

Neural-Символическое Обучение и Рассуждая: Экспертиза и интерпретация

Тарек Р. Besold

tarek-r.besold@city.ac.uk

Факультет компьютерных наук, города, Университет Лондона

Артур д'Авила Garcez

a.garcez@city.ac.uk

Факультет компьютерных наук, город, университет Лондона

Sebastian Bader

sebastian.bader@uni-rostock.de

Факультет компьютерных наук, Университет Росток

Говард Боумен

h.bowman@kent.ac.uk

Школа Компьютерной, Университет Кента

Pedro Домингуш

pedrod@cs.washington.edu

Факультет компьютерных наук и инженерии Университета штата Вашингтон

Паскаль Hitzler

pascal.hitzler@wright.edu

Факультет компьютерных наук и инженерии, Государственный университет Райта

Кай-Уве Kuhnberger

kkuehnbe@uni-osnabrueck.de

Институт когнитивной науки, Университет Osnabrück

Луис С. Лэмб

luislamb@acm.org

Instituto de Informatica, Федеральный университет штата Рио-Гранде-ду-Сул

Daniel Lowd

lowd@cs.uoregon.edu

Факультет компьютерных и информационных наук, Университет штата Орегон

Присцилла Мачадо Виейра Лима

priscilamvl@gmail.com

НТБ, Федеральный университет штата Рио-де-Жанейро

Лео де Пеннингговская

leo.depenning@illuminoo.com

Illuminoo BV

Гади Пинкасова

pinkas@gmail.com

Центр академических исследований и научно-исследовательский центр Gonda мозга, Бар-Илан, Израиль

Hoifung Пун

hoifung@microsoft.com

Исследования Microsoft

Герсон Zaverucha

gereson@cos.ufrj.br

Сорре, Федеральный университет штата Рио-де-Жанейро

Абстрактные

Изучение и понимание человеческого поведения имеет отношение к информатике, искусственному интеллекту, нейронным вычислениям, когнитивной науке, философии, психологии и ряд других областей. Предполагая познание как основа поведения, в том числе наиболее известных инструментов в моделировании поведения расчетно-логические системы, Коннекционистские модели познания, и модель неопределенности. Недавние исследования в области когнитивной науки, искусственным интеллектом и психологии подготовили ряд когнитивных моделей мышления, обучения и языка, которые подкрепленной вычислений. Кроме того, электронная ФФ орты в информатике исследования привели к развитию когнитивных вычислительных систем

интеграции машинного обучения и автоматизированного мышления. Такие системы показали обещание в ряде приложений, в том числе вычислительной биологии, диагностики неисправностей, обучения и оценки в тренажерах, а также программного обеспечения кванта Вери фи. Это совместное исследование рассматривает личные идеи и мнения ряда исследователей на нейронно-символического познания и рассуждения. Статья состоит из трех частей: во-первых, мы обрамляем масштабы и цели нейронных-символьных вычислений и взглянуть на теоретические основы. Затем мы перейдем к изложению реализаций нейронных-символьных вычислений, систем и приложений. Наконец, мы представляем проблемы, стоящие перед зоной и направления для дальнейших исследований.

1. Обзор

Изучение человеческого поведения является важной частью компьютерной науки, искусственным интеллектом (ИИ), нейронных вычислений, когнитивной науки, философии, психологии и других областях. Предполагающая, что поведение, как правило, определяется и руководствуясь познания и умственной обработки, одним из наиболее известных инструментов в моделировании поведения являются расчетно-логические системы в основном адресации рассуждения и мысли процессов на высоком уровне (классическую логику, немонотонную логику, модальную и временной логики), Коннекционистские модели познания и мозг в основном адресация динамики низкого уровня и возникающие процессы (прямой связь и рекуррентные сети, симметричная и глубокие сетей, самоорганизующиеся сети), а также модель неопределенности адресация зачастую расплывчатый или вероятностный характер многих аспектов когнитивный обработки (байесовская сеть, марковские процессы, марковские логические сети, вероятностные логики программы индуктивных).

Недавние исследования в области когнитивной науки, искусственным интеллектом и психологии подготовили ряд когнитивных моделей мышления, обучения и языка, которые подкрепленных вычисления (Пинкер, Новак, & Lee, 2008; Шастри, 2007; Sun, 2009). Кроме того, недавние е ФФЫ орта в информатике привела к развитию когнитивный вычислительных систем, интегрирующих машинному обучение и автоматизированное рассуждение (Garcez, Брод, и Gabbay, 2002; Garcez, Lamb, и Gabbay, 2009, Valiant, 2000). Такие системы показали обещание в диапазоне применений, включая диагностику неисправностей, вычислительную биологию, обучение и оценки в тренажерах, а также программного обеспечения Вери фантастическом катион (де Пеннинг, д'Авили Garcez, Lamb, & Meyer, 2010, 2011).

Предваряя изложение теоретических основ в разделе 2, интуиция мотивации нейронной-символической интеграции в качестве активного поля исследования заключается в следующем: В нейронных вычислениях предполагаются, что ум эмерджентного свойство мозга, и что вычислительные когнитивный моделирование может привести к уважительной теории познания и о эр и следа понимания некоторых когнитивный процессов (Sun, 2009). Отсюда в своей очереди, предполагается, что коннекционизм должен быть в состоянии о ФФ эр соответствующего изобразительном языке искусственного интеллекта, а также. В частности, коннекционистская вычислительная теория разума должна быть в состоянии повторить параллелизм и виды адаптивных процессов обучения замеченных в нейронных сетях, которые являются общепринятыми в качестве ответственной за необходимую прочность и конечного е и след ectiveness системы в борьбе с здравыми знаниями. В результате, чисто символический подход не был бы су FFI Cient, как утверждают Valiant в (Valiant, 2008).

С другой стороны, логика фи твердо установлена в качестве основного инструмента при моделировании мышления и поведения (Kowalski, 2011; Pereira, 2012) и многими рассматривалась в основном как «исчисление информатики». В связи с этим, часто также неклассические

Логики играют важную роль: временная логика, например, была знаменательно влияние косяка как в науке и промышленности (Pnueli, 1977), и ди-и след различны модальных логик стали лингва-франка для, среди прочего, спецификацией и анализ знаний и общения в многоагентные и распределенные системы (Феджин, Хальперн, Moses, & Варди, 1995). Исследования практического рассуждения в ИИ доминировали немонотонным формализм. Интуиционистская логика может обеспечить адекватную логическую основу для нескольких ключевых областях теоретической информатики, в том числе теории типа и функционального программирования (Van Dalen, 2002). Наконец, описание логик-которые похожи на Крипку модель, играет важную роль в изучении семантической сети (Баадер, Calvanese, Макгинессы, Nardi, и Patel-Schneider,

2003).

Тем не менее, при создании моделей, которые сочетают в себе обучение и рассуждения, один должен примирять методологий различных областей, а именно преимущественно статистических данных и логическому для того, чтобы объединить соответствующие преимущества и обойти недостатки и ограничения. Например, методология нейронной-символическая система стремится перенести принципы и механизмы между (часто неклассическими) логикой на основе вычислений и нейронных вычислениями. В частности, он рассматривает, как принципы символического вычисления могут быть реализованы с помощью коннекционистских механизмов и как subsymbolic вычисления могут быть описаны и проанализированы в логических терминах. Здесь коннекционизм обеспечивает аппаратное обеспечение, на котором уровни ди далее различны абстракции могут быть построены в соответствии с потребностями приложения. Эта методика выглядящий на принципах, механизмах и приложениях, доказал плодотворный путь развития исследований в области нейронных-символической интеграции на протяжении более двух десятилетий, о чем свидетельствует, например, результаты обобщены в обзоре по (Bader & Hitzler, 2005), собранные в книгах (Hammer & Hitzler, 2007) и (Garcez и др., 2009), и сообщили в настоящем обзоре.

Например, в (Garcez и др., 2009), описанный подход привел к прототипичным коннекционистским системам для неклассических рассуждений в попытке Чтобы найти адекватный баланс между сложностью и выразительностью. В этом контексте известные как нейронный-символической система искусственные нейронные сети (ИНС) обеспечивает механизм для параллельных вычислений и надежного обучения, в то время как логика дает необходимые пояснения для сетевых моделей, что облегчает необходимое взаимодействие с миром и другими системами. В интегрированной модели, не конфликт не возникает между непрерывным и дискретным компонентом системы. Вместо этого, плотно соединено гибридная система существует, что непрерывно по своей природе (ИНС), но имеет четкую дискретную интерпретацию (ее логику) на различных уровнях абстракции.

С более практической точки зрения, рациональные агенты часто осмысляются как выполнение приобретения концепции (как правило, без присмотра и статистическое) и манипуляции концепции (как правило, под наблюдением и символическую) в рамках постоянного цикла восприятия и действия. Вопрос о том, как примирить статистический характер обучения с логическим характером мышления, с целью создания таких надежных вычислительных моделей, интегрирующие приобретение концепции и манипуляции, был идентифицирован как ключевой исследовательской задачей и фундаментальной проблемы в информатике (Valiant, 2003) , На этом фоне мы видим нейронную-символическую интеграцию как способ решения проблемы заявили через механизмы перевода знаний и извлечение знаний между символическими логическими системами и сетями subsymbolic.

Есть также важные применения нейронной-символической интеграции с высоким уровнем значимости для промышленного применения. Слияние теории (известный как фоновые знания в ма-

скула обучение) и данные обучения (т.е. обучения на примерах) в ИНС было показано больше е Ф.Ф. еptive, чем чисто символические или чисто коннекционистских систем, особенно в случае реального мира, шумных, неструктурированных данных (де Пеннинговская и др., 2010, 2011 ; Тоуэллы & Shavlik, 1994). Здесь успешно решал сценарии применений включают моделирование бизнес-процессы, сервис-ориентированные вычисления (доверительное управление и предотвращение мошенничества в области электронной коммерции), синхронизацию и координацию в крупных многоагентных системах, а также мультимодальные обработки и интеграцию.

В мультимодальной обработке, например, существует несколько форм рассуждений: сцена, классифицируемый катон может быть достигнут с помощью хорошо обученной сеть дает немедленный ответ после ряда допущений. Изменение в сцене, однако, может потребоваться более специфичные временные, немонотонные рассуждения и обучение на основе данных (на основе, например, на величине изменения в сцене). Некоторые предположения, возможно, должны быть пересмотрены, информацию аннотации изображения может обеспечить контекст различны ди Ф.Ф., похищение и подобие рассуждения пересекающегося сетевые ансамблей может быть необходимо, вероятностные распределения, возможно, придется рассуждали о, и так далее. Интегрированная система должна быстро реагировать, пересмотреть свои ответы в присутствии новой информации, а также контролировать неизбежное накопление ошибок, полученных от реальных данных (то есть доказать свою надежность). Это дает прекрасную возможность для применения нейронных-символической системы.

Остальная часть статьи структурирована следующим образом. В разделе 2 мы опишем принципы нейронного-символического вычисления, пересмотреть теоретические основы о начинании, и дать прототипичный пример того, как можно сочетать обучение и рассуждения в интегрированном виде. В разделе 3 показано применение методики с использованием NSCA, нейронный-символический агент наделен обучение и мыслительные способностями, как первый подробный пример. Раздел 4 относится концепции, лежащих в основе теории сознания в психологии и когнитивный науке и их аналоги в нейронных-символьных вычислениях, перед тем Раздел 5 описывает работу адресации связывающих проблем (введенные в предыдущем разделе), что дает возможность вывод первого порядка вычисляется neural- символические системы. Раздел-затем выдвигает на первый план более технические основы первого порядка (предикат) логика обучения в коннекционистских системах, а затем в разделе 7, введение в марковскую логику и соответствующие сети в сочетании между логикой и графическими моделями. Это приводит к концептуальным соображениям, приведенным в разделе 8, которые касаются нейронного-символического вычисления последних событий в AI и возрожденный интерес к (ре) создание потенциала человеческого уровня с искусственными системами, в частности, в последнее время, например, в области моделирования языка и обработка. Наконец, в раздел 9 краткое описание выбранного в настоящее время существующие и широко признанные подходы к решению основных вопросов нейронной-символической интеграции, возникающей из соседних исследований и след ортов и дисциплин до секций 10 и 11 настоящее предложило направление дальнейших исследований и завершить обследование.

2. Прологомены Neural-символьные вычисления

Цели нейронного-символическое вычисление должны обеспечить последовательное, объединяющее представление для логики и коннекционизма, внести свой вклад в моделирование и понимание познания и, тем самое, поведение, а также для создания более вычислительных средств для комплексного машинного обучения и рассуждений. С этой целью, логические и сетевые модели изучаются вместе в качестве интегрированных моделей вычислений. Как правило, алгоритмы перевода с символическим для

коннекционистское представление и наоборот используются для обеспечения либо (я) нейронная реализация логики, (б) логическая характеристика нервной системы, или (III), система обучения гибрида, который объединяет особенности от коннекционизма и символическом искусственного интеллект.

С теоретической точки зрения, эти электронных ФФЫ орта появляется обоснованной. Согласно нашему текущему знанию и пониманию, как символические / познавательные и суб-символическому / нейронные модели, особенно при фокусировке на физически реализуемые и осуществимые системы (т.е. конечных автоматов физических), а не строго абстрактные модели вычислений, вместе с полученным физическим и концептуальные ограничения-кажется, формально эквивалентны в очень общем смысле: несмотря на то, частично ди далее тч теоретические аргументы, такие как дано (Фавор, 2009), обе парадигмы рассматриваются на практике эквивалент относительно вычислимости (Siegelmann, 1999). Кроме того, с точки зрения сговорчивости, например в (ван Rooij, 2008), эквивалентность на практике в отношении классических размеров анализа (т.е. взаимозаменяемости за полиномиальные накладные расходы, за исключением) была установлена, дополняя и поддерживая предварительное теоретическое предположение об эквивалентности в широкое признание инвариантности Диссертации (ван Эмд Боас, 1990). И, наконец, (Leitgeb, 2005) представил *в принципе* результат существования, показывая, что нет существенной ди далее разностный в изобразительной или решении проблем власти между динамическими системами с распределенными представлениями и символическими системами с немонотонными возможностями рассуждения.

Но в то время как эти фантастические ndings обеспечивают прочную основу для попыток ликвидировать разрыв между коннекционизмом и логикой, многие вопросы, тем не менее, остаются без ответа, особенно при переходе через из области теоретических исследований до внедрения и применения, в частности переключения с композиционных символов, обозначающей идеализированную реальность в практически вещественные векторы, полученных от датчиков в реальном мире: Несмотря на то, вводя основные связи и взаимные зависимости между символическим и subsymbolic парадигмой, уровни анализа довольно грубые и почти все результаты экзистенциальны только характером. Например, при установлении *в принципе* Эквивалентность описана выше, (Leitgeb, 2005) не дает конструктивные методы для того, как на самом деле получить соответствующий символический аналог к югу символической модели и наоборот.

Тем не менее, за последние десятилетия несколько попыток были сделаны при разработке общей нейронной-символической основы, как правило, пытаются применять самый популярные методы их соответствующего время, например, как в настоящее время модульных глубоких сети. Растущее внимание уделяется в последнее время глубоких сетей, где есть надежда, что на высоком уровне абстрактные представления будут выходить из необработанных данных низкого уровня (Hinton, Osindero, и Дэ, 2006). Большинство современной нейронная-символическая система использует прямую связь и рецидивирующий сети, но плодотворную работу в области используемой симметричные сетей (Пинкасова, 1991b) из вида, приложенный в глубоком обучении, и последние работы начинает решать реальные приложения симметричной нейронной-символической сеть (де пеннингговская и др., 2010). Там, в общем каждом уровне нейронной-символической системы представляет собой эволюцию знаний нескольких агентов с течением времени. Каждый агент представлен в сети это кодирование уровня здравого смысла (немонотонному) знаний и предпочтений. В сети / агенты на уровнях различны ди и след могут быть объединены вверх для представления реляционных знаний и вниз, чтобы создать специализации, после того, что известно как Сеть- фи привести методологию (Garcez & Gabbay, 2004).

-Расслоение, которые будут продолжать служить в качестве общего примера в оставшейся части этого раздела-о и далее ERS принципиального способом объединения сетей и может рассматриваться как один из

общие методики нейронного-символической интеграции. Основная идея сети Интернет принести проста: фантастические разводили сети могут состоять из соединенных между собой нейронов, как обычно, но и других сетей, образуя рекурсивную структуру. Фи привести функцию де фи NES, как эта сетевая архитектура ведет себя; она де фи определяет, как сети должны быть связаны друг с другом. Как правило, фи принести функции позволит активацию нейронов в одной сети, чтобы повлиять изменение веса в другой сети В. Наглядно это можно рассматривать как учебную сеть

В в то же время, что сеть бежит. Albeit будет сочетанием простого и стандартного ИНС, фантастические разводили сети могут аппроксимировать любую полиномиальную функцию в неограниченной области, таким образом, быть более выразительным, чем стандартное Feedforward сетей.

Расслоения является лишь один пример того, как принципы из символьных вычислений (в данном случае, рекуррентные), могут быть использованы коннекционизма для продвижения исследований в этой области. В оставшейся части этого раздела мы рассмотрим более подробно, и в пути манифеста с кратким изложением нашего точки зрения (ы) на нейронных-символической интеграцию-принципы, механизмы и приложениях, которые ведут исследования в нейронной-символической интеграции.

2.1 Принципов Neural-Символических интеграции

С начала коннекционизма (McCulloch & Pitts, 1943) -arguably первая neuralsymbolic системы булевых логических-самых neuralsymbolic систем сосредоточились на представлении, вычислительной техники, и изучение языков, кроме классической логики (Browne & Sun, 2001; Cloete & Zurada, 2000; Garcez и др, 2002, 2009; Holldobler & Kalinke, 1994; Шастри, 2007), с много е фф Орт быть посвящена представлению фрагментов классической логики первого порядка. В (Garcez и др., 2009), был предложен новый подход к представлению знаний и рассуждений, создание Коннекционистский неклассические логики (в том числе коннекционистского модальный, интуиционистской, височной, немонотонной, эпистемных и реляционной логики). Совсем недавно было показано, что основы аргументации, абдуктивного рассуждения и нормативные многоагентных систем также могут быть представлены одним и тем же сетевым рамками. Это поощрение в той степени, что разнообразие форм рассуждения могут быть реализованы одним и тем же, простую структуру сети, которая специализируется на ди и далее различными способами.

Ключевой характеристикой многих нейронных символических систем, является модульность. Один из способов построения нейронных сетей-символических через бережную инженерию сетевых ансамблей, где модульность затем выполняет важную роль для понятности и технического обслуживания. Каждая сеть в ансамбле может быть ответственна за специфичны задачи или логики, с общей моделью потенциально очень выразительна, несмотря на его относительно простых компонентов. Тем не менее, хотя и является довольно распространенным, модульность не является строгой необходимостью: в качестве альтернативы, как описано в разделе 5, можно начать с неструктурированной сети (т.е. не модульной), и пусть изменения веса формировать свою способность обрабатывать символические или subsymbolic представления.

