершая Вселенскими соборами, которые принимали окончательное решение.

Если изобразить эту систему графически, получится нечто очень похожее на многослойный пресептрон.



В дальнейшем через университеты, возникшие в Европе в средние века, такая организация обсуждения проблем и принятия решений была перенесена в науку, которая постепенно приобретала привычную нам форму.

Схожесть с пресептроном не случайна; и в том, и в другом случае требуется собрать информацию из множества источников, взвесить ее значение по ряду параметров и принять решение.

В прошлые века коммуникация и принятие решений о истинности или ложности мысленных конструкций в логических социальных сетях занимала многие годы, но с введением научных и технических журналов, распространением библиотек и университетов, положение значительно улучшилось.

В настоящее время создание Интернет ускорило коммуникацию до предела восприятия; научная статья (в форме working paper) может быть размещена в сети и прочитана через минуты после ее завершения.

Возникло большое количество научных онлайн-журналов, в которых автор платит за публикацию, такие публикации помогают быстро увеличивать цитируемость, в отличие от модели, где платят за чтение статей.

Созданы сайты, где ученые могут размещать свои публикации или ссылки на них, создавать группы по интересам, такие как Researchgate, Academia, Mendeley и др.

Эти усовершенствования научной коммуникации, при всей своей эффективности, создали другую проблему: количество доступной ученому научной информации стало слишком велико для обозрения. По названию статьи и даже по аннотации часто не понять, нужна ли она в действительности, приходится читать, теряя время.

Исправить ситуацию возможно за счет применения элементов искусственного интеллекта, который позволил бы осуществить более точный мониторинг появления новых материалов и их отбор. При современном состоянии ИИ полная автоматизация невозможна, но интеллектуальные агенты, натренированные на конкретную предметную область могут быть вполне эффективны.

Таким образом, возникает общее представление о гибридной когнитивной сети, в которой люди-эксперты работают над решением проблемы с поддержкой систем искусственного интеллекта.

Реализация

Гибридная когнитивная сеть должна включать следующие элементы:

Стандартная часть:

- Собственно цифровая сеть коммуникации. Поскольку многие проекты будут требовать создания режима секретности, такая сеть может быть реализована в форме VPN, через протокол интернет, усиленный стойким шифрованием. Существует open source проект OpenVPN.

- Обычные сетевые средства коммуникации. Можно выделить средства коммуникации, ориентированные на разные временные интервалы публикации и ответа на нее.

Самая медленная – это книги. Написание книг и реакция на них занимает многие месяцы и годы. Реализуется чаще на обычных CMS (Drupal, Joomla! и т.п.) с дополнительными модулями. Основной функционал – поиск; аналоги авторских и предметных каталогов, поиск по ключевым словам.

Более быстрая – статьи. Пишутся за недели и месяцы, время реакции тоже недели и месяцы, хотя некоторые статьи создают дискурс на десятки лет. Реализуется подобно библиотекам.

Еще более быстрая – заметки в Интернет. Время написания – минуты или часы, реакция до нескольких дней. Реализуется средствами интернет-форумов или досок объявлений. Существует множество open source «движков», phpBB, Vanilla и т.д.

Самый быстрый – службы мгновенных сообщений (чат), самый известный пример – твиттер. Известные open source системы – XMPP, его развитие Ignite Realtime, OnlyOfficeJabber и другие.

Элементы искусственного интеллекта:

Персональные интеллектуальные агенты. Помощники эксперта, способные к семантическому поиску.

Реализации: Aigents.com, Open-source JAVA Agent DEvelopment Framework, Open ADK.

Средства обработки естественного языка (natural language processing – NLP). Требуется для постоянного мониторинга контента, порождаемого в сети, поиска дискурсов и возникающих понятий, определения групп интересов. Существуют open-source реализации: GATE и Apache UIMA (Unstructured Information Management applications) project.

База знаний: системы хранения знаний с семантическим поиском. Реализация – Semantic MediaWiki .

Средства моделирования

Набор средств для создания математических моделей различных процессов.