

«Естественный» синапс — это место контакта нейрона (или нейронов) и принимающей сигнал «рабочей» — эффекторной — клеткой (например, мышечной). Искусственный пока имитирует лишь саму «механику» его работы.

Технически устройство, описанное в новой статье в PNAS, представляет собой микроканал конической формы. Через который при получении электрического импульса мигрируют ионы, содержащиеся в растворе. Меняя, таким образом, заряд по разные стороны канала. При этом параметры импульса, естественно, прямо влияют на разницу потенциалов. Иными словами, это достаточно достоверная имитация базового механизма передачи нервных сигналов.

Однако достаточно достоверная не значит идеальная. Во-первых, пока это очень медленный «синапс»: частота передачи сигналов в эксперименте составила 1,7 в минуту. Во-вторых, по биологическим стандартам, устройство получилось чрезвычайно грубым: размеры конуса — 150 на 200 микрометров. Для сравнения: диаметр «биологического» синапса — 1-2 микрометра и меньше. Иными словами, разница примерно в сотню раз. Впрочем, первый транзистор отличался от современных не в 100 раз, а в 200 миллионов раз.

При этом размеры современных транзисторов уступают корейско-голландскому устройству уже в десятки тысяч раз, что автоматически вызывает вопрос о смысле исследования. Первое и простейшее «оправдание» — перспективы создания биологически совместимой вычислительной техники, причем совместимой практически идеально. Очевидно, устройство потенциально способно не только обмениваться с нейронами «естественными» сигналами, но и работать в естественной среде.

Однако в действительности применимость «биоморфных» технологий много шире. Посмотрим на классический компьютер — построенный из кремниевых транзисторов по архитектуре фон Неймана.

Транзистор имеет всего три контакта. По одному из которых поступает управляющий сигнал, а два других посылают соседям один из возможных «ответов» (ток либо проходит через транзистор дальше, либо нет). Как нетрудно догадаться, это автоматически порождает проблемы. Во-первых, для выполнения сколько-нибудь осмысленной деятельности нужно очень много транзисторов. При этом энергопотребление и нагрев растут прямо пропорционально их числу.

Естественным путем решения этой проблемы стала миниатюризация. Однако

сегодняшние кремниевые транзисторы почти достигли физического предела уменьшения. При этом даже смена материала не сможет отодвинуть этот предел действительно радикально.

Во-вторых, задачи в транзисторных «цепях» могут выполняться только «линейно», что резко замедляет вычисления. Иными словами, мейнстримовая миниатюризация в итоге работает с минимальной эффективностью.

Другая особенность классического транзистора — принципиальная неспособность запоминать информацию. Отсюда необходимость во внешней памяти, загружаемой заново каждый раз, что также энергоемкий и затратный с точки зрения вычислительных мощностей процесс. При этом доступ к памяти в каждый конкретный момент ограничен пропускной способностью канала. Иными словами, ортодоксальный вариант компьютера словно специально создан для того, чтобы потреблять энергию и выделять тепло при очень ограниченной отдаче.

Как прямое следствие, за каждым обработанным запросом в ChatGPT стоит расход порядка 2,5 киловатт-часа, а суммарное энергопотребление эквивалентно 17-тысячному городу. При этом на охлаждение тратится две тысячи тонн воды в год.

Между тем пока мы видим лишь прелюдию к экспансии нейросетей, аппетиты которых будут расти. В итоге средний сценарий их развития говорит о триллионах запросов и энергопотреблении на уровне Нидерландов или Аргентины уже к 2027 году.

Иными словами, при использовании традиционной технологии об автономных роботах с интеллектом «чата» можно забыть. И не только из-за габаритов — нетрудно представить себе объем «запросов» при полноценной обработке визуальной информации.

Между тем число нейронов у взрослого мужчины лишь примерно вдвое меньше, чем искусственных нейронов у ChatGPT. При этом в стандартном режиме он обходится примерно 0,03 киловатт-час в сутки и еще примерно 0,07 киловатт-час расходуется на все остальное. Иными словами, суточный расход энергии мозгом — примерно в 83 раза меньше, чем при обработке одного запроса нейросетью.

Безусловно, важнейшую роль здесь играет принципиально другая конструкция человеческих нейронов. Которые, например, имеют в среднем пять тысяч контактов и являются очень условным аналогом не транзисторов, а мемристоров, то есть способны не только обрабатывать, но и хранить информацию. Это очевидным образом снижает

число рабочих «элементов», необходимых для решения конкретной задачи. Однако и потребление энергии в расчете на один биологический «псевдотранзистор», включая передачу сигнала, — в тысячи раз меньше.

Иными словами, если мы хотим построить «процессор» с минимальным энергопотреблением, потребуется так или иначе воспроизвести весь спектр биологических лайфхаков. При этом использование медленной, но экономной передачи сигнала для поддержания работы процессора при ультранизком расходе энергии может быть во многих случаях разумной ценой.