Другое общее организационное свойство нейронных-символические сети дельное с глубокими сетями-их вообще иерархической организацией. Сеть нижнего уровня принимает необработанные данные в качестве входных данных и производит модель данных. Сеть следующего уровня занимает выход первой сети в качестве входных данных и производит некоторое высокоуровневое представление информации в данных. Сеть следующего уровня затем дополнительно повышает уровень абстракции модели, и так далее, пока некоторое представление высокого уровня не может быть изучено. Идея заключается в том, что такие сети могут быть обучены независимо друг от друга, возможно, также сочетая неконтролируемым и поднадзорное обучение на ди уровней и далее различными иерархий. Результирующий

параллельная модель вычислений может быть очень мощной как это о Ф.Ф. ERS дополнительной выразительность требует сложных приложений при сравнительно низких вычислительных затратах (Garcez и др., 2009).

2.2 Механизмов Neural-Символических интеграции

Мы разделяем мнение, что представление изначально предшествует обучению. Нервная-символическая сеть может представлять собой диапазон выразительных логик и реализовать некоторые важные принципы символьных вычислений. Однако, нейронные-символьные вычисления не только о представлении. Механизмы распространения активации и других методов передачи сообщений, градиент-спуска и других алгоритмов обучения, рассуждения о неопределенности, массовым параллелизмом, отказоустойчивости и т.д. являются важной частью нейронной-символической интеграции. Проще говоря, нейронные-символической сети являются электронные FFI Cient вычислительные модели, а не представительные инструменты. Именно механизмы в виде алгоритмов FFI Cient e, которые позволяют вычислительные возможности нейронных-символической системы.

Возвращаясь к примеру фи принести, также фи разводили сети вычислительные модели, а не только графические модели или математические абстракции, такие как графики или сети в целом. Нервная-символическая сеть может быть отображена непосредственно на аппаратное обеспечение, которое обещает, например, реализация в очень-крупномасштабной интеграции (VLSI) чипе быть простым и экономически е и далее еptive. Главное архитектурное ограничение, что здесь мозг вдохновил, что нейронная-символическую система должна реплицировать и специализироваться простые нейронные структуры, в которой один алгоритм может быть применен электронным FFI ciently на ди уровней и далее различны абстракции, с получившаяся система способна проявляющее поведения эмерджентного ,

Именно это появление в общей отличительной черте коннекционистских подходов, которые диктуют необходимость механизмов извлечения знаний. Таким образом, цикл нейронной-символической интеграция включает в себя (I) перевод символических (фоновый) знаний в сеть, (II) изучение дополнительных знаний из примеров (и обобщения) по сети, (III) выполнение сети (то есть рассуждения) и (IV) символическое извлечение знаний из сети. Extraction обеспечивает объяснение, и облегчает обслуживание и инкрементные или передачи знаний.

В общей нейронной-символической системе, сеть ансамбль A (представляющее, например, временная теория) может быть объединена с другим сетевым ансамблем B (представляющее, например, эпистемическое состояние агента, в). Снова используя фи принести в качестве иллюстративного механизма, метауровень знание в одной сети может быть интегрировано со знанием объекта уровня в другой сети. Например, один может рассуждать (в мета-уровне) о действиях, которые необходимо предпринять (в объектном уровне) для решения несоответствий в базе данных. Реляционная знания также могут быть представлены таким же образом, с отношениями между понятиями, закодированных в различных (объектного уровня) сети потенциально быть представлены и узнали через сеть мета-уровня. Более конкретно, если две сети обозначают понятия $P(X, Y)$, а также $Q(Z)$ содержащий переменные X, Y а также Z , соответственно, сеть мета-уровень может быть использован для отображения представление μ а также Q на новую концепцию, скажем, $R(X, Y, Z)$, таким образом, что, например, соотношение

$P(X, Y), \mu Q(Z) \rightarrow R(X, Y, Z)$ правомочно (Garcez и др., 2009).

Рисунок 1 иллюстрирует прототипичную нейронную-символическую систему. Модель, в самом общем виде, позволяет количество сетевых ансамблей, чтобы быть объединены в ди уровнях и далее различны абстракции посредством, например, фи принести. В рисунке, каждый уровень представлен элементом

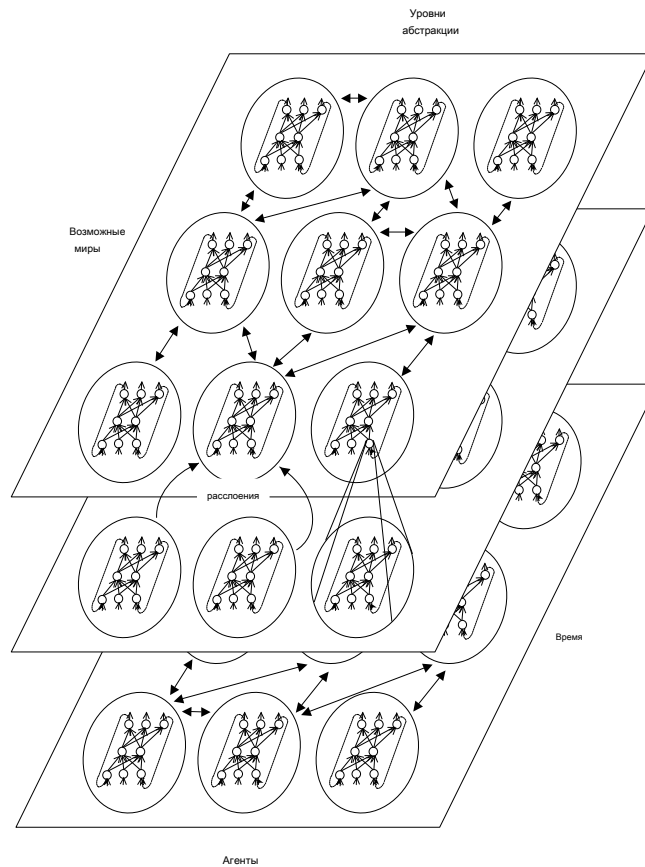


Рисунок 1: (. Garcez и др 2009) Общий обзор концепций нейронной-символической системы.

сеть ансамбль в горизонтальной плоскости, в то время как сеть фи приносит происходит по вертикали между сетями в ди Ф.Ф. различны ансамблях. Специализация происходит вниз, когда нейрон является фи выведен на сеть. Реляционная знание представлено вверх, когда несколько сетей объединены на сети мета-уровня. Эволюция знаний с течением времени происходит на каждом уровне, как и альтернативные результаты, и немонотонное и эпистемическое рассуждения для нескольких, взаимодействующих агентов. Модульное обучение происходит внутри каждой сети, но также применяется в нескольких сетей в ансамбле. Же структура мозга стиле реплицируется по всей модели таким образом, что один алгоритм применяется на каждом уровне и на разных уровнях.

2.3 Применение нейронные-символьного интегрирования

Как это часто бывает, что в конечном итоге будет через успешные внедрения и реальные приложения, что полезность и важность нейронной-символическую интеграция будут общеизвестными. Практическое применение является важным компонентом и является постоянным

Особенность исследования нейронных-символьные вычисления. В теории бокового параллельно с поисками понимания взаимосвязи и интеграции между коннекционизма и символики в качестве основы познания и интеллекта, мы заинтересованы в находка пределы представления и разработки более эффективных методов машинного обучения, которые могут открыть новые возможности применения рассмотрение. На практике уже в текущем состоянии исследований, реальные приложения возможны в районах с общественной значимости и / или потенциально высоким экономическим последствиям, как биоинформатика, семантической сети, диагностики неисправностей, робототехника, программное обеспечение систем VERI фи катиона, моделирование бизнес-процессов, предотвращение мошенничества мультимодальные обработки, шумовой анализ текстов, а также обучение и оценка в тренажерах.

Все эти области применения ди Ф.Ф. различны имеют нечто общее: требуется вычислительная система, которая способна учиться на собственном опыте и который может рассуждать о том, что было изучено (Browne & Sun, 2001; доблестный, 2003). Для этого обучения, рассуждения процесса, чтобы быть успешным, система должна быть устойчивой (таким образом, так что накопление ошибок, вытекающих из внутренней неопределенности, связанной с предметной областью можно управлять (Valiant, 2003)). Метод, используемый в исследованиях на нейронной-символической интеграции, чтобы включить некоторые из вышеупомянутых приложений были (я) использовать алгоритмы перевода между логикой и сетями (применяя соответствующие доказательства эквивалентности), (б) для изучения систем эмпирического пути тематические исследования (после практической мотивации от статистического машинного обучения и нейронных вычислений) и (III), чтобы сосредоточиться на потребностях приложения (отметив, что некоторые потенциально интересные приложения требуют только даже элементарных логических представлений, например, (Becker, Mackay, & Dillaway, 2009)). Пример нейронной-символической системы, которая уже предоставляет вклад этого типа проблем в области биоинформатики и диагностики неисправностей является Коннекционистским Индуктивным обучением и логическое программирование (CILP) Система (Garcez & Zaverucha, 1999; Garcez и др, 2002). Точно так же, NSCA Система обсуждается в разделе 3 была успешно применена, например, для моделирования поведения в тренажерах.

2.4 Нейронные-Символическая интеграция в Nutshell

Таким образом, нейронные-символьные вычисления включают в себя интеграцию высокого уровня познавательных способностей (включая индукционные, дедукцию, и похищение), а также исследование того, как мозг делает ментальные модели (Тэгард & Stewart, 2011), среди прочего, также покрывающего моделирования эмоций и внимание / полезности. На вычислительном уровне, он обращается к изучению интеграции логики, вероятности и обучения, а также разработка новых моделей вычислений, сочетающих надежное обучения и электронной FFI Cient рассуждения. Что касается приложений, успехи были достигнуты в различных областях, включая моделирование, биоинформатики, диагностика неисправностей, разработка программного обеспечения, проверки модели, обработки визуальной информации и предотвращения мошенничества.

В следующем разделе, в целях дальнейшего ознакомления читателя с общим мышлением мы будем иметь детальный взгляд на одной такой систему, а именно нейронной-символической архитектуры когнитивного агента под названием NSCA -aligning представление с тремя категориями принципов, методов и приложений.

3. NSCA как в примере приложения для Neural-символьных вычислений

По общему мнению, что е и далее е effective интеграции автоматизированного обучения и познавательного мышления в реальных приложениях является ди FFI Культ задачей (Valiant, 2003). Как правило, большинство приложений работают с большими объемами данных, наблюдаемых в реальном мире, содержащих ошибки, пропущенные значения, и несоответствия. Даже в контролируемых условиях, таких как тренажеры, интегрированное обучение и рассуждения не очень успешно (Сандеркок, 2004; Heuvelink, 2009). Хотя использование моделируемых сред упрощается сбор данных и знаний, это все еще очень ди FFI культ построить когнитивную модель (интеллектуального) агента, который способен справиться со многими сложными отношениями в наблюдаемых данных. Например, когда дело доходит до оценки и подготовок высокого уровня сложных познавательных способностей (например, руководство, тактическое маневрирование, безопасное вождения и т.д.) обучение по-прежнему руководствуются или делается человеческими экспертами (ван ден Бош, Riemersma, Ши FFL ETT, Эллиот, Салас, и Coover, 2004). Причина заключается в том, что эксперт поведения на высоком уровне познания является слишком сложным для моделирования, выявить и представить в автоматизированной системе. Среди прочего, там может быть много временных отношений между низкими и высоким порядком аспектами учебной задачей, человеческое поведение часто недетерминировано и субъективное (т.е. предвзятые личным опытом и другими факторами, такими как стресс или усталость), и то, что известно часто описывается неопределенно и ограничивается явной (т.е. «объяснимым») поведения.

Несколько попыток были сделаны для решения этих проблем. Например (Fernlund, Гонсалес, Georgiopoulos, и DeMara, 2006) описывает целый ряд систем, которые используют машинное обучение, чтобы узнать сложные отношения из наблюдений экспертов и стажеров во время выполнения задачи. Хотя эти системы являются успешными в обучении и обобщения, им не хватает выразительности символических систем и, следовательно, ди FFI культ для интерпретации и проверки (Smith & Kosslyn, 2006). В качестве альтернативы можно было бы добавить вероятностного рассуждения к logicbased систем (Heuvelink, 2009). Эти системы работают лучше выражать свое внутреннее знание, поскольку они используют явные символические представления и способны справиться со многими из наиболее распространенных типов несоответствий в данных, рассуждая с вероятностями. К сожалению, когда дело доходит до представления знаний и моделирования, эти системы все еще требуют либо статистического анализа больших объемов данных или представления знаний вручную. Таким образом, оба подхода время дороги и не подходят для использования в приложениях реального времени, которые требуют обучения в режиме онлайн и рассуждения.

3.1 Принципов Neural-Символической интеграция Exempli фантастическом изда

Построение е и след е effective моделей когнитивный агентов является давним исследованием усилия в искусственном интеллекте, когнитивный науке и многоагентных системах (Валиант, 2003; Wooldridge, 2009). Одной из главных задач в направлении достижения таких моделей является предоставление комплексных познавательных способностей, таких как обучение, рассуждения и представления знаний. Нервная-символическая система стремится сделать это в рамках парадигмы нейронных вычислений, интегрируя индуктивное обучение и дедуктивный (ср (Garcez и др 2009; Lehmann, Bader, и Hitzler, 2010) для примеров). В таких моделях ИНС используется для изучения и рассуждать о знании (мысленном агента) о мире, в лице символической логики. Для того чтобы сделать это, алгоритмы отображения **логических теорий (или знания о мире) T в сети N который вычисляет логические последствия T . Это обеспечивает также систему обучения в сети, которые могут быть обучены примерами использования T в качестве фоновых знаний. В агентах, наделенных нейронным вычислением, индукция обычно рассматриваются как процесс изменения веса**

сеть таким образом, чтобы отражать статистические свойства набора данных, что позволяет для обобщения более невидимых примеров. В том же месте, вычит нейронная вычисление выходных значений в ответ на входные векторы (кодирование стимулов из окружающей среды) с учетом определенного набора весов. Такие сетевые вычисления были показаны диапазон эквивалентных временных логических формализмов (Агнец, Borges, & Garcez, 2007). Основываясь на таком подходе архитектура агента называется Neural Символическое Когнитивная агент (NSCA)

был предложен в (де Пеннинговская и др., 2011). NSCA использует временную логику как теорию T и больцманово машина Ограниченной (RBM) в ИНС H . УКР частично соединен ИНС с двумя слоями, видимый слой V и скрытый слой $ЧАС$, и симметричные соединения W между этими слоями (Смоленского, 1986).

УКР де фи определяет распределение вероятностей $P(V = v, H = \text{час})$ по парам векторов v а также час закодированы в этих слоях, где v кодирует входные данные в двоичных или реальных значениях и час кодирует вероятность задней $P(H/v)$. Такая сеть может быть использована для вывода или реконструкции полных векторов данных на основе неполных или противоречивых входных данных и, следовательно, реализовать автоассоциативная память или автоассоциатор. Она делает это путем объединения распределения вероятностей задних генерируемое каждый блок в скрытом слое с распределением условной вероятности для каждого блока в видимом слое. Каждый скрытый блок ограничивает ди Ff различны подмножество размеров в многомерных данных, представленных на видимом слое и поэтому называется эксперт по некоторой функции во входных данных. Вместе эти скрытые блоки могут сформировать «продукт экспертов» модель, которая сдерживает все размеры в исходных данных.

NSCA способен (I), выполняющее изучение сложных временных отношений от неопределенных наблюдений, (б) рассуждения вероятностно о знании, что было изучено, и (III), представляющее знание агента в формате логической основой для целей проверки. Это достигается за счет использования нейронных обучения для выполнения надежной обучения и адаптации, а также символические знания представляют собой качественные рассуждения. NSCA было подтверждено в тренажере, используемой в реальных сценариях, иллюстрирующий использование еptive е ФФ подхода. Результаты показывают, что модель агент способен научиться выполнять автоматическую оценку водителя от наблюдения данных моделирования в реальном масштабе времени и оценки по инструкторов по вождению, и что эти знания могут быть извлечены в виде временных правил логики (де Пеннинговская и др. , 2011).

3.2 Механизмы Neural-символьного интегрирования Exempli фантастическом изд

Использование рекуррентной Temporal Restricted больцмановский машина (RTRBM) (Sutskever, Хинтон, & Taylor, 2009) в качестве специализированного варианта общей концепции УОР, NSCA кодирует временные правила в виде гипотез о верованиях и ранее применяемых правил. Это возможно из-за повторяющиеся связи между скрытыми единицами активациями на время T и активации на время $T - 1$ в RTRBM. На основе байесовского вывода механизма RTRBM, каждый скрытый блок $ЧАС$ представляет собой гипотезу о гр правиле Специфического p которая вычисляет вероятность того, что заднее правило предполагает определенное соотношение в верованиях b наблюдается в видимом слое B , с учетом ранее применяемых правил p_{T-1} (т.е. $P(R/B = B, R_{T-1} = p_{T-1})$). Из этих гипотез RTRBM выбирает наиболее применимые правила p используя случайную выборку Gaussian распределения вероятностей задней (т.е. $p \propto P(R/B = B, R_{T-1} = p_{T-1})$) и вычисляет условную вероятность или вероятность всех верований, учитывая выбранные правила

применяются (т.е. $P(B/R=p)$). -Ди-разностный и далее между наблюдаемыми и предполагаемыми верованиями может быть использован NSCA обучать RTRBM (т.е. обновить его вес), чтобы улучшить гипотезы о наблюдаемых данных. Для обучения, RTRBM использует комбинацию сопоставительной дивергенции и обратное распространение во время.

В духе (Братман, 1999) 'ы убеждения, желание, намерение (BDI) агенты, наблюдаемые данные (например, данные моделирования или человеческие оценки) кодируются как верования и ди и след разностной между наблюдаемыми и предполагаемыми верованиями фактических последствий или намерения агента на его окружение (например, адаптируя результаты оценки в тренажере). Значение веры представляет собой либо вероятность наступления какого-либо события или состояния в окружающей среде (например, *Льется = TRUE*), или реальное значение (например, *Скорость = +31,5*). Другими словами, NSCA рассматриваются как двоичные и непрерывные данные, например, с помощью видимого слоя непрерывной стохастической (Chen & Murray, 2003). Это улучшает способность агента к модели асимметричной информации, которая, в свою очередь, является очень полезной, так как измеренные данные, поступающие из тренажера часто асимметричны (например, учебные задачи, как правило, имеют место в ограниченной области моделируемого мира).

Из-за случайный характер функций активации сигмовидной, верования можно рассматривать как нечеткие множества с гауссовой функцией принадлежности. Это позволяет представить неопределенные понятия, как *быстро* а также *медленный*, а также приближения изученных значений, что полезно при рассуждении с неявным и субъективным знанием (Sun, 1994).

Когнитивный временная логика описана в (Лэмб и др., 2007) используется для представления знаний предметной области в терминах верований и ранее применяемых правил. Эта логика содержит несколько модальных операторов, которые расширяют классическую логику модальной с понятием прошлого и будущего. Для того, чтобы выразить верования на непрерывных переменных, эта логика распространяется с использованием равенства и неравенства формул (например, *Скорость < 30*, *Льется = истина*). В качестве примера рассмотрим задачу, когда стажер диски на городской дороге и подходы к перекрестку. В этом случае обучаемый должен применить *Уступать - в - правильно* править. Используя расширенную временную логику, можно описать правила об условиях, сценарии, и оценки, связанные с этой задачей. В правилах (1) до (4) в таблице 1, обозначает « правда когда-нибудь в будущем»и $A \wedge B$ обозначает « A

было верно, так как возникновение B ».

условия:

(1) (*Погода > хорошо*)

что означает: погода по крайней мере, хорошо

Сценарий:

(2) *ApproachingIntersection* \wedge ? (*ApproachingTra FFI* $c = \text{справа}$)

что означает: автомобиль приближается к перекрестку, и когда-нибудь в будущем тра FFI с приближается с правой стороны

(3) ((*Скорость* > 0) \wedge *HeadingIntersection*) S (*DistanceIntersection* $< \text{Икс}$) \rightarrow *ApproachingIntersection* смысл: если автомобиль движется и движется к перекрестку, поскольку она была сочтена близко к перекрестку, то автомобиль приближается к перекрестку.

Оценка:

(4) *ApproachingIntersection* \wedge (*DistanceIntersection* $= 0$) \wedge (*ApproachingTra FFI* $c = \text{правильно}$) \wedge (*Скорость* $= 0$) \rightarrow (*Оценка = хорошо*)

это означает: если автомобиль приближается к перекрестку и прибывает на перекрестке, когда тра FFI с грядет справа и останавливается, то обучаемый получает хорошую оценку

Таблица 1: Ситуации и оценки из ведущего сценария имитатора.

Правило (4) является примером неопределенного понятия, что является весьма субъективным (расстояние *Икс* в котором человек рассматривается как приближение к перекрестку зависит от ситуации

и личный опыт). Когда это правило кодируется в RTRBM, становится возможным узнать более объективное значение *Икс* на основе наблюдаемого поведения диалогов людей и след различен в различных сценариях. Это exempli fi es основной цели объединения рассуждений и обучения.

Временные правила логики представлены в NSCA путем установки весов соединений в RTRBM. Поэтому правила должны быть переведены в форму, которая относится только к непосредственно предыдущему временному шагу (обозначаемый временному оператор \bullet). Алгоритм преобразования для этого также описано в (Лэмб и др., 2007). Тогда мы можем закодировать любое правило в качестве стохастической связи между скрытым блоком, который представляет собой правило, видимые блоки, которые представляют верования и предыдущие скрытые активации единицы, представляющая применяемые правила в предыдущем шаге времени. Например, правило $\alpha S \beta$ могут быть переведены по следующим правилам: $\beta \rightarrow \alpha S \beta$ а также $\alpha \wedge (\alpha S \beta) \rightarrow \alpha S \beta$, где α а также β моделируется видимыми блоками, $\alpha S \beta$ скрытой части, и $\wedge (\alpha S \beta)$ моделируются рецидивирующим подключением к одной и той же скрытой единице. (Пинкасова, 1995) показывает, как отобразить эти логические на основе положения в энергетической функции симметричной сети и как иметь дело с неопределенностью, вводя понятие конъюнкции, называется штраф, как описано ниже.

3.3 Применение Neural-Символическое Интеграция Exempli фантастическом изд

NSCA был разработано в рамках трехлетнего исследовательского проекта по оценке в тренажерах. Он реализован как часть платформы мульти-агент для виртуального обучения (де Пеннингговский, загрузки, и Карп'е, 2008) и был использован для эксперимента на симуляторе вождения. В этом эксперименте пять студентов приняли участие в испытании вождения, состоящей из пяти тестовых сценариев каждый. Для каждой попытки всех данные из симулятора (т.е. 43 измерений, как относительные положения и ориентация все тра FFI с, скорость, шестерня и оборотов в минуту автомобиля студента) и численные оценки баллы по нескольким навыкам вождения (т.е. управление транспортного средства, экономическое вождение, тр FFI с поток, социал- и безопасное вождение), которые были предоставлены три инструкторов по вождению, присутствующих во время попыток наблюдало NSCA в настоящее время. NSCA был в состоянии извлечь из этих наблюдений и вывести результаты оценки, которые похожи на те из инструкторов по вождению (де Пеннингговский и др., 2010, 2011).

NSCA Система также была применена в рамках системы Визуальное Intelligence (VI), которая называется CORTEX в программе Eye DARPA в Мышления (Донлон, 2010). Эта программа преследует способность учиться вообще применимые и генеративные представлениями действий между субъектами в сцене (например, лица, автомобили, объекты) непосредственно из визуальных входов (т.е. пикселей), а разум об изученных представлениях. Ключевое различие между глазом Мышления и внедренный в VI в том, что последний сделал большой прогресс в осознании широкий круг лиц и их свойства (которые можно было бы рассматривать как существительных в описании сцены). Глаз Мышления стремится добавить возможность описать и причину о действиях (которые можно было бы рассматривать как глаголы в описании сцены), что позволяет более полно описательную визуального опыта. Эта цель поддерживается

NSCA поскольку она стремится научиться описывать действия в терминах сущностей и их свойств, обеспечивая объяснение рассуждения позади него. Результаты показали, что система способна изучать и представлять основные семантики действий от наблюдения и использовать это для нескольких VI задач, как распознавание, описание, обнаружение аномалий и gap-заполнение (де пеннингговского, ден Холландер, Bouma, G., & Garcez, 2012).

Совсем недавно, NSCA был включен в Интеллектуальной транспортной системы с целью сокращения выбросов CO₂ (де пеннинговская, д'Авила Garcez, баранина, Stuiver, & Meyer, 2014). Результаты показывают, что NSCA способно распознавать различные стили вождения от реального времени данные датчиков в автомобиле и обгоняет внедренный в этой области. К тому же, NSCA способно описать эти сложные стили вождения с точкой зрения временных правил логики на основе многоуровневых для человеческой интерпретации и проверок экспертов.

3.4 NSCA в двух словах

В итоге, NSCA является примером для когнитивной модели и агента архитектуры, о ФФ ERS электронной и далее еctive подход интегрирующей символическое мышление и нейронную обучение в уни модели фи-е изд. Такой подход позволяет агенту, чтобы узнать правила о наблюдаемых данных в сложных реальных условиях (например, поведение экспертов для подготовки и оценок в тренажерах). Узнал поведение может быть извлечено, чтобы обновить существующие знания домена для проверки, отчетности и обратной связи. Кроме того, подход позволяет знания предметной области должны быть закодированы в модели и имеет дело с неопределенностью в реальных данных. Результаты, описанные в (де Пеннинговская и др., 2010, 2011) показывают, что агент способен изучать новые гипотезы из наблюдений и извлечь их во временную логическую формулу. Но хотя результаты являются многообещающими, модель требует дальнейшей оценки вождения экспертов. Это позволит дальнейшей проверке модели в оперативной обстановке с большим количеством сценариев, большим стажером населением, а также нескольких оценок по инструкторам по вождению. Другие текущей работы включает в себя исследования по использованию Деер больцмановские машины (Салахутдинов & Hinton, 2009; Tran & Garcez, 2012) Ф.И. правила й высокоуровневых и применение в RTRBM для облегчения адаптивного обучения. В целом, эта работа иллюстрирует прикладную модель для представления знаний, обучения и рассуждения, которые на самом деле могут привести к реалистическим вычислительным моделям когнитивный агентов, что позволяет решать задачи, выдвинутые в (Valiant, 2003; Wooldridge, 2009).

4. Neural-Символическая интеграция в и когнитивной науки: Строительство Психические модели

До сих пор мы разработали на общих теоретических и концептуальных основ neuralsymbolic интеграции и обсудили NSCA В качестве примера успешной реализации и развертывания приложений нейронной-символической системы. В дальнейшем, мы теперь вернемся более подробно некоторые из наиболее актуальных теоретических и прикладных основных вопросов, в том числе обязательной задачи, коннекционистский логики первого порядка (ВОЛП) обучения, сочетание вероятностей и логики в марковских логических сетях, и взаимосвязь между нервным-символическим мышлением и человеческим уровнем AI-и описать наше нынешнее состояние знаний, а также возможные будущие события, в следующих разделах. Мы начнем с введения в тесной связи между нейронными-символической интеграции и основные темы из когнитивной науки.

Для того, чтобы полностью объяснить человеческое познание, мы должны понять, как мозг построить и манипулировать ментальные модели. Клари фи катион этого вопроса проинформирует развития (искусственно) интеллектуальных систем и, в частности нейронного-символических систем. Понятие психической модели имеет долгую историю, возможно, прежде всего, использованных Джонсон-Лэрд, чтобы обозначить свою теорию человеческого мышления (Johnson-Laird, 1983). Принимая частичное вдохновение (Тэгард, 2010) и (Нерсисян, 2010), мы интерпретируем ментальную модель как когнитивная

представление о реальной или воображаемой ситуации, отношения между частями ситуации, в и, возможно, даже, как эти части могут действовать друг на друга. Эти модели будут обязательно сохранять ограничения, присущее тому, что изображено, а их оценку и манипуляция Утверждались быть ключом к человеческому разуму (Johnson-Laird, 1983). Например, вид пустой дороги и россыпью разбитого стекла, может колдовать мысленное представление (т.е. модель) уголовного разбитое стекло Вашего автомобиля, что привело вас к разуму, что ваш автомобиль был украден. В более абстрактном примере, когда размышлял о том, почему проблема исключающее ИЛИ линейно неразделимы в контексте сетевого обучения, можно представить себе двухмерную плоскость, содержащая четыре регулярно размещенные кресты (две для истинно, два несовпадения) и мысленно рассмотреть размещение из Классификации граничных линий фи катионными. Изучая такую ментальную модель, можно убедиться в том, что линия деления истинные и ложные выходы не могут быть найдены.

В частности, были утверждать, что строительство и вывод на ментальных моделях могут играть важную роль в абдуктивном рассуждении и, даже, творчество (Тэгард & Stewart, 2011), наша кража автомобиля примером является случаем первого. Точно, как ментальные модели и их манипуляции являются neurally реализованы, однако, остается в значительной степени неопределенным. Центральное место в построение мысленной модели является формированием комбинированного (вышестоящего) всего представление из наборов (подчиненных) часть представлений. Например, в примере кражи автомобиля, композитный ваш автомобиль и капюшон фигура разбив окна могут быть созданы из памяти и фантазии автомобиля и рисунка в изоляции.

Самый основное требование нейронной реализации такой комбинации представления является решением так называемым «связывающей проблема», то есть отождествление общего нейронного механизма для представления которых отдельные узлов активных нейронов (например, те, которые представляют красный Форд и те, кто представляет капюшон фигуру) связаны и которые не являются (также см раздел 5.2). Многие предложения были сделаны для механизма связывания мозга, например, временной синхронности (Engel & Singer, 2001), присоединительные коды (O'Reilly, Басби, & Soto, 2003; Bowman & Wyble, 2007; Риготти, Рубин, Ван, & Фуси подходы 2010) и даже свертка на основе (Тэгард & Stewart, 2011). Из них, присоединительные кодов, кажется, включают наималейший шаг от традиционных курсов кодировки ИНСА, например, всплески и колебательная динамика не предполагается, и, таким образом, кажется наиболее естественным интегрируемым с нейронной-символической системой в настоящее время сформулированы.

Под конъюнктивные кодами приближаются, то предполагается, что блоки, которые избирательно реагируют на coactivation нескольких представлений элементов доступны. Такие присоединительные аппараты могут находиться в общедоступном связывании ресурса, такие как, например, связывание бассейн в (Bowman & Wyble, 2007). Две проблем в конъюнктивный код связывании подходов (и, соответственно, к нейронному представлению ментальных моделей) являются (I) масштабируемости, и (II) новые конъюнкции:

- Привязка ресурс, который исчерпывающе перечисляет все возможные комбинации представлений в качестве уникальных (localist) единиц не масштаба. Таким образом, реалистичные конъюнктивные коды обычно предполагают, что связывание ресурсов используют распределенные представления, которые обеспечивают более компактными, и действительно масштабируемым, кодировки действующих ассоциаций, смотрите, например, на изысканиях этого вопроса в (Риготти-d и др., 2010) и (Wyble & Bowman, 2006).

- Мы часто испытываем комбинации представительских, что мы никогда не испытывали раньше. Стандартный пример является пресловутым синим бананом или, в контексте примера кражи автомобиля, вполне вероятно, что ментальная модель conjoining с капюшоном фигуры, красный Ford, и конкретный подъездным путем никогда бы не была ранее испытывала. (На самом деле, тот факт, что каждый опыт обязательно обязательно иметь на появление первого делает точку.) В результате, присоединительные подходы кодирования предлагают случайного PRECON фи gured связывания ресурсов с такой дискриминацией способностью, что они могут е слы estively соединить любое сочетание элементов представлений, которые могут возникнуть в будущем; модель фронтальной функции лепестка в (Риготти-d и др., 2010) является таким подходом.

Нервные присоединительные методы кодирования такого рода были бы, то, как представляется необходимым условием для brainbased теории ментальных моделей.

Очевидно, что способность строить мысленные модели имеет ограниченную ценность, если она не помогает нам принимать решения и, в целом, разумеется, например, для определения того, что преступление было совершено, и что полиция должна быть вызвана. Есть два способа, в которых современная когнитивный нейронаука предполагает такое рассуждение может возникнуть через любого из статистических выводов или совещательной разведки. Первый из них может использовать статистический вывод, основанный на сопоставлении с образцом, в манере классического ИНС, как и отраженного в использовании дескрипторов «ассоциативного обучения» в этом контексте. В отличие от этого, последний будет более явно совещательный характер и, соответственно, будет включать в себя серийный исследование пространства гипотетических возможностей. Такой процесс может быть самым естественным образом оформлен с точкой зрения классических символьных вычислений и, возможно, поиск в производстве системных архитектур (Newell, 1994). Кроме того, статистический процесс вывода можно рассматривать как более непрозрачны для осознания, чем обсуждение.

Двойные теории процесса рассуждения, например, (Evans, 2003), утверждают, что подразделение такого рода является основой человеческого мышления. Что касается нашего основного интереса здесь, хотя, второй из этих форм мышления, совещательной разведки, кажется наиболее естественным образом, чтобы охватить понятие ментальной модели, с коннотацией формирования сознательных гипотетических представлений.

Современная когнитивная неврология также подчеркивает роль эмоций / тельцам государственных оценок в рассуждении. В частности, эмоция может рассматриваться как обеспечение оценки оценочным качества ситуации или альтернатив, многое в манере (Дамасио, 2008) 'ы соматических маркеров. Но главное, теории ТФ т.д., такие, как, скажем, теория оценки (Шерер, 1999), предлагают многомерное пространство эмоционально заряженных чувств, богаче, чем требуется простое вознаграждение и наказания или функции полезности. Например, как отвращение и страх отрицательно valenced, маркировка, если вы хотите, наказывая опыт; однако, аверсивный ответы на два, как правило, довольно ди Ф.Ф. различны. В частности, только страх, скорее всего, заниматься весьма срочных FI ответами IGHT или фи ОТВ, связанных с непосредственной физической опасностью. Действительно, можно утверждать, что ключевая функция эмоций выделить типы реагирования и, по крайней мере, в том смысле, различия между эмоциями по многим параметрам когнитивно значимы. Это относится к когнитивным и neuroscienti фи гр счетов нейронного-символической интеграции в том, что ментальные модели, возможно, то, можно рассматривать как «штампованные» с оценками и далее estive, которые будут связаны (в порядке, ранее обсуждалось) до нона и след estive участков ментальной модели, что дает богатую ТФ ЕСТ-цветной высокой размерной представление.

4.1 Когнитивные нейронауки и Neural-символьные методы

Когнитивный неврология (Газзанига, Иврите и др. 1998) является одним из основного тока Научным проекта, с объяснением познавательных способностей (например, восприятием, памятью, вниманием, и языком) с точки зрения их реализации в мозге, как главная цель. Численное моделирование (особенно когда ИНС на основе) является ключевым элементом программы исследований. Концептуально, такое моделирование обеспечивает «клей» между познанием и мозгом: она позволяет бетон (disproveable) объяснения должны быть сформулированы, как когнитивные способности может быть сгенерирован мозгом. В этом смысле, модель, объясняющая познавательное поведение, а также данные нейровизуализаций сделать особенно хорошую работу по преодолению между когнитивный и нервным. Примеры таких моделей включают в себя (Bowman 2006; Chennu, Craston и соавт 2009; Коуэлл и Cottrell 2009; Craston, Wyble и др, 2009), а весь проект Dynamic Причинная Modeling (. Friston, Харрисон и соавт 2003) можно увидеть иметь эту цель.

Затем это выдвигает на первый план пригодности ди Ф.Ф. различны когнитивных парадигм моделирования в когнитивной неврологии; то есть, как следует вычислительные модели в когнитивной нейронауки быть сформулированы? Если рассматривать когнитивный и мозг наук в очень широком определению, хотя вычислительные модели многих сортов были использованы, есть две доминирующие традиции. Первый из них является символическим моделированием, которые **наиболее часто формулируются в производственной системе архитектур, такие как SOAR (Newell, 1994), EPIC**

(Kieras, Мейер, Мюллер, и Сеймура, 1999) и АКТ-Р (Андерсон и Lebiere, 2014; Андерсон, Fincham, Цинь, и Stocco, 2008). В противоположность этому, вторая традиция моделирование сети, начиная от абстрактных Коннекционистский, например (Rumelhart, McClelland, и PDP Research Group, 1986), в нейрофизиологически детальные подходы, например, (Hasselmo & Wyble, 1997; Frans'en, Tahvildari, Еропов, Hasselmo, и Алонсо, 2006). С самых ранних вычислений на основе предложений для изучения ума и мозга, которые, вероятно, датируются 1950-х годов, главенство колебалось между символической и нейронной сети на основе. Тем не менее, современный интерес в мозге и картографической познания к ней, привело к длительному периоду нейронной сети превосходстве и, конечно, в когнитивной нейронауки области, символическое моделирование теперь редко; наиболее известные исключения, являющиеся (Anderson и др 2008; Taatgen, Juvina, Schipper, Borst, & Martens, 2009).

Там, хотя остаются в меньшинстве, кто не согласен с нейронных сетей выразительности. Некоторые из них приняли очень сильно анти-нейронных сетей точки зрения, например, (Fodor & Пилишина, 1988), утверждая, что мысль принципиально символично-лингвистическое действительно, ср (Fodor, 2001) -И, что она должна, таким образом, обязательно изучаться на этом уровне. Компьютерные системы часто принимаются как оправдывает катион для этой позиции. Следует отметить, что программное обеспечение является интересным уровнем, который определяет функциональное поведение компьютера в любой момент. В отличие от этого, утверждаются, что отображение для аппаратной реализации, например, компиляция до ассемблера, является неподвижным предопределенным определены преобразованием, которое не нужно перепрограммировать, когда алгоритм бежится изменяются. То есть, алгоритм выполняется определяется программой, а не отображение в аппаратных средствах, или перефразировать, в компьютерных системах, программное обеспечение, где функциональная сложность заключается. Ключевой момент в том, что программное обеспечение является символическим. Если один вид познание, как отмечено его функциональное богатство, можно затем сделать вывод, что познание наиболее целесообразно изучено символический и что отображение в мозг является фиксирован неинтересным шаг реализации, который не отражает богатство и разнообразие познавательного опыта и поведения.

Это положение делает, конечно, зависеть от аналогии между строением компьютерных систем и структурой ума мозга. Многие отвергнут эту аналогию. В самом деле, большинство из тех, которые выражают несогласие с сетевым моделированием, принимает менее экстремальную линию. Признавая актуальность осуществления мозга, они утверждают, что ИНС выразительно ограничены в качестве метода когнитивного моделирования, или по крайней мере подчеркнуть, что ИНС не приводят к естественному моделированию некоторых когнитивных функций. Ниже приведены случаи, в точке.

(Я) решение проблем Правила наведения: Дункан изучил человеческую способность кодировать и применять последовательности правил, регулирующих как выполнять умственные задачи (Duncan, Парр, Вулгар, Томпсон, Брайт, Кокс, епископ, и Ниммо-Смит, 2008). такие *правила задачи* утверждает, что FUL фи lment из набора свойств мандатов ответа конкретного вида. Кроме того, эти свойства и отношения между ними могут быть довольно сложными, например, «если все буквы на одной стороне экрана находятся в столицах, а стрелка точку по диагонали вправо вверх на другой стороне, то реагировать в направлении, указанном большой треугольник». Дункан и его коллеги показали, что производительность их правила задачи экспериментов сильно коррелируют с флюидным интеллектом и, таким образом, с IQ. Производительность также связано с активацией областей мозга, как полагают, участвует в е и след ortful задачи поведение, подчиненное (Dumontheil, Thompson, & Duncan, 2011). Таким образом, кажется, эти эксперименты правила задачи будет выявление нейрокогнитивных основ интеллекта и решения проблем; ключевая проблема для науки. Естественный способ думать об этих правилах задач является как символические операции с логическими предпосылками и действиями, которые определяем конкретный ответ. В этом контексте участники стремятся правильно оценивать предпосылки в отношении предъявляемых стимулов и применять соответствующие (правильные) действия. Ясно, что эта точка зрения может быть непосредственно отраженными в модели производственной системы. Наконец, один из ключевых вопросов для исследования правил задачи состоит в том, характеризующем, как делаются ошибки. В значительной степени, ошибки, как представляются, возникают из предварительных условий и / или действий, мигрирующих между операциями (т.е. между правилами). Опять же, такие ошибки, естественно, могут быть смоделированы в разработке производственной системы, в которых предварительные условия и действия могут misbind друг с другом.

(Б) Центральные Исполнительные функции: Понятие централизованной системы управления, которая направляет мысль (центральную исполнительную) часто встречается в теории обработки информации познания, ср Baddeley и модель рабочей памяти НІТСН (в Baddeley, 2000), наблюдательный внимани система Shallice (в Norman & Shallice, 1986) и Центральный двигатель в Барнарде Cognitive подсистемы взаимодействующей (Барнарда, 1999). Центральный Исполнительный является, например, гипотеза прямых сенсорных, перцептивных и двигательных систем, в частности, наиважнейшая доминантных (стереотипных) ответов. Соответственно, он будет участвовать в отображении Организмы текущих целей в стратегии для реализации этих целей (а, так называемые, поставленная задача). Кроме того, было бы навязать эту задачу установить при обработке мозга через централизованный контроль. Действительно, Центральный Исполнительный будет участвовать в кодировании и применение правил задач, подобных только что обсуждали, а также сознательное изучение последовательностей альтернатив в поисках стратегии для получения цели, то есть то, что в AI терминов будет называться планирование. Нейрофизиологическая локализация центрального органа исполнительной власти была сосредоточена на лобной доле и решение де дефицитов, возникающих от лобного повреждения лепестка (Norman & Shallice, 1986), в частности проблемы. Ряд наиболее выдающихся работников в этом поле сформулировал свои модели в символических системах производства, а не ИНС. Одним из таких подходов использует COGENT Моделирование обозначение высокого уровня (Shallice & Купер, 2011), в котором компонента, закодирована в производстве

системы, выполняются параллельно, при условии межкомпонентного взаимодействия. В частности, символические методы искусственного интеллекта, кажется, хорошо подходит для моделирования некоторые классические ЦЕНТРАЛЬНОГО ИСПОЛНИТЕЛЬНОГО (лобная) задачи, такие как, Башни Ханоя Башни Лондона и планирования в целом (Shallice & Cooper, 2011).

(III) синтаксические структуры: Длинным работают дебаты в психолингвистике противопоставили объяснения на основе правил и ассоциации на основе языковой морфологии, с Пинкером обычно переноской фла AG для бывших и McClelland и со-пропагандистов последнего (Пинкер и Ульмана, 2002). Самые горячие дебаты были сосредоточены на особенно отличительный (U-образной) профиль данных развития, возникающих, когда дети изучают английский язык в прошедшем времени в отражении; то есть, когда они определяют, что добавление «эд» су FFI х навязывает прошедшее время интерпретации (например, «пребывание» до «остался»), при условии соблюдения многих исключений (например, «бежать», чтобы «бежал»). Хотя сосредоточено на одном специфическом аспекте языка, эта дискуссия служит soundingboard для более широкого рассмотрения: являются регулярными синтаксическими и морфологическими аспектами языка в целом, лучше всего рассматривать как символические правила или выходящий из подготовки параллельной распределенной обработки (PDP) -стиль сети. Эти дебаты специфичны к регулярным аспектам языка; есть еще мнение, что исключения могут быть, естественно, обрабатываются PDP сетей. Критическая точка для connectionists является то, что, в случае PDP, нет никакого смысла, к которому любое конкретное правило или, на самом деле, абстрактное понятие правило, встроен в сеть априори; сеть, в некотором смысле, «открывает» закономерность; на самом деле, если вы хотите, понятие регулярности. Хотя правила против объединения дебаты, безусловно, бежать и бежать, с точкой зрения нашего внимания здесь, это демонстрирует, в наименее воспринимаемом-несоответствии между классическими коннекционистскими подходами и характеристиками когнитивного поведения. Кроме того, предлагаемый ответ на это несоответствие является символическим моделирование метафоры, которые, естественно, могут быть реализованы в классических методах искусственного интеллекта, такие как, производственные системы, логическое программирование и т.д.

(IV) Композиционность: Одна характеристики проблемы коннекционизм су и след ERS является то, что, естественно, не отражает репрезентативную композиционность присущих многих познавательных способностей (Фодор & Пилишин, 1988). Например, когда части объединяются в целое, те части, по крайней мере, в значительной степени, несут их интерпретации с ними, например, «Джон» в «Джон любит Джейн» есть, якобы, тот же смысл, как это делает в «Джейн любит Джон». Хотя ди Ф.Ф. различны классы ИНС придерживаться этой характеристики для ди и след различны степени, это, конечно, кажется, что классические модели PDP-стиль не автоматически генерировать репрезентативную композиционность. По крайней мере, когда принято по номинальной стоимости, PDP моделей демонстрируют позиционно-SPECI схем кодирования с фи; то есть, тот же элемент (например, «John») должны быть отдельно узнали в позиции X и Y позиции, без автоматического интерпретационного переходящий из одного положения в другое. Следовательно, элемент может иметь очень ди Ф.Ф. интерпретации различны в двух положениях и, в самом деле, что стало известно о «Джон» от его появления на позиции X (где это, возможно, было замечено часто) не будет автоматически информировать понимание «Иоанна» в положении Y (где можно было увидеть редко, если вообще). Это не фи т с субъективным опытом: время первых мы видим, «Джон» в той или иной грамматической позиции, увидев его в другом месте часто, не, казалось бы, проявляются как полное незнакомство с понятием «Джон». Хотя это и не всегда явно формулировались с точки зрения композиционности, прерывистые нападения на PDP подходы сторонники localist схем нейронные кодирования имеют сходную флavour (страница,

2000). Действительно, критика localists' прорезь кодирования используется в моделях PDP слова чтения (Bowers, 2002) очень критик позиционно-специфичным кодирования.

Эта неудовлетворенность коннекционизм предполагает символическую перспективу, так как символические модели foundationally композиционные в моде ИНС аргументированы не быть. Если принять эту массу аргументов, один остается с точки зрения, что в то время как ссылка на мозг, конечно, критически (на самом деле, мозг не есть только нейроны, синапсы, и т.д., для обработки с), он не может быть практичным карта непосредственно нейронные реализации. Скорее всего, когнитивное моделирование на промежуточном символическом уровне может быть более целесообразным.

В частности, ИНС, безусловно, сделал глубокий вклад в понимание познания. Это хотя и было наиболее выражено для специфичного набора познавательных способностей. Например, были разработаны убедительные модели сенсорных процессов, например, этапы «что» визуальной обработки пути (т.е. вентральный поток), Cf (Li, 2001; Raizada & Гроссберг, 2003; Серр, Олива, & Поджио, 2007) ; и лицо SPECI обработки фи гр в пределах того, что токопровод, сра (Burton, Брюс, и Джонстон, 1990; Коуэлл и Cottrell, 2009). Зрелые нейронные модели зрительного внимания через пространство и время также были предложены, например, (итти, Koch, & Нибур, 1998; Heinke & Хамфрис, 2003; Bundesen, Habekost, & Kyllingsbæk, 2005; Bowman & Wyble, 2007). Есть также сложные нейронные модели систем памяти и обучения в целом, например, (Bogacz & Brown, 2003; Norman & O'Reilly, 2003; Davelaar & Usher, 2004). Кроме того, многие из этих моделей являются убедительными в их нейрофизиологических деталях.

Однако, как оправданно выше, есть области, где нейронное моделирование имело меньший успех. В общих чертах, нейронные модели архитектурного уровня не хватает, то есть модели, которые broadscope, общее назначение, и утверждают, что актуальность за пределами специфичного когнитивного феномена. Есть, однако, многие модели символической архитектуры уровня ума, например, SOAR (Newell, 1994) - возможные причины этой ситуации были обсуждены, в частности, в (Bowman & Su, 2014). Он также может быть так, что нейронные объяснения становятся менее убедительными, как когнитивные функции становятся более высокого уровня; то есть, они хорошо подходят для характеристики «периферийные» систем, таких как, зрение, слух и двигательных действия, но они делают меньше, а также один шаг вверх по иерархии обработки от входа и выхода. В частности, то, что Sloman бы назвать совещательным обработки (Sloman & Logan, 1999), которая могла бы охватить Дункана правило задач экспериментов, планирования и, по крайней мере, элементы, функции центрального исполнительного, менее естественно моделируется с ИНС. Для того, чтобы еще раз подчеркнуть, мы не говорим, что такие познавательные способности не в каком-то фундаментальном смысле быть neurally смоделированы, а мы предлагаем это может быть не самый естественный способ описания.

Как уже отмечались, возможный ответ на эту ситуацию, чтобы рассмотреть вопрос о том, чтобы отображение нейронного субстрата может быть ускорено путем разрыва проблемы моделирования в два этапа: Первое символическое моделирование когнитивных поведения и второе отображение из символических нейронных. С философской точки зрения, можно было бы спорить, что символический посредник более непосредственно отражательной ектов функциональные характеристики познавательной способности в вопросе, что приводит моделист к каким-то образом, более верную модель на когнитивном уровне. Кроме того, предложение было бы, что такая функциональная характеристика также сильно ограничивает нейронную реализацию и наоборот, через прямое отображение в ИНС, т.е. е и далее ectively стадии компиляции (которая, тем не менее, может оказаться необходимой в вычислительной для электронной FFI тивности).

Во-первых, когнитивное поведение символической модели, шаг может очень естественно дать некоторую форму расчетной логической модели, возможно, сформулированный в системах архитектуры производства, такие как SOAR (Newell, 1994), АКТ-Р (Андерсон & Lebiere, 2014) или EPIC (Kieras и др., 1999), или в виде логической программы. В любом случае, второе отображение будет включать некоторую форму вычислительной логики в ИНС преобразование-точно тему нейронного-символический проект. Таким образом, в связи с этим, применение результатов нейронного-символическим исследования может сделать значим вклад в Cognitive Neuroscience.

Тем не менее, такое применение вызывает целый ряд научно-исследовательских тем. Пожалуй, наиболее значимы из них является разработка отображений из вычислительной логики нейрофизиологически правдоподобных нейронных сетей. Это требует рассмотрения более сложных уравнений активации, которая непосредственно модели ионных каналов и динамику мембранного потенциала (например, на основе уравнений Ходжкина-Хаксли (Huxley & Ходжкина, 1952)); он также предлагает подходы, которые лечат возбуждение и торможение отчетливо и использовать более биологически реалистичное обучение, такие как сопоставительный Хебб обучение, так как возникает, скажем, O'Reilly и Munakata в обобщенной Рециркуляции алгоритм. Хороший источник для таких нейронных теорий (O'Reilly & Munakata,

2000).

5. Ввод Машины для работы: Переплет и первого порядка Умозаключение в Neural-Символическое Framework

Несмотря на то, что должно было стать ясно, что кажется правдоподобным предположение, что человеческое познание производит и обрабатывает сложные комбинаторные структуры с использованием нейронных сетей, на компьютерных науках и инженерные систем стороны такого систематическое и динамическое создание этих структур представляют проблемы для ИНСА и теории нейро -cognition в целом (Fodor & Пилишина, 1988; Marcus, 2003; Ван дер Вельде и De Kamps, 2006; Feldman, 2013). Теперь мы будем увеличивать на конкретной работе, связанной с проблемой представления таких сложных комбинаторных структур ИНС-с акцентом на ВОЛП representations- а е FFI ciently вычислительно обучения и их обработки для выполнения когнитивных задач на высоком уровне. Похожи на презентацию NSCA в разделе 3 в качестве примера применения для теоретических и концептуальных соображений из раздела 2, нижеследующее-, кроме представления соответствующего органа работы с нейронной символической интеграции, обеспечивают практическую коллегу и заземления для некоторых из познания и когнитивной нейронауки в центре обсуждения из предыдущего раздела 4.

Конкретно, в этом разделе описан, как логика предикатов могут е F ciently быть представлена и обработано в типе рекуррентной сети с симметричными весами. Подход основан на (переменной) связывания механизма, который кодирует компактно логические выражения Fi первого порядка в активации. Обработка затем осуществляется путем кодирования символических ограничений в весах сети, тем самым заставляя решение к задаче вывода появляться в активации.

5.1 Вычислительного Вид на Фодоре, Пилишина, и вызов (ы) для моделирования познания

Одним из наиболее вычислительно влиятельного и соответствующих положениях, касающихся познания и нейронную-символического вычисления был разработан (Фодором & Пилишина, 1988). Там, две характеристик считаются необходимыми для любой парадигмы, которая направлена на моделировании познания: (I) комбинаторной синтаксис и семантику психических представлений и (б) чувствительность структуры

процессов. Первая характеристика позволяет для представления рекурсивно построено из атомных единиц таким образом, что семантика не-атомной структуры является функцией семантики ее синтаксических компонентов. Вторая характеристика относится к способности основывать применение операций на структурном представлении, на его синтаксической структуре.

Несколько работ частично ответили на их вызов, но никто не мог захватить полную выразительную мощь ВОЛПА: Даже с ограниченной выразительностью, большинство попыток преодолеть Фодора и проблемы Пилишина являются чрезвычайно *localist*, не имеют устойчивости к повреждению нервов, имеют ограниченную способность к обучению, имеют большой размер - сложности, и требуют специальной инженерных сетей, а не использовать общий механизм назначения, который был бы независим от соответствующей базы знаний (KB) (Анандан, Letovsky, и Mjolsness, 1989; Ballard, 1986; Garcez и др 2009; Holldobler, 1990; Holldobler & Kurfess, 1992; Stolcke, 1989; Ван дер Вельде и De Kamps, 2006). Кроме того, лишь немногие системы использовали распределены представления и даже те, которые делают обычно Су Ф.Ф. эр от потери информации и мало возможностей обработки (Plate, 1995; Поллак, 1990; Stewart & Eliasmith, 2008).

5.2 Связывание Проблема вычислительном Revisited

Несмотря на большой объем работы (не исчерпывающе) обобщены в разделе 4 и соответствующие многочисленные попытки приблизиться критику Фодора и Пилишина'S, вычислительная нейронная моделирование познавательных задач высокого уровня, и в особенности обработки языка по-прежнему считается жестким вызовом. Многие из проблем при моделировании принимаются очень вероятно, связано с уже описанным ранее «связывания проблемы». После де определению пути (Feldman, 2013), короче общее связывание проблема касается того, как элементы, которые кодируются в различных схемах с массовым параллелизмом вычислительного устройства, могут быть объединены в сложных способах, для различных когнитивный задач. В частности, связывание становится чрезвычайно сложным в обработке языка и абстрактного мышления (Barrett, Фельдман, и Dermed, 2008; Jackendor Ф.Ф., 2002; Маркус, 2003; Ван дер Вельде и De Kamps, 2006), где необходимо иметь дело с более требовательными , особый случай: переменные связывания. Основной вопрос сужается до: как простых составных частей, склеенных к переменным, а затем использовать несколько отдельных частей и из массивно параллельной системы, в ди и след роли различны? После (Jackendor фф,

2002), связывание проблема может быть дополнительно разложить на четыре конститутивную задачу: (I) массивность связывания проблемы, (б) проблема нескольких экземпляров факта, (III) проблема связывания значений переменных и (IV) соотношение между креплениями в рабочей памяти кратковременной (WM), в отличие от привязок в долговременной памяти (LTM). Когда совместно глядя на все четыре, механизм связывания, как представляется, основное препятствие, поскольку каждый отдельный элемент зависит от него. Но даже если механизм связывания должен существовать и дает представление сложных структур в нейронных подобных единицах, не ясно, как они обрабатываются, чтобы достичь цели и как процедурные знания, необходимые для такой обработки инкапсулируются в синапсах. Конкретно, в вычислительном случае, когда FOL выражение должно быть обработано с использованием только нейронные единицы с двоичными значениями, как должны быть закодированы соответствующие формулы? Как нейронные-ансамбли, представляющие собой простые составляющие, склеиваются, образуя сложные структуры? И как эти структуры манипулируют для универсального фантастического и в конечном счете для рассуждений?

В качестве шага первой, механизм связывания требуется, что позволит «клей» вместе элементы в KB, таких как предикаты, функции, константы и переменные, а затем применить его

в манипулировании этих структур. Соответственно, несколько попыток было сделано для решения переменной связывания проблемы (Анандан и др., 1989; Barrett и др, 2008; Browne & Sun, 1999; Шастри & Ajanagadde, 1993; Ван дер Вельде и De Kamps, 2006). Но в то время как соответствующие подходы внесли большой вклад в наше понимание возможных механизмов на работе, оказались, что каждый из них по-прежнему имеет означающий муравей недостатки, связанный с ограниченной выразительностью, высокими требованиями к пространству, или требований центрального управления.

5.3 фи катионы Умозаключение Конкретных в виде фиксированные точки ИНСА

Вместо разработки снизу вверх, модульное решение для выполнения рассуждения с ИНС, давайте рассмотрим противоположный подход: определение того, что должно быть возможными результатами процесса рассуждения как фиксированные точки (или устойчивыми состояниями) ИНС. В каждом ИНСЕ, что достигает равновесия, сходящееся на стабильное состояние (вместо осциллирующего), то фиксированы точек сети данных зажатые входов-может быть связана с решениями для рассматриваемой проблемы. В частности, ИНС с симметричной матрицей весов, такими как, например, в уже обсуждались Больцмановские машины из раздела 3, выполняют градиентный спуск в функции энергии, чья глобальные минимумы может быть связаны с решением проблемы. Хотя симметричные (минимизация энергии) сети являются простым выбором архитектуры, может быть использована и другая (несимметричная) архитектура, пока устойчивые состояния возникают и те устойчивые состояния соответствуют решениям задачи под руку. При использовании симметричных сетей, общая стратегия связать проблемно-решения с глобальными минимумами уравнения энергии, соответствующей сети. В частности, как «проблемы» является ВОЛП умозаключение, обоснованное умозаключение цепи являются решениями, которые отображаются в глобальные минимумы функции энергии.

Было показано, что симметричный ИНС высокого порядка (с мультипликативными синапсами) эквивалентны «стандартным» симметричного ИНС (с попарно синапсами) со скрытыми единицами (Пинкасова, 1991b). В соответствующих имитациях, варианты высокого порядка Больцмановских машин поэтому используются, так как они быстрее моделировать и имеют меньшее поисковое пространство, чем «стандартную» машину Больцмана со скрытыми единицами и только попарными соединениями. Декларативные спецификации (ограничения), характеризующие действительные решения составляются (или уроков) в весах высокого порядка Больцман машины. Первый порядок КБ либо компилируется в весах, или она зажимается на активацию некоторых из видимых нейронов. Запрос, краткосрочные факты, и / или дополнительные сигналы также могут быть закреплены на WM. После того, как КБ находится на своем месте (хранятся в виде долгосрочных весов) и запросы или сигналы накладываются на WM, ИНС начинает выполнять его градиент алгоритма спуска, поиск решения для задачи логического вывода, который удовлетворяет ограничения, сохраненные в качестве весов. Когда сеть оседает на глобальный минимум, такое стабильное состояние интерпретируется как ФОЛ цепь вывода, доказывающие запросов с использованием шагов логического вывода разрешения на основе (Пинкасова, 1991a; Lima, 2000; Пинкас, Лима, & Cohen, 2012, 2013).

5.3.1 ИНС Спецификации Использование Весовые логических ограничений

Первый шаг в этом процессе является спецификацией задачи (например, ФОЛ логического вывода) в виде набора взвешенных пропозициональных формул логики. Эти формулы представляют собой логические ограничения, которые обеспечивают выполнение нейронов, чтобы получить ноль / один назначение, которые точно соответствуют действующим решениям ВОЛПА вывода.

Besold и др.

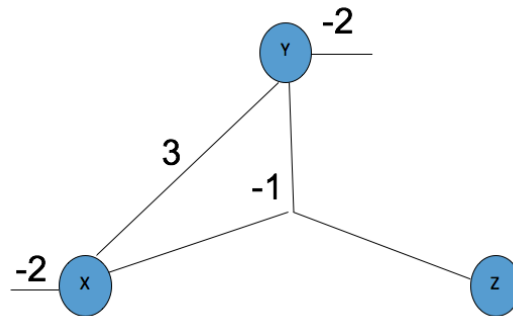


Рисунок 2: Симметричная сеть 3-го порядка, характеризуются функцией энергии 3-го порядка: $XYZ - 3XY + 2X + 2Y$. Минимумы этой функции энергии фиксированы точки сети. Эта сеть выполняет поиск решений для удовлетворения взвешенной конъюнктивной нормальной формы (УТС): $(\neg I_{кс} \vee \neg Y \vee Z) \wedge (I_{кс} \vee Y)$. Обратите внимание, что пункты весового CNF дополнены штрафами что отражает важность каждого ограничения.

Удовлетворяя пропозициональную формула представляет собой процесс присвоения значений истинности (истина / ложь) к атомным переменным формулы, таким образом, что формула оказывается правдой. Предельно удовлетворяющий взвешенного множество таких формул находки истинность значение, которые минимизируют сумму весов нарушенных формул. Ассоциируя истинностное значение истинного с 1 и ложное с 0, квадратичная функция энергии создается, которая в основном вычисляет взвешенную сумму нарушенных ограничений. Сочетания выражены в виде мультипликативных терминов, дизъюнкции как мощность объединения, и отрицание выражены в виде вычитания из

1. Для того, чтобы намекает сеть о том, как близко это к раствору, конъюнкция весовых формул рассматриваются как сумма весов булевых ограничений, которые не удовлетворяются под ред. Можно связать проблему максимальных взвешенным удовлетворяться способностями исходного набора весовых формул с находкой глобального минимума для уравнения энергии, которое соответствует задаче спецификации, (Пинкас, 1991b). Другими слова, не удовлетворяющий пункта будет нести увеличение энергии системы. Один интересная и мощная стороны е и далее ЭСТ такого отображения является возможностью частично удовлетворяющим иначе unsatisfiable способного набор формул. Это явление используются для определения оптимальных и предпочтительных решений, как в: находке кратчайшего объяснения или доказательства (скупость), находка наиболее общих Уни катионов F_i , находка наиболее вероятный вывод или наиболее вероятное объяснение и т.д. Когда уровни различны ди ФФ весов связаны с формулами в КБ, немонотонной механизм логического вывода получается на основании Пенальти логики. В самом деле, в (Пинкасе, 1995) эквивалентности была показана между минимизацией энергии симметричного ИНСА и удовлетворяться способностями одиннадцатиметровый логики, то есть, каждый штраф логической формула может быть е FFI ciently переводятся в симметричный ИНС и наоборот. На рисунке 2 приведен пример для сетевого экземпляра задачи удовлетворяться способности.

На основании этих результатов, компилятор был построен, который принимает решения спецификаций как форма индексных количественно ц и изда штрафных логических формул и генерирует ИНС, которые стремятся к максимально удовлетворяющим решениям Конкретной ред формул. Этот компилятор может быть использован для создания ИНС для различных поиска ВОЛП, уни фи катиона и вывода проблем, с участием тысяч нервных узлов и десятки тысяч синапсов. В качестве альтернативы для компиляции,

эта спецификация ограничение также может быть извлечено (неконтролируемым) с использованием анти-Хебб обучения, когда точка неподвижной обнаруживаются, что не удовлетворяют некоторые ограничения (Пинкасова, 1995).

5.3.2 ИНС Характеристики ВОЛП Умозаключение Цепи

Есть два возможных подход к логическим рассуждениям, тот, который использует модель теорию и другую теорию доказательств. Неформально, теория модели подтверждает, что заявление может быть выведенные путем проверки всех возможных моделей, которые удовлетворяют КБ, в то время как доказательство теории применяет меры логического вывода для создания доказательства заявления, о котором идет речь. Большинство попыток делать FOL выводы в ИНС-или с использованием пропозициональную удовлетворяться способности в целом, используют modeltheoretic методы, основанные на заземленных KBS (Clarke, bière, Рэйми, и Zhu., 2001; Garcez и др., 2002; Domingos, 2008; Holldobler & Kurfess, 1992). Тем не менее, результаты заземления в любом экспоненциальном взрыве числа булевых переменных (нейроны) или жестких ограничения к выразительной силе языка ВОЛПА (например, нет функциональных символов и не экзистенциальных кванторов \exists ERS). Методы доказательства теории можно использовать незаземленные FOL формулы и являются более электронной FFI Client в размере. Другим следствием того, что доказательство для оператора запроса является то, что это доказательство можно рассматривать как объяснение этого заявления, что не всегда можно получить при использовании теоретико-модельные методы. Доказательство теоретико-подход, например, используется (Пинкас, 1991a) и позволяет для различных правил вывода (например, разрешения, Modus Ponens и т.д.) либо с зажимает КБ или с КБ хранится в весах. В (Лима, 2000), теорема ФОЛ пружер была изучена с разрешением и базой знаний, которая зажимается полностью в активации. В (Пинкасовой и др., 2013) Наиболее общий унифицированный с проверкой вхождения осуществляется с использованием более поздним связыванием механизма с пониженной сложностью размера.

После того, как запрос и / или сигналы были зажаты на WM, сеть настроена для запуска до тех пор, пока раствор не выступает в качестве глобального минимума функции энергии сети. Это устойчивое состояние может быть интерпретировано как доказательство или цепочка FOL положений. Каждый пункт получают либо путем извлечения его из LTM, путем копирования его из предыдущих стадий (необязательно с коллапсу), либо путем разделения двух предыдущих положений в цепочке. На каждом этапе, необходимость унифицировать разрешенные литералы может привести к литералам, чтобы стать более специфичным с переменным связыванием. Можно построить ИНС, который ищет доказательства по опровержению (полная стратегия, производное противоречие, когда инвертированный запрос добавляется в КБ) или непосредственно вывести заявление. Непосредственно выводя заключение, хотя и не полная стратегия, имеет больше сходства с человеческим разумом, и кажется более правдоподобным познавательно.

Вариации выше основного использование механизма логического вывода конкуренции между доказательствами и могут быть использованы для выведения наиболее вероятного объяснения для запроса, наиболее вероятное объяснение, наиболее вероятно, логического вывода, наименее дорогостоящий план или наиболее предпочтительные действий (как в роботизированных законах Азимов).

5.4 отказоустойчивого Механизм динамического связывания

Компактный механизм связывания лежит в основе описанного кодирования ВОЛПА и логический вывод двигателя общего назначения. Техника связывания является основным механизмом, на котором строится WM и именно этот механизм, который позволяет извлекать необходимые знания из LTM и затем унифицированный фантастического и умозаключений.

В отличие от височного связывания (Шастри & Ajjanagadde, 1993), он использует пространственное (конъюнктивный) связывание, которое захватывает композиционность в стабильном состоянии нейронов и не чувствителен к временной синхронизации или временной интервал распределения. Хотя были предложены несколько присоединительных (не временные) вяжущие механизмов, один в (Пинкасова и др., 2012) компактно используют только несколько ансамблей нейронов, каждый из которых может связывать любого из многих элементов знаний, хранящихся в LTM. Это свойство имеет три боковые e Ff ЕКТСА: (I) снижение числа нейронов (единицы), так как связующие вещества только необходимо для представления доказательств (а не захватить все КБ), (II) отказоустойчивость, так как каждый связывающий ансамбль общее назначение и может занять место любого другого неудачного связывания ансамбля, и, наконец, (III) высокая выразительная мощность, так как любая меченный направлена структура графа может быть представлена этим механизмом (ВОЛП как частный случай).

Эта привязка применяется метод ригели нейронов связать FOL объектов (разделы, литералы, предикаты, функцию и константу), используя специальные нейронные ансамбли под названием общего назначения (GP) связующие вещества. Эти GP-связующие действуют в качестве косвенных указателей, то есть два объекта связаны тогда и только тогда, когда через GP-связующее указывает на них обоих. Так как GP-связующие могут указывать на другие GP-связующие, вложенная иерархия объектов может быть представлена с использованием таких связующих веществ, указывающие, в то же время к объектам / концепциям и другим связующим. В самом деле, произвольные оргграфы могут быть представлены и, как частные случаи, вложенный друг в друг точка, литералы FOL, статьи, и даже полные доказательства разрешения на основе (Пинкасова, 1991а; Лима, 2000). Манипулирование таких сложных представлений осуществляется с помощью взвешенных булевых ограничений (составитель или изученные в синапсах), которые вызывают GP-связующие быть динамически распределены ВОЛС объекты, извлеченные из LTM. Так как GP-связующие могут быть использованы для связывания любых двух или более объектов, количество связующих веществ, которые на самом деле необходимы пропорциональна размеру доказательства. В результате действительно необходимы только несколько таких связующих веществ для познавательного-правдоподобного умозаключения цепи, что приводит к сложности размер WM быть значительно уменьшен до $N \cdot \log(N)$ где N

доказательство длины и K это размер КБ (Пинкасова и др., 2013). Интересно, что отказоустойчивость этой WM ВОЛП является результатом общего характера связующих веществ, что связано с избыточностью в синапсах, а не к избыточности в единицах. GP-связующие случаться для решения большинства задач, заявленных (Jackendor Ф.Ф., 2002) и обобщенно выше.

5.5 Примеров умозаключений, кия, и дедуктивные ассоциации

Для того, чтобы проиллюстрировать GP-связующий подход, предположим, КБ, состоящий из следующих правил:

- Для каждого i , существует мать U который также является родителем i : *Родитель (мать (i), i)*
- Для всех x, y, r : *Родитель (x, y) ∧ Родитель (y, r) → Grandparent (x, r)*
- Больше детали КБ: факты и / или правила, возможно, дополненные веса

Общее назначение разрешения на основе логического вывода и WM реализован в вариации машин Больцмана компилируются (или научились) с указанным выше КБ, хранящимся в ее синаптических соединениях. При зажатии запроса: *Grandparent (V, Энн)* на некоторые из видимых блоков в WM, при решении на глобальном энергетическом минимуме, сеть способна связывать переменную V в перспективе, представляющий «мать матери Анны» и находка полного логического вывода цепи в качестве доказательства для этого запроса:

1. Получение и объединяющее правило 1: *Родитель (мать (Аппе), Энн)*
2. Получение и объединяющее правило 2: \neg *Родитель (мать (мать (Аппе)), мать (Анна))* \vee
 \neg *Родитель (мать (Аппе), Энн) \vee GrandParent (мать (мать (Аппе)), Энн)*
3. Решение 1 и 2: \neg *Родитель (мать (мать (Аппе)), мать (Анна))* \vee
GrandParent (мать (мать (Аппе)), Энн)
4. Получение правило 1 раз, но с ди далее различны уни фи катиона (несколько экземпляров):
Родитель (мать (мать (Аппе)), мать (Анна))
5. КЭД путем разрешения 3 и 4: *GrandParent (мать (мать (Аппе)), Энн)*

FOL положение, литералы, атомы, и термины представлены в виде графиков, направленных ациклических в то время как GP-связующие вещества выступают в качестве узлов и помеченных дуг. GP-связующие динамически распределены без центрального контроля, и только необходимые правила извлекаются из КБ, для того, чтобы вывести запрос (цель). Другие правила, которые хранятся в КБ (в LTM) не получены в WM, потому что они не нужны для вывода.

Запрос может рассматриваться как своего рода кием в WM, которая направляет в сети «внимание» к поиску для прародителей Анны; тем не менее, другие сигналы могут быть предоставлены вместо или в дополнение к запросу, например, мы можем ускорить сходимость системы к правильному решению, указав в СУС, что Правило 1 и / или правило 2 следует использовать, или, еще лучше, путем зажима этих правил на активацию WM до начала работы в сети.

Этот механизм логического вывода ANN также может быть истолкован как интеллектуальная ассоциативной память, поскольку сигналы генерировать цепочку шагов дедукции, что приводит к сложным ассоциациям. Например, вместо запроса, если только «Аппе» и «матери» зажаты в репликах в WM это составляет попросив систему, чтобы вывести что-то о человеке, «Энн» и понятие «матерях». В этом случае система будет сходиться к одному из нескольких возможных выводов (каждый с известной вероятностью). Например, система может фи обнаружит, что Энн имеют родитель, который является ее матерью. С другой стороны, это может фи обнаружит, что Энн имеют прародитель (мать ее матери), или, что мать Анны есть мать и что мать тоже есть мать. Вероятность возможных возникающих ассоциаций определяются такими факторами, как веса LTM (т.е. штрафов из пунктов ВОЛПА КВ, хранящихся в синапсах) и глубина вывода цепей.

5.6 Может Нейроны логично? Основные отрицательные и положительные результаты

Только что описано работа свидетельствует о том, что неограниченное (но память переплета) умозаключение ФОЛА может быть представлены и обработано относительно компактным, аттрактора на основе ИНСА таких, как ранее введенные машины Больцмана. Тем не менее, в (Пинкасова & Dechter, 1995) также было показано, что нейронные механизмы распределены-среди которых также машины Больцмана кол-могут иметь патологические колебания в ИНСЕ, которые имеют циклы и может никогда не обнаруживает глобальное решения. Связанный положительный результат в том, что, когда ANN не имеет циклов, можно иметь нейронную функцию активации, которая гарантирует находку глобального решения. После того, как цикл добавляется к сети, ни одна функция активации не способна устранить все патологические сценарии выше. К сожалению, описанные FOL сети имеют много циклов и поэтому обречены быть неполными, даже с границей на доказательстве

длина. К счастью, с некоторыми правдоподобными предположениями о демоне, можно гарантировать, что такие патологические сценарии не будут длиться вечно. Другим серьезным препятствием для достижения хорошей производительности является то, что Больцмановские машины и другие симметричные сети используют стохастические локальные алгоритмы поиска для поиска глобального минимума их энергетической функции. Теоретически и практически, локальные минимумы могут досаждают пространство поиска, поэтому такие ИНСЫ не гарантируется Чтобы найти решение в конечное время. В настоящее время исследование пытается решить эту проблему, используя множество вариаций Хебба методов обучения и вариации машин Больцмана изменить синапсы ИНСА таким образом, что локальные минимумы устранены и глобальный минимум найден быстрее. Как выясняется, недавний успех глубокого обучения в основном возникает также из одних и тех же изменений, но с более практической точки зрения. Изучение стилей этого типа называются «скорость практиковать», как они стремятся использовать обучение в качестве стратегии для самосовершенствования.

6. Коннекционистский первого порядка Логика обучения Использование метода сердечника

Как уже стало ясно, в предыдущих разделах, одна из центральных проблем, лежащие в основе нейронной-символической интеграции является представление и изучение фи правила первого порядка (то есть реляционные знания) в коннекционистской обстановке. Но есть вторая мотивация к тому же стимулы выходят из когнитивной науки и когнитивной нейронауки и направленных на создание моделей и (пере) познавательных процессов высокого уровня в правдоподобной нервной обстановке: Цель создания новой, neurally-вдохновленное вычислительная платформа для символьных вычислений, которые, помимо всего прочего, будут массово распространены без центрального управления, надежного удельных неудач и самосовершенствующегося со временем.

В этом и следующем разделе мы возьмем более техническую точку зрения-в определенную степень противопоставлении сильно познавательно мотивированные разделы 4 и 5-и сосредоточиться на два прототипических примера для работы по информатике стороны нервной-символической интеграции, а именно (я) приближение **непосредственного оператора следствия T_L связанный с программой ВОЛП π в нейронную контексте обучения, и (II)** марковской логики в качестве вероятностного расширения ВОЛП сочетающей логики и графических моделей. В то время как понижая разработки по последней теме в следующем раздел 7, мы начинаем с вопросом оператора следствия.

Предположим, что программа FOL, содержащую тот факт, $p(X)$. Применяя Ассоциированные T_L -Оператор один раз (для произвольной начальной интерпретации), приводит к результату, содержащему в бесконечно много атомов, а именно: все $p(X)$ - Атомы для каждого $Икс$ в пределах, лежащей в основе Эрбрана вселенной U_L .

Таким образом, необходимо аппроксимировать оператору, потому что даже одно приложение может привести к результатам в конечных. И хотя в этом простом примере можно было бы все еще быть в состоянии Определит представление конечного, проблема может стать произвольным комплексом для других программ. Так называемые рациональные модели были разработаны для решения этой проблемы (Bornscheuer, 1996).

К сожалению, нет никакого способа, чтобы вычислить верхнюю границу на размер результирующего рационального представления. Потому что мы не знаем любого другого представления конечным, то мы будем в следующем концентрате на стандартном представлении с использованием Эрбрановы интерпретации. В принципе, есть два варианта аппроксимации данного оператора. С одной стороны, можно развить аппроксимирующее Коннекционистское представление для заданной точности. Это приближение в пространстве приводит к увеличению числа скрытых блоков слоя в результате сетей (Bader, 2009a; Bader, Hitzler, Holldobler, & Witzel, 2007). В качестве альтернативы,

коннекционистская система может быть построена, что приближает однократный применение T_n тем лучше, чем дольше он работает. Это приближение времени было использовано в (Bader & Hitzler, 2004) и (Bader, Hitzler, и Garcez, 2005). Сети, построенные по приближению в пространстве, являются более или менее стандартной архитектуры с большим количеством единиц скрытых слоев, в то время как другие, как правило, очень небольшой, но использование нестандартных архитектур и устройств.

6.1 Целесообразность первого порядка Основной метод и Погружение первого порядка правил в нейронных сетях

Хорошо известно, что многослойные опережающие сети являются универсальными аппроксиматорами для всех непрерывных функций на компактных подмножествах (Фунахасье, 1989). Если подходящий способ представления функции интерпретации первого порядка (от конечных векторов) вещественных чисел можно найти, то прямая связь сеть может быть использована для аппроксимации функции смысла таких программ. Необходимо, чтобы такое представление совместимо с обоими, логика программирования и парадигмы ANN. Для этого, один де фи определяют гомеоморфное вложение из пространства интерпретаций в компактное подмножество действительных чисел. (Holldobler, Kalinke, & Storr, 1999) показала наличие соответствующей метрики гомеоморфизма, и (Hitzler, Holldobler, & Седа, 2004) впоследствии расширила этот результат на характеристику топологического изоморфизма. После этих прозрений, можно использовать отображение уровня для реализации этого вложения. Для некоторой интерпретации Эрбрана \mathcal{A} , отображение уровня биективен $|\cdot| : B_L \rightarrow \mathbb{N}^+$,

а также $b > 2$, функция вложения $\eta : B_L \rightarrow \mathbb{R}$ и ее расширение $\eta : I_L \rightarrow \mathbb{R}$ которые определяются как следующим образом:

$$\eta : B_L \rightarrow \mathbb{R} : \gamma \rightarrow b \cdot |\gamma| \quad \eta : I_L \rightarrow \mathbb{R} : \gamma \rightarrow \sum_{A \in \mathcal{A}} \eta(A)$$

Вот, $S_b = \{\eta(I) \mid I \in L\} \subset \mathbb{R}$ используется для обозначения множества всех встроенных интерпретаций. Это приводит к «бинарному» представлению данной интерпретации \mathcal{A} в системе счисления с основанием b . После того, как вложение интерпретации и глядя на ее представление в рамках этой системы счисления один делается вывод A в каждой позиции i для всех $\mathcal{A} \in \mathcal{A}$. Как уже упоминалось выше, вложение требуется гомеоморфными (т.е. быть непрерывным, взаимно однозначным и иметь непрерывный обратный), потому что, будучи гомеоморфно гарантирует, что η по крайней мере перспективный кандидат, чтобы преодолеть разрыв между логическими программами и коннекционистскими сетями. Можно построить вещественную версию e_n непосредственного оператор следствия T_n следующим образом:

$$e_n : S_b \rightarrow S_b : \text{Икс } \gamma \rightarrow \eta(T_n(\eta^{-1}(\text{ИКС})))$$

Потому как η гомеоморфизм, e_n это правильное вложение T_n в том смысле, что сохраняется вся структурная информация. Более того, η как известно, быть непрерывным, что позволяет сделать вывод о том, что e_n непрерывна для непрерывного T_n . На основе этих прозрений, можно построить Коннекционистские сети для аппроксимации T_n . Как уже упоминалось выше, есть в принципе два варианта аппроксимации в пространстве и во времени. вложение η

введенная выше позволяет одновременно, как описано подробно в (Bader, Hitzler, & Witzel, 2005; Бадер. и др, 2007; Bader & Hitzler., 2004; Бадер и др, 2005). В (Bader, 2009b), методы были описаны в полной детализации для построения стандартных сигмоидальных сетей, а также радиальную базисной функция сети для любой заданной точности аппроксимации. То есть, если

интерпретация y подается в сети, в результате чего выходной сигнал может быть интерпретирован как интерпретация J и расстояние между J а также $T_L(J)$ ограничена заданной ,

До сих пор мы имели дело с приближением однократного применения. Прямое следствие только оператора. Более интересный является то, сходится ли повторное применение аппроксимирующей функции в точку фиксированной, относящейся к модели базовой программы. Конвергенция можно показать, показывая, что функция является сжимающим и за счет использования принципа сжимающих отображений Банаха. К сожалению, получается, что аппроксимирующие функции, описанные выше, не сжимающие на p в целом, но они сжимающие, если ограничены C . Таким образом, можно сделать вывод, что не только функция смысловой T_L может быть аппроксимирована с помощью Коннекционистской сети, но и итерации приближения сходится к состоянию, соответствующему модели из основной программы при условии, что T_L сам по себе является сжимающей.

6.2 Обучение в Настройка первого порядка

В пропозициональной установке, обучение было показано, что возможно (Garcez и др., 2002) и иметь преимущество (Bader, Holldobler, & Marques, 2008). Как и в случае пропозициональном, подготовка данных используются для адаптации весов и структуры случайно инициализированных сетей. Подход построен на основе вектора сетей. Первый был предложен в (Bader и др., 2007), а также с конкретно с учетом метода обучения. Для того, чтобы показать применимость этого подхода, сеть была случайно отформатирована и пары интерпретаций и соответствующих последствий (т.е. пара y а также $T_L(J)$), использовались в качестве обучающей выборки. После сходимости процесса обучения сеть была оценена путем подачи случайных интерпретаций в сеть и итерации вычисления путем подачи выходных сигналов обратно во входной слой. была найдена. Полученная последовательность интерпретаций сходится к модели базовой программы P . Отсюда можно сделать вывод о том, что сеть действительно удалось узнать непосредственным следствием оператора su FFI довольно точным.

С более практической точки зрения (признавая выше ди FFI culties, но признавая их Научные, репрезентативную ценность), некоторые из них выбрали для подключения нейронных вычислений с работой, проделанной в индуктивной логике программировании (Muggleton & de Raedt, 1994; Базилио, Zaverucha, и Барбос, 2001). Первоначальные результаты, что воспользоваться методами ILP для электронной FFI Cient ВОЛП обучения под названием propositionalisation (Kroegel, Rawles, Zelezny, Флэч, Lavrac, и Врубель, 2003), только начали появляться, но с многообещающими результатами (см, среди прочего, Коннекционистская система ИЛПА по формуле (Франк, Zaverucha, & Garcez, 2014), или работы (Šourek, Aschenbrenner, Železný, & Kuzelka, 2015; de Raedt, Керстинг, Натараяно, & Poole, 2016)). Кроме того, в рамках более практического подхода, расширение выше фиксированном-точка метод к нескольким сетям, как показано на рисунке 1, было предложено в (Garcez и др.,

2009). Такое расширение, называется коннекционистского модальная логика, представляет интерес здесь по отношению к FFI обучения, потому что он доказал способность представлять и обучение пропозициональные модальные логики, которые, в свою очередь, эквивалентны фрагменту ВОЛП с двумя переменными. Логика высказываний модальной была показана, робастно разрешимы (Варди, 1996) и его можно рассматривать, как результат, кандидат на о т ч и далее адекватной середине между требованиями способности представления и изучением электронной FFI тивности.

английский	Первый порядок Логика	вес
Курение вызывает рак	$\forall x: Sm(x) \rightarrow Ca(x)$	1,5
Друзья разделяют привычки к курению $\forall Икс$	$\forall y: PT(x, y) \wedge Sm(x) \rightarrow Sm(y)$	1,1

Таблица 2

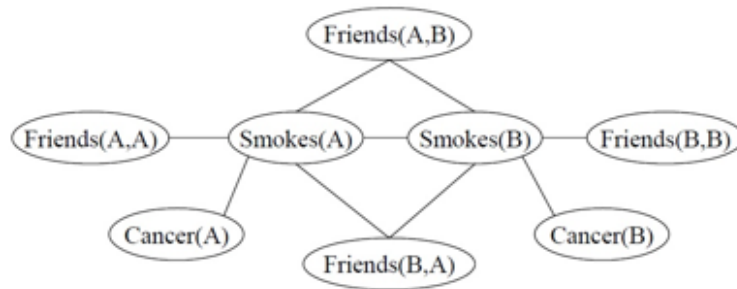


Рисунок 3: Сеть Маркова с двумя константами Анна (A) и Боб (B).

7. Марковские логические сети Комбинирование Вероятности и первого порядка логики

В качестве второго примера прототипе для более технической стороны нервной-символической интеграции помимо коннекционистского первого порядка применения обучения в предыдущем разделе, мы обратим наше внимание на логических сети Маркова. Маркова логика является вероятностным продолжением ВОЛП (Richardson & Домингос, 2006; Домингос & Lowd, 2009). Классическая логика может компактно моделировать зависимости высокого порядка. Тем не менее, в то же время он является детерминированным и, следовательно, хрупким в присутствии шума и неопределенности. Основная идея марковской логики преодолеть эту хрупкость путем объединения логики с графическими моделями. С этой целью, логические формулы размягчаются с весами, а ϕ и далее активно служат в качестве шаблонов функций для сети Маркова или марковского случайного поля (MRF) (Taskar, Chatalbashev, & Koller, 2004),

т.е. неориентированный графическая модель, которая позволяет произвольные потенциалы высокого порядка, аналогичные (но потенциально гораздо более компактным, чем) высокого порядка распределения Больцмана.

В частности, логика сеть Маркова (MLN) представляет собой набор весовых формул FOL. Вместе с набором констант, что ϕ определяет марковской сети с одним узлом на атом земли (предиката со всех переменных заменены на константу) и одну особенность на первом формуле. Вес особенностью является вес формулы первого порядка, который произошел его. Вероятность состояния $Икс$ В такой сети задается $P(x) = \frac{1}{Z} \exp(\sum_{\phi} \phi(x) \cdot \phi)$, где Z

постоянная нормировки, ϕ это вес ϕ -й формула, $\phi = 1$, если ϕ -й формула верна, и $\phi = 0$ в противном случае.

В качестве примера, таблица 2 показывает MLN с двумя весовыми формулами, а на рисунке 3 показывает соответствующую сеть Маркова с двумя константами, Анна (A) и Боб (B).

Точные и приближенные методы вывода для MRFs и других графических моделей также могут быть применены к MLNs. Популярными вариантами включают переменные элиминации или клеммные деревья для точного вывода и передачи сообщений или цепь Маркова метода Монте-Карло для приближенных умозаключений. Однако, по сравнению со стандартными MRFs, MLNs ввести новые вызовы и

возможности для электронной FFI фективных рассуждений. Одна задача состоит в том, что даже скромно таком размере домены могут дать исключительно большие и сложные сети заземления Маркова. Например, в области социальной сети, такие как таблицы 2, количество дружбы квадратично растет с числом людей, так как любая пара людей может быть друзьями. Хуже того, простые такие понятия, как транзитивность (друзья ваших друзей могут быть ваши друзья, а) приводят к кубическому числу основных формул. Таким образом, домен 1000 человека имеет 1 миллион возможной дружбы и 1 миллиард возможных транзитивности отношений. Стандартные методы логического вывода не очень хорошо масштабироваться до этих размеров, особенно когда формула веса являются большими, представляющие сильные взаимодействия между переменными. Возможность, что эти большие модели имеют большую симметрию, структуру, которая может быть использована «поднимаемым» алгоритмы логического вывода (см (Браз, Amir, & Roth, 2005; Singla & Domingos, 2008)). Отменено умозаключение работает путем группирования переменных или переменные конфигураций подтвер- вместе, что приводит к более простым задачам логического вывода. Например, когда нет никаких доказательств того, ди и далее erentiate Анны и Боб, вероятность того, что каждые дымы должны быть одинаковыми. В зависимости от структуры МЛН и доказательств, подняли вывод можно вычислить точные или приблизительные вероятности более миллионов переменных без инстанцирования или рассуждений о всей сети. Другой метод масштабирования умозаключение, чтобы сосредоточиться на только подмножество сети, необходимой для ответа на конкретный запрос или набор запросов (Riedel, 2008). Связанные методы включают в себя ленивым логического вывода (Singla & Domingos, 2006; Пун, Domingos, & Самнер, 2008), которая только инстанцирует необходимую часть сети, и крупнозернистого ТО- умозаключение тон- (Kiddon & Domingos, 2011), который использует более сложные модели повторно определяют результаты более простых моделей.

Значения параметров в качестве МЛН имеют решающее значение для его e и следа $ectiveness$. Для некоторых приложений, просто установив относительные веса $ди$ различны и след формул $су FFI Cient$; например, специалист в предметной области может быть в состоянии идентифицировать некоторые формулы как жесткие ограничения, другие, как сильные ограничения, и другие, как слабые ограничения. Во многих случаях, однако, наилучшая производительность достигается путем тщательного настраивая их к $фи T$ ранее наблюдаемые данные обучения. Как и с другими вероятностными моделями, обучение веса МЛН основано на максимизацию (штрафуется) вероятности обучающих данных. Учитывая полностью Наблюдаемые подготовки данных, это выпуклая задача оптимизации, которая может быть решена с помощью градиентного спуска или методов оптимизации второго порядка. Тем не менее, вычисление градиента требует запуска логического вывода в модели для вычисления ожидаемого числа удовлетворяться изд заземлений каждой формулы. Кроме того, градиент часто очень плохо обусловленной, так как некоторые формулы имеют гораздо больше удовлетворительных мели, чем другие. Стандартное решение состоит в аппроксимации градиента с коротким пробегом цепь Маркова Монт-Карло, аналогичен стойкими контрастными методами дивергенции, используемых для подготовки ограниченных Больцмана машин (Тилеман, 2008), а также обуславливать градиент путем деления ее дисперсией в каждом измерении (Lowd & Домингуш, 2007). Популярная альтернатива оптимизация $pseudolikelihood$ (Besag, 1975), который не требует работает умозаключения, но может обрабатывать длинные цепи доказательств плохо.

Совсем недавно, новая парадигма возникла путь тесной интеграции обучения с умозаключениями, так как $exempli$ $фи$ ред сетями суммарной продукции (СПБАЯ) (Пуны & Домингосом, 2011). Развитие SPNs мотивировано ищут общие условия, при которых функция распределения является послушной. SPNs направлены ациклические графики с переменным листьев, суммами и продуктами в качестве внутренних узлов и взвешенных ребер. На рисунке 4а показан пример SPN, реализующие распределительную дерево с кластерами (*Икс 1* и *Икс 2*) и (*Икс 1* и *Икс 3*) и *separa-*

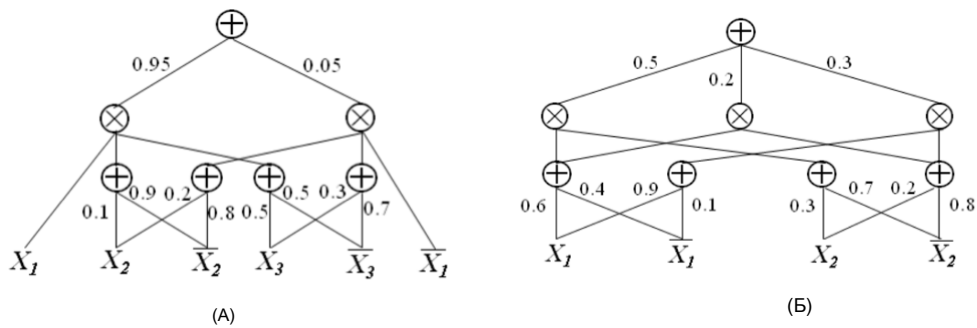


Рисунок 4: (а) Пример реализации SPN, распределительное дерева кластеров (*Икс 1* и *Икс 2*) и (*Икс 1* и *Икс 3*) и сепаратор *Икс 1*. (б) Наивная модель Байеса с переменными *Икс 1* и *Икс 2* и три компонента.

тор *Икс 1*, и 4б наивной модели Байеса с переменными *Икс 1* и *Икс 2* и три компонента.

Для того, чтобы вычислить вероятность частичной или полной переменной конфигурацией, устанавливают значения переменных листа и вычислить значение каждой суммы и произведения узла, начиная с нижней части СПН. Например, для вычисления $P(X_1 = \text{правда}, \text{Икс}3 = \text{ложь})$ в SPN на рисунке 4а, установить значение *Икс 1* а также *Икс 3* 0, так как эти значения переменных противоречат Специфическим-е изд конфигурации и установить все другие листья до 1. Каждого из четырех нижних узлов суммы является взвешенной суммой двух значений листа; слева направо, они оценивают в 1, 1, 0,5, и

0,3. Узлы продукта над ними оценки 0,5 и 0,3, а общий корень равных

$0,95 \cdot 0,5 + 0,05 \cdot 0,3 = 0,49$. Таким образом, вывод следует структуре SPN, позволяя произвольные запросы вероятности дать ответ в линейное время.

Было показано, что если SPN удовлетворяет очень общие условия называются полнотой и последовательностью, то она представляет статсумма и все маргинальные некоторую графическая модель. По существу все податливые графические модели могут быть отлиты в SPNs, но SPNs также строго более общий характер. SPNs достичь этого, используя ключевое понимание динамического программирования, где захватываются промежуточные результаты (по внутренним суммам и продуктам) и использовать повторно. В этом аспекте, SPNs похож на многие существующие формализмы, такие как И / ИЛИ

графики и арифметические схемы. Однако, в то время как этот формализм, как правило, используются в качестве мишеней для составления графических моделей, SPNs являются общий вероятностная модель класс в своем собственном праве, с помощью методов обучения Cient x FFI, разработанных на основе обратного распространения и EM. Эксперименты показывают, что умозаключение и обучение с SPN, может быть как быстрее и точнее, чем с помощью стандартных глубоких сетей. Например, SPNs выполнить завершение изображения лучше, чем у многих внедренных глубокие сети для решения этой задачи. SPNs также интригующие потенциальные подключения к архитектуре коры головного мозга (см (Пуны & Домингоса, 2011)).

Реляционные сети суммарной продукции (RSPNs) расширить SPNs для обработки реляционных доменов, таких как социальные сети и клеточные пути (Нат & Domingos, 2015). Как MLNs, RSPN представляет собой распределение вероятностей атрибутов и отношений между набором объектов. В отличие от MLNs, умозаключение в RSPN всегда послушное. Для достижения этой цели, RSPNs Определим набор классов объектов. Класс также указывает возможные атрибуты каждого объекта, классы его составных частей, а также возможные соотношения между частями. Например, на рисунке 5 показан частичный класс Специфическое катиона для простой политической области. А «Регион» может

Класс Регион:

сменная часть нации отношение Смежны (нации,
нация) отношение конфликты (нации, нация)

Класс Nation:

уникальная часть правительства
сменная часть лица атрибут
HighGDP
отношения Опора (Person, правительство)

Рисунок 5: Частичный класс де определению для реляционной сети суммы произведения в простой политической области.

содержать любое число стран, и две страны в пределах областей могут быть смежной в конфликте, или оба. А «Nation» содержит правительственный объект и любое количество людей объектов, каждый из которых может или не может поддерживать правительство. Класс также указывает на послушное распределение по атрибутам, частей и отношений. Структура и параметры SPN, можно извлечь из данных. RSPNs были применены к социальной сети и автоматизированных систем отладки, где они были намного быстрее и точнее, чем MLNs.

Таким образом, в связи с достигнутой комбинацией между символическими представлениями и subsymbolic (в данном случае, графические) функциями и формами мышления, MLNs и связанным с ними формализмом обнаружил растущее число приложений и приобрел популярность в AI в целом. Поэтому их можно рассматривать как истории успеха символично-численного интегрирования тесно связан с исходной областью neurally-правдоподобных или -inspired реализаций высокого уровня когнитивного мышления, иллюстрирующими ценной теорией работы с высоким отношением к компьютерной науке и, таким образом, дополняя вышеупомянутые примеры-среди других заданных NSCA и Penalty логико-а соответствующие нейронные-символические приложения на системном уровне.

8. Что касается Neural-символических систем для человеческого уровня искусственной разведки

После довольно технической экспозиции последних двух разделов, мы теперь вернемся к более концептуальна мотивированным соображениям, направленным на соотношении между нейронной-символической интеграцией и исследованием в области человеческого уровня искусственным интеллекта (HLAI) перед определением соответствующих проблем и исследовательские возможности.

HLAI, понимаемый как стремление к искусственно созданным объектам с человекоподобными способностями, было преследуемым человечеством с моментом изобретения машин. Он также является движущей силой в создании искусственного интеллекта как дисциплина в 1950-х годах. 20-го века А.И., однако, переросла в более узком направлении, делая упор все больше и больше специального назначения и одного метода управляемых решений для проблем, которые были когда-то (или все еще) считаются сложным, как игра игры, решение проблем, естественного языка, компьютерное зрение, когнитивная робототехника, и многие другие. По нашему мнению, 20th-

Поэтому век ИИ может быть воспринят в качестве эксперта AI, производство и проводит решения соответствующих конкретных задач с помощью специализированных методик, предназначенных для решения задач, о которых идет речь. Мы не ставим своей целью, чтобы сказать, что это плохо развитие, совсем наоборот, мы считаем, что это был (и остается) очень полезным и очень успешным начинанием достаточно (и в некоторых случаях хорошо зарекомендовавшей) возможности для значительного влияния на общество.

Однако, несмотря на то, был один из основных идей в первые дни AI-стремление к HLAИ сокращается в 20-м веке, предположительно потому, что первоначальное видение создания систем с требуемыми возможностями оказалось гораздо труднее понять, чем ожидается в начале. Тем не менее, в последние годы омоложение оригинальных идей стало очевидно, объясняется, с одной стороны, пониманием того, что некоторые сложные задачи, которые выходят за рамки специализированных систем, а с другой стороны, открытие новых методологий соответствующего решения трудно проблемы общего интеллекта. Некоторые примеры таких новых методик являются многочисленными подходами к машинному обучению, неклассические логические рамки или вероятностных рассуждениям, лишь некоторые из них. Кроме того, быстрое развитие нейронаук на основе изобретения, по существу, повторно определены средства записи и анализа нейронных моделей активации в головном мозге в фл и епсед этого развития. Они сопровождается междисциплинарными е и след ортой в рамках когнитивного научного сообщества, в том числе психологов и лингвистов с подобными видениями.

Совершенно очевидно, что реализация HLAИ требует междисциплинарной интеграции идей, методов и теорий. Действительно, мы считаем, что такие дисциплины, как (узкая) искусственная интеллекта, неврологии, психологии и компьютерной лингвистики будет существенно сходить, прежде чем мы можем надеяться реализовать подобный человеку искусственно интеллектуальные системы. Одним из центральных вопросов в этом стремлении, таким образом, мета-вопрос: Какие конкретные направления исследований, которые могут быть отслежены в ближайшем будущем для того, чтобы продвинуться в правильном направлении, и чтобы избежать прогресса, в настоящее время заметно обсуждаются, хотя и весьма спекулятивные -Меньше желательные технологические или социальные сценарии? Общее видение не дает никаких ответов на это, и в то время, очевидно, что нам потребуются некоторые великие всеохватывающим междисциплинарные теории для HLAИ, мы не можем надеяться достичь этого в один гигантский скачок. Для практических целей отказа от чистой необходимости, так как мы не можем, и по принципиальным соображениям не следует даже пытаться, кромсать наше Научное наследство (Besold, 2013) -мы требует отождествления следующих шагов, в частности тем, которые являются достаточно узкими, чтобы они может осуществляться, но достаточно универсален, чтобы они могли продвигать нас в правильном направлении. Это подводит нас к теме данной обзорной статьи, нейронные-символической интеграции.

Предложение для нейронной-символической интеграции в таком направлении исследований начинается с двух наблюдений. Они соответствуют два принципиальным перспективам пересекающихся и сочетающих в нейронной-символической интеграции:

(Я) Физическая реализация нашего ума на основе нервной системы, то есть на сеть нейронов, как идентифицируется и исследовано в нейробиологии. Если мы надеемся достичь HLAИ, мы не можем игнорировать это нервный или subsymbolic аспекта биологических интеллектуальных систем.

(Б) Формальное моделирование сложных задач и человека, в частности, абстрактное мышление основано на манипуляции символов, сложные символические структуры данных (как графы, дерева, форма и грамматик) и символической логики. В настоящее время не существует жизнеспособной альтернативы символических подходов для того, чтобы закодировать сложные задачи.

Эти две точки зрения, однако, нейронная и символические-существо ортогональны друг к другу с точки зрения внедренный в соответствующих дисциплинах. Как свидетельствует также в предыдущих разделах, символически понимание нервной системы трудно и требует значимы количества повторно определены теории и техники, все еще далеки от предоставления общих решений. Кроме того, совершенно не ясно, в настоящее время, как символическая обработка в целом возникает из нейронной активности сложных нейронных систем. Таким образом, символическое представление знаний и манипулирование сложных структур данных, используемых для представления знаний на уровне, необходимый для HLAИ является путем вне сферы тока искусственных нейронных подходов.

В то же время люди, используя их нейронную Wetware, то есть их мозг-способен успешно справляться с символическими задачами, манипулировать символические формализмы, для представления знаний, используя их, и решать сложные задачи, основанные на них. В целом, существует значительное несоответствие между человеческой нейрофизиологией и познавательными способностями как образцы для подражания для HLAИ с одной стороны, и теорией и вычислительными моделями для нейронных систем и символической обработкой с другой стороны.

Мы убеждены в том, что СИГВВП прогресс фи косяк в HLAИ требует согласования нейронных и символических подходов с точки зрения теории и вычислительных моделей. Мы считаем, что это примирение столько же центральной для продвижения HLAИ как для когнитивной науки и когнитивной нейронауки (как описано в разделе 4). Мы также считаем, что стремление к этому примирению является своевременным и целесообразным на основе текущего состояния искусства, и будет продолжаться кратко упомянуть некоторые недавние изменения в нейронной-символической интеграции, которые мы считаем, что особенно важно для обсуждения HLAИ. Дополнительные указатели могут быть найдены, например, в (Bader и др, 2005; Сеп и Hitzler, 2007; Garcez и др., 2009).

Линия исследования мы хотим отметить берет свое начало от расчетных моделей в (узком) AI и машинного обучения. Он пытается формально определяют и практическую реализацию системы на основе ИНСА, которые способны к обучению и работе с логическими представлениями сложных областей и умозаключения по этим представлениям. Хотя это может быть прослежена до (McCulloch & Pitts, 1943) «ы вехой бумаги о соотношении логики и двоичным порога ИНС, это было в значительной степени бездействующим до 1990-х, где (как уже упоминалось в предыдущих разделах) Первый neural- были реализованы, сравните символические системы обучения на основе этих идей, в частности, (Тоуэллы & Shavlik, 1994; Garcez & Zaverucha, 1999; Garcez и др., 2002). В то время как эти первоначальные системы были еще кон определен в пропозициональные логики, в последние года система с аналогичными возможностями на основе фантастической логики первого порядка, была реализовано, сром., Например, (порывы, "K

uhnberger, и Geibel, 2007) или Раздел 6.

Следует отметить, однако, что эти системы, несмотря на то, что они обеспечивают концептуальный прорыв в обработке символов по ИНС-прежнему сильно ограничены в их области и применимости, а также улучшения в этих направлениях не кажутся просто на всех, Для того, чтобы преодолеть такие проблемы, мы также требуем новых идей, заимствованных из других дисциплин, в целях установления нейронной-символической системы, которые приводятся в движении видения HLAИ. Результаты когнитивной психологии на особенностях человеческого мышления, которые, как правило, не охваченные стандартными логическими методами, должны быть включены. Последние парадигмы для ИНС, которые более сильно вдохновили от нейробиологии-как описано, например, в (Мааса, 2002) -need быть исследованы на нейронную-символической интеграции. Помимо этого, мы требуем новых творческих идей, заимствованных, в частности, из теории динамических систем или органических вычислений для дальнейшего тему.

Представлен выбор считается иллюстративными и есть несколько других электронных FF орты, которые могут быть упомянуты. Тем не менее, мы выбрали эту линию исследований для представления в последнем абзаце, как нам представляется, быть типичным и представителем в том, что он приводится в движение с помощью компьютерной науки, машинного обучения и классического исследования А. И. пополнен идей когнитивной науки и нейронауки.

9. (Most) Последние изменения и незавершенное от Neural-символического соседства

Прежде чем завершить эту самоуверенную съемку с двумя разделами, посвященных будущим проблемами и направлениями (раздел 10) и в некоторые короткие всеобъемлющих замечания (раздел 11), в дальнейших мы будем выделить последние работы всплывать вне традиционного ядра поля, но в тематической близости от нервного -symbolic интеграция с высокой значимости и потенциального воздействия на соответствующие основные вопросы.

9.1 Другие Парадигмы в подсчетах и представление сужая Neural-символьной Gap

Мы первая хотим обратить внимание читателя на недавно предложенной нейро-вычислительный механизм, называемый «Концепторы» (Jaeger, 2014). Вообще говоря, Концепторы являются Предложенная математическая, вычислительная и нейронная модель объединения двух основных идей, а именно о том, что режимы обработки рекуррентной нейронной сети (РНН) можно охарактеризовать геометрией соответствующих государственных облаков, и что режим переработки попутного выбран и стабилизируется, если состояния соответствующего RNN являются фильтровали, чтобы оставаться в пределах определенного соответствующего состояния облака. Когда вместе взятые, это в основном сводится к возможности контролировать множество режимов обработки одного единственного RNN, представляя собой форму контроля сверху вниз к снизу вверх коннекционистской сети. В то время как основные математические соображения довольно сложные, общий подход можно представить как два шага подход. Во-первых, ди-FF различны ellipsoids enveloping ди далее erently формованного состояния облакам ведомое экспонаты РНН при воздействии ди далее различны динамическая входных паттернов каждый дает ди Ff концептор различны. После того, как шаблоны вождения были сохранены в сети (то есть, сеть научилась реплицировать шаблон возбужденного состояния последовательности в отсутствии стимулов водителя, выполняя тип самостоятельного моделирования), они могут быть выбраны и стабильно регенерируют вставить соответствующие идейные фильтры (т.е. геометрия эллипсоидного служит в качестве решения механизма фильтра) в цикле обновления сети. Это частично геометрический характер подхода также о и далее ERS еще одно преимущества в том, что Концепторы может быть объединено с помощью операций родственных (и частично со-экстенциональных) в булеву логику, таким образом, вводя дополнительное измерение семантической интерпретации, особенно обслуживающее символической сторону нейронной-символической интеграции ,

Хотя Концепторы являются подход к оснащению суб-символической модели в виде RNNs с определенными символическими возможностями и свойствами, второй недавний результат исследований, мы хотим отметить, направлена на коннекционистского реализации вычислительной архитектуры фон Неймана (Von Neumann, 1945), как прототип символично компьютерная модель. «Нейронные тьюрингова машины» (HTM), введенные в (Graves, Wayne, & Danihelka, 2014), в основном пару RNNs с большим, адресуемой памятью, дающая от конца до конца ди Ff дифференцируема и gradientdescent обучаемого аналога памяти-лента Тьюринга обогащенного бесконечна -состояние машины. С точки зрения

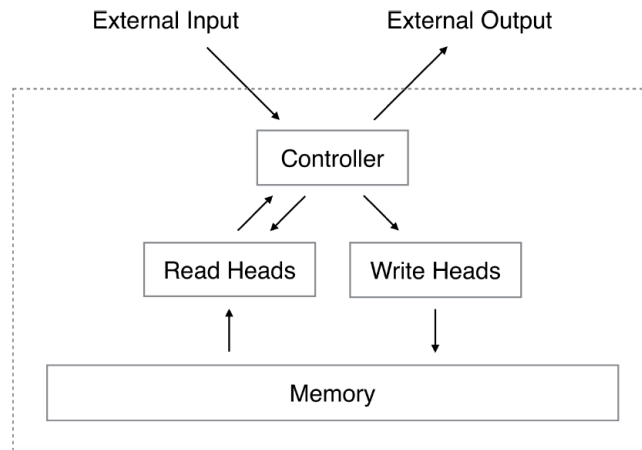


Рисунок 6: (. Взято из (Graves и соавт., 2014)) Базовая структура нейронной машины Тьюринга: Сетевой контроллер принимает внешние входы и выходы излучают в течение каждого цикла обновления. Одновременно он читает и пишет из матрицы памяти через параллельные чтения и записи головы.

архитектура, НТМ состоит из нейронной сети и контроллером банк памяти (смотрите рисунок 6). Контроллер управляет взаимодействием с окружающей средой с помощью входных и выходных векторов и взаимодействует с матрицей памяти, используя селективную операции чтения и записи. Главная особенность НТМА является их впритык ди далее erentiability, который позволяет для обучения с градиентным спуском подходами. Для того, чтобы ввести эту способность, операции чтения и записи были де определены таким образом, что позволяет им взаимодействовать с ди и след различны степени со всеми элементами в памяти (модулированного механизмом фокусировки, ограничивающим фактический диапазон взаимодействия), в результате чего в вероятностной режиме чтений и записей (ди и след от различны высоко специфичного дискретных операций чтения / записи Тьюринг машин выполняет). Кроме того, взаимодействие с памятью имеет тенденцию быть очень редкими, смещающей модель в стороне без помех хранения данных. Какое место в памяти для участия в определяются специализированными сетевых выходами параметризации операции чтения и записи по матрице памяти, де Fi Ning нормализованного взвешивания над строками матрицы. Эти выходы сети, таким образом, можно сосредоточить внимание на резко, определяемую одностанционную памяти или распространение его слабо охватывают память в нескольких соседних местах. Несколько экспериментов также сообщалось в (Грейвса и др.,

2014) показывают, что НТМ, как правило, способны к изучению простых алгоритмов из примера данных, а затем обобщить эти алгоритмы вне тренировочного режима.

Другой подход к комбинируя учебные возможности коннекционистских сетей с формой Read- и записываемой памяти, а также вывод мощностями, был представлен в виде так называемые «сети памяти» (Weston, Чопр, и Борд, 2015). Сеть памяти состоит из массива индексированных объектов (таких как массивы или векторы) m составляющий компонент памяти, компонент карты ввода y преобразовывать входящий вход к соответствующему внутреннему представлению функций, обобщение g обновление (и / или, возможно, сжатие и обобщающие) старые воспоминания дали новый вход, картографическую выход функции O производить новый вывод дал новый вход и текущее состояние памяти, и, наконец, в ответ p

преобразования выходного сигнала в формат фи NAL. за исключением m , остальные четыре компоненты y ,

Γ , O , а также p потенциально могут быть извлечены. На этом фоне, если вход $I_{кс}$ прибывает в сети памяти, $I_{кс}$ преобразуется в внутреннее представление признака $I(x)$ Перед обновлением памяти m в m данный $I_{кс}$: $\forall y \in G(m, I(x), m)$. Затем выходные функции

$o = O(I(x), m)$ вычисляются на данные новый вход и память, и функция вывода декодируется в ответ ϕ конечного $R = R(O)$. Поток модели остается одинаковым во время обучения и тестирования, с единственным ди Ф.Ф.

разностная в том, что параметры модели \mathcal{Y} , Γ ,

O , а также p остаются неизменными во время тестирования. В связи с этим очень общей характеристикой, сети памяти могут быть гибки в отношении возможных механизмов реализации для \mathcal{Y} , Γ , O ,

а также P , как правило, позволяет использовать стандартные модели от машинного обучения (например, RNNs, SVMs или деревьев решений). В (Weston и др., 2015), также отчет о качестве конкретизации общей модели дается, описывающая реализацию сети памяти для применения в текстовой вопрос ответа (QA) настройки с использованием нейронных сетей в качестве компонентов. В результате чего системы затем применяется и оценивала отдельно в крупномасштабном QA и настройке моделируемого QA (последний также включая некоторые QA с ранее не словами), перед объединением обоего установок в ϕ конечном гибридном крупномасштабном и моделировала систему QA, выступающую в качестве доказательства концепции для возможных будущих возможностей подхода.

9.2 Другие применения системы Сужение Neural-символьной Gap

После предыдущего обсуждения нейронных-символический соответствующие достижений на архитектурно-методическом уровне, в этом пункте мы остановимся на несколько последних прикладных системах, которые частично или полностью на основе коннекционистских подходов, но успешно применяются для задач преимущественно символического характера.

Область применения Первый, который в последнее время наблюдается быстрое развитие и систем, работающих на ранее не достижения, качественно новый уровень производительности является областью зрения на основе задач, таких как семантической классификации изображений в соответствии с их изобразительной содержанием. В (Vinyals, Тошев, Bengio, и ERHAN, 2014) и (Karpathy & Fei-Fei, 2015) были представлены два подхода к связанным проблемам. В то время как порождающая, глубокая РНН модель в (Vinyals и др., 2014) порождает естественный язык предложение, описывающее содержание входных изображений, (Karpathy & Фей Фей, 2015) представляет собой подход к признанию и маркировка задачам для содержания ди и далее различны изображений регионы (соответствующие объекты различны ди и следа и их свойства в изображениях). Модель из (Vinyals и др., 2014) сочетает в себе входную сверточную нейронную сеть, используемую для кодирования изображения (сеть предварительно подготовленная для задачи ϕ катионного изображения Classi, и последний скрытый слой служит в качестве интерфейса к следующему естественному языку декодера), с выходом РНН, генерирующий естественные предложения языка на основе обработанных данных входной сети. Результатом является полностью стохастической градиентной-обучаемым нейронной сетью объединения в своих подсетях видения и языковая модель, и, если положить в приложениях-значимо превосходят предыдущие подходы и сужая разрыв в работоспособность человека от используемых тестовых наборов. В (Karpathy & Фей Фей,

2015), область сверточная нейронная сеть применяются к отождествлению объектов (и соответствующим областям) в изображениях и их последующее кодирование в мультимодальном пространстве вложения, вместе с двунаправленным RNN, который вычисляет представления слов из входных фраз, а также встраивает их в то же пространство. После того, как это отображение изображений и предложений в пространство вложения было достигнуто, объективное выравнивание в терминах балла изображения предложения (в зависимости от индивидуальных счетов-области слова) вводятся, связывая

вместе обе части модели. В результате чего системы снова работают лучше, чем предыдущие подходы, показывающее состояние выполнения искусства в изображении-предложении ранжирования критериев и превосходя предыдущие базовые уровни извлечения в полнокадровых и области на уровне экспериментов. Обе модели и соответствующие системы реализации замечательны тем, что классически задач, связанных с семантические описания были связаны с базами данных, содержащими некоторую форму фоновых знаний, а также соответствующие изображения обработки / компьютерного зрения подходы в течение длительного времени также полагались на методы, основанные на правилах. Описанная работа дает доказательство того, что этот тип задача-которая явно концептуально еще работает на символическом уровне, как правило, также может быть решен на практике в чисто коннекционистских архитектурах.

Другой тесно связан тип задачи является визуальной аналогии решений, как, например, применение, например фи-е изд в преобразовании изображения запроса в соответствии с примером пары взаимосвязанных изображений. В (Рид, Чжан, Чжан и Ли, 2015), авторы описывают способ, как бороться с этой проблемой снова применяя глубокие сверточные нейронные сети. В своем подходе первая сеть узнает, чтобы работать в качестве глубокого функционального кодера отображения изображений вложения пространства (который из-за свою внутреннюю структуру позволяет выполнять некоторые формы рассуждения о аналогиях с помощью довольно стандартными математическими операций), до того, как вторая сеть служит в качестве глубокая функция декодера отображение вложения обратно в пространстве изображений. Полученные модели способны учиться аналогиями решений на основе внешнего вида, вращение 3D поза, и атрибуты несколько объектов, охватывающий самые разнообразные задачи, начиная от завершения аналогии, передач анимации между видео игровых персонажами и 3D представляют перенос и поворот. Кроме того, здесь, в общих моделях аналогии ранее были в основном символический и генеративные модели изображений пришлось кодировать значимы до фоновых знаний и ограничивают допустимые преобразования, в результате чего общий домен задача в основном символический характер.

Домен-ди-FF различны адресуется работы, описанной в (Ферстер, Ассаэль, де Фрейтас, & Whiteson, 2016). Там, глубоко распределенный рецидивирующий Q-сеть успешно применяется для изучения протоколов связи для групп агентов в координации задачах в частично наблюдаемой обстановке многоагентной. Этот тип задачи требует агентов координировать их поведение максимизации общих Пайо и далее, имея дело с неопределенностью в отношении оба, скрытого состояния окружающей среды и информации о состоянии и будущие действия членов своей команды. Для того, чтобы решить эту проблему, так называемые «глубоко распространены периодические Qnetworks» (DDRQN) были introduced. Эти сети основаны на независимых глубоко Q-сеть агентов с Long кратковременной памятью (LSTM) сети (Hochreiter & Schmidhuber, 1997) -а в настоящее время очень популярный тип RNN для выполнения задач, связанных с обработкой последовательных данных, дополнена (я) подачи каждого агента с его предыдущим действием в качестве входных данных на следующий временный шаг (позволяя агент, чтобы приблизить их собственную историю действия по наблюдению), (б) разделение весов в одной сети в период между всеми агентами и кондиционированием на ID агента (что позволяет быстро обучение и разнообразно поведение), и (III) отключение опыт воспроизведение (избегая проблемы с нестационарностью, возникающей в результате одновременного обучения нескольких агентов). Полученная система оказывается способной к решению (для ограниченного числа агентов) два классических логических загадок обязательно требующего совместное общения и обмен информации между игроками в своих оптимальных стратегиях решения. В то время как будущая работа будет по-прежнему должна доказать, масштабируемость подхода, работа в (Ферстер и др., 2016), тем не менее, является доказательством концепции для способности коннекционистских сетей, чтобы узнать коммуникационные стратегии в сценариях-которых многоагентных является еще экземпляр подлинно символической емкости, которая была успешно реализована в нейронной сети архитектуры.

Фи NAL пример успешной прикладной системы адресации аспектов neuralsymbolic интеграции, которые мы хотим упомянуть, является AlphaGo Go система игры представлены в (серебро, Huang, Мэддисон, Guez, Sifre, ван ден Driessche, Schrittwieser, Antonoglou, Panneershelvam, Ланкто, Dieleman, Греве, Nham, Kalchbrenner, Sutskever, Lillicrap, Лич, Kavukcuoglu, Graepel, и Hassabis, 2016). В то время как Go в течение длительного времени доказала свою вычислительно неосуществимо решить за счет быстро растущего дерева игры и соответствующего пространства поиска, авторы предлагают архитектуру оснащенную глубокие нейронные сети на основе выбора перемещения и позиционирования функций оценки (научились, используя комбинацию руководил учась экспертных игр человека и усиление обучения с массовым параллелизмом самостоятельной игры). Интеграция этих, в виде политики и стоимости сетей с дерева поиска опережения подхода Монте-Карло, в результате высокой производительности дерево поиска позволяет по электронной FFI ciently исследовать дерево игры и играть в игру на уровне наравне с лучшим человеком в мире игроки. Ди Ф.Ф. отличны от предыдущих примеров из зрительной области, в результате архитектуры ближе к подлинно нейронный-символическим в сочетании между обученной коннекционистскими сетями для обрезки и сужения пространства поиска и последующий алгоритма Монта-Карло, выбирающие действия по поиску опережения дерева. Благодаря ранее не достижение производительности AlphaGo представляет собой яркий пример возможностей, введенных путем объединения Коннекционистских и символические подходов и методы.

9.3 Другие Исследования Е ФФ орты и программы сужая Neural-символьной Gap

В этом Fi NAL пункте мы хотим сосредоточиться на три ди Ф.Ф. различны направления работы, которые выделяются среди обычной работы в нейронную-символической интеграции в том, что они либо начинают с точки зрения ди Ф.Ф. различны чем нейронные вычислений или логики на основе моделирования, или применять методы и подходы которые обычно не используются в поле. Примеры таких «неклассических нейронных-символический исследовательские программы» можно найти, например, в некоторых расследованиях методов и механизмы многоагентных систем, в эмпирических попытках понимания RNNs с помощью модели языка в качестве стендового предоставления данных для качественного анализа и в теоретической информатике и теории сложности проектов с ориентацией на практической ди Ф.Ф. егences между символическими и коннекционистскими подходами с теоретическими средствами. Мы опишем три последние предложения как прототипические случаи в точке.

(Я) Анкерное Знание взаимодействия в многоагентных системах: Типичная установка в многоагентных системах включает в себя средство изучением окружающей среды, изучение от него, рассуждения о его внутреннем состоянии с учетом ввода, действуя и, при этом, изменении окружающей среды, исследуя далее, и так далее, как часть постоянного цикл (Cutsuridis, Хуссейн, и Тейлор, 2011). Одна из самых больших проблем в этом контексте является (двунаправленная) интеграция между зондированию низкого уровня и взаимодействия, а также формирование знаний на высоком уровне и дальнейших рассуждений. В (Besold, K

uhnberger, Garcez, Sa FFI Отти, Fischer, &

Банди, 2015) междисциплинарная группа исследователей, охватывающих от когнитивной психологии через робототехнике к формальному ремонтного онтология и мотивировочной эскизе концептуально (что составляет) сильно познавательного мотивированной нейронной-символической архитектуры и модели приобретения знаний, расположенного агента через взаимодействие с окружающей средой в постоянном

цикл обучения на основе опыта, более высокий порядок обсуждения и формирования теории и пересмотра.

В соответствующих исследованиях по программам дополнительно введены и фактически оформлены в терминах нейронной-символической интеграции (Besold & Kuhnberger, 2015) - ди ФФ различны темы также мужчин- tioned в настоящей обзорной статье объединены для достижения этой цели. Исходя из вычислительной неврологии и сетевого уровня когнитивного моделирования (как представлено, например, в рамках Концепторов, описанные выше) в сочетании с психологическими соображениями о воплощенном взаимодействии как части формирования знаний, представления низкоуровневых зондирований агента создаются в качестве выходных земля слой предусмотренной архитектуры / модели. Этот выходной сигнал затем поступают во второй слой, выполняющего расширенную форму закрепления (Coradeschi & Ca FFI Отти, 2000) не только заземляющие символов, относящихся к воспринимаемым физическим объектам, но и динамически адаптироваться и ремонт приобретенных отображений между окружающей средой и внутренним представлением. Улучшенные представления низкого уровня как выход крепежного слоя затем, в своей очереди подают в-в терминах нейронной-символической интеграция наиболее классический подъем слоя, т.е. до архитектурного модуля многократного применения методов, сочетающих нейронных обучения и представление временного знаний в стохастические сети (как,

например, уже обсуждалось в виде RBMS в контексте NSCA рамки в Разделе 3) непрерывно дополнительно повысить уровень абстракции обрабатываемой информации в направлении символического представления высокого уровня их. После того, как су FFI ciently общий уровень был достигнут, фи NAL слой сочетая сложные когнитивно-вдохновленные методы рассуждения на символическом уровне (такие как, например, аналогии решений, концепция смешивания и концепции реформирования), а затем используется для поддержания модели агент и мир на уровне явных декларативных знаний (Bundy, 2013), который, тем не менее предназначен кормить обратно в нижние слои-среди других, служа в качестве руководства для дополненного основного процесса.

(Б) Визуальное и Понимание Периодические сети: LSTMs выделяются между текущей глубиной сетевыми моделями из-за их способность хранить и извлекать информацию в течение длительных периодов времени, используя встроенный в постоянных каруселях ошибок, управляемых с помощью явных механизмов стробирования. Тем не менее, в то время как будучи использован в нескольких высоких про-фи ле прикладных исследований, это пока только мало известно о точных свойствах и представления возможностей LSTMs. В (Karpathy, Johnson, & Фей-Fei, 2015), первый шаг в сторону фи прояснения некоторых соответствующих вопросов описаны и представлены результаты, которые также имеют отношение с нейронной-символической точки зрения. Авторы эмпирически пролить свет на аспектах представления, предсказания и типов ошибок в LSTM сети, выявить наличие интерпретируемых клеток внутри сетей, которые отслеживают зависимости дальнобойных в соответствующих входных данных, и предполагают, что замечательное исполнение LSTMs может быть из-за структурные зависимости дальнего действия.

Для того, чтобы получить представление о внутренней работе LSTM сетей, язык символов на уровень модель используется в качестве интерпретируемого тестового стенда для нескольких вычислительных экспериментов. В теории, путем использования его ячеек памяти LSTM должна быть способна запоминать LongRange информацию и отслеживать ди Ф.Ф. различны атрибуты соответствующего ввода его просто обработки. Кроме того, используя обработку текста в качестве примера, вручную настройки ячейки памяти таким образом, что он отслеживает, является ли он внутри строки в кавычках или нет не является чрезмерно сложной задачей. Тем не менее, будь то LSTMs на самом деле прибегать к этому типу внутренней структуры (т.е. формирование интерпретируемых клеток с выделенными, более абстрактными задачами) не было ясно.

Эмпирический findings в (Karpathy и др., 2015) показывает, что такая «клетка задачи» действительно существует, в случае программной обработки коды в качестве текстового ввода в диапазоне от клеток проверки для круглых скобок после того, как если заявление клеток, действующих в качестве счетчика длины линии или отслеживание отступов блоков кода. На втором этапе, используя конечный горизонт l -грамм модель как *comparandum* показана, что сети LSTM удивительно хорошо отслеживать взаимодействия и зависимостей большой дальности в пределах входных данных.

От нервного-символической точки зрения, работа сообщили в (Karpathy и др., 2015) имеет важное значение, по крайней мере, двумя способами. С одной стороны, фантастические findings, связанные с существованием интерпретируемых дальностей и связанными с ними элементами в настройках сети на самом деле взять LSTMs один шаг ближе к уровню символических интерпретируемости и от полностью распределенной и нелокализованной классической коннекционистской обстановки. С другой стороны, общий подход и набор методов, применяемых в соответствующих эмпирических исследований и анализов (в том числе качественных экспериментов визуализации, статистических оценок клеточных активаций, и сравнение к l -граммные модели) предложить возможные методы, которые могут быть использованы в более детальных последующих исследованиях о извлечении знаний из LSTMs и аналогичных оценок других типов рецидивирующих нейронных сетей в целом.

(III) выявление и изучение Differences в сложности: Контрастный подход к мульти-агент, описанному в (I) и эмпирический анализ LSTM сетей в (II), второй программе исследований, предложенной в (Besold & Kuhnberger, 2015) подходы дифференциалов между ИНС и логик от угла теоретической информатики и теории сложности. В предлагаемой линии работы последних достижений в области моделирования и анализа коннекционистских сетей и новых разработок и инструментов для исследования ранее необдуманные свойства символических формализма должны быть объединены в попытке предоставления объяснения эмпирических дифференциалов и следствий с точки зрения применимости и эффективности между нейронные и логические подходы очень быстро появляется в сценарии приложений (которые, кажется, резко противоречит формальной эквивалентности, как описано в разделе 2).

С этой целью предполагается, что форма и характер полинома над головой, как расчетно-сложности дифференциалов и следствий разностные между парадигмами должны быть рассмотрены более подробно. Основная надежда состоит в том, что пристальный взгляд на этих накладных расходов может пролить свет на ранее непродуманные факторах, как до сих пор большинство результатов сложностей были созданы с использованием исключительно **ВРЕМЯ** а также **ПРОСТРАНСТВО** как классические ресурсы для этого вида анализа, и большинство анализов опустили более точные исследования оставшегося многочлена над головой после установления удобства манипулирования. На этом фоне, рабочие гипотезы для программы являются то, что для полного информативного анализа других ресурсов, возможно, должны быть приняты во внимание на коннекционистской стороне (например, как количество шипов сеть подсадки требует в процессе вычислений, или число образцов, необходимом сходимости из начального состояния сети до стационарного распределения), и что дополнительные формальные инструменты и методы должны быть применены на символическом стороне (такие как методы параметризуются-сложности, которые принимают во внимание проблемно и приложения-SPECI свойств фи гр проблемных классов, или теория описательно-сложность, которая также считает, что полиномиальное время и логарифмическое время иерархии).

Результаты этих дифференциалов и следствий обещают внести свой вклад в решении некоторых из основных теоретических и практических напряжений, возникающих при сравнении пригодности и выполнимости применения некоторых формализмов к типам дифференциалов Ф.Ф. различны задач и областей. Которые в свою очередь,

то, как ожидается, также последовательно предоставлять информацию об общих проблемах (и их потенциальных средствах правовой защите), которые нужны были бы преодолеть для продвижения к нейронной-символической интеграции на формальную и теоретической стороне.

10. Задачи и будущие направления

Во втором предпоследнем разделе нашего исследования мы даем обзор того, что мы считаем, что в настоящее время наиболее актуальные и / или, если успешно решены, потенциально наиболее полезными теоретические и практические вопросы и темы на пути к преодолению разрыва между коннекционистскими вычислениями и на высоком уровне рассуждений.

Как описано в этой статье, если коннекционизма должна быть альтернативной парадигмой логики на основе искусственного интеллекта, ИНС должен быть в состоянии вычислить символическое рассуждение эффективно и далее активно, сочетающий в себе отказоустойчивость в коннекционистском компоненте с «хрупкостью и жесткостью» ее символический аналог. На более архитектурно ориентированной стороне, за счет интеграции Коннекционистские и символические систем, используя некоторые из ранее эскизных или аналогичных подходов, гибридные архитектуры (Sun, 2002) стремятся к решению этой проблемы и о ФФ эр принципиальным способом вычислений и обучение со знанием различных типов и на ди-Фф уровни различны абстракции. Исходя из нескольких модульных подходов, таких как, например, CLARION архитектура (BC, 2003) в когнитивного моделировании или видении символических / суб-символические несколько контекстных систем (Besold & Мандли, 2010) в интеграции данных и рассуждениях, у них есть потенциал, чтобы открыть путь и служить основой для конвергенция между парадигмами и развитием полностью интегрированной монолитной нейронной-символической системой. Тем не менее, разработка таких систем в настоящее время все еще находится в зачаточном состоянии, и часто мотивируется гораздо более конкретными инженерными императивами на индивидуальной основе случая, чем координированные и целенаправленные исследования и след ортов, направленные на проверку общих подходов и понимание общих механизмы и ограничения на работе.

На более позитивной ноте, с более формально-центрированной точки зрения можно выделить несколько текущих точек аппроксимации и потенциала кон фл ие: глубокие доверительные сети основаны на машинах больцмановских в ранней симметричной модели ИНС и связанных с байесовской сетей. Некоторые нейронные-символический рецидивирующий сети очень похожи на динамическую байесовскую сеть. Коннекционистский логик модальный использует модальную логику как язык для рассуждений о неопределенности в коннекционистских рамках. Марковские логические сети комбинируют логические символы и вероятность. В заключение можно, таким образом, следует отметить, что цели, на самом деле, кажется, сходятся. Тем не менее, здесь также текущее состояние ТФ кондиционирования еще не удовлетворительны, как слишком часто слишком мало внимания уделяется интеграции на уровне механизмов, что позволяет им оставаться варьировалась вместо запуска сходимости также вдоль этого (потенциально даже более актуально) измерение.

В общем, много вопросов и пределы интеграции до сих пор не решены, а иногда даже без внимания. В настоящее время в центре внимания по-прежнему в основном лежит на любом обучении или рассуждениях, с е и след ортами, направленных на улучшение его с помощью внедренных моделей от других, а не по-настоящему интегрировать как в принципиальном, как с вкладом в оба областях. В дополнение к, например, (Garcez, Besold de Raedt, Földiák, Hitzler, Icard, Kuhnberger, Lamb, Mäkeläinen, & Silver, 2015) мы теперь обратимся к некоторым из (многих) нерешенных проблем: мы видели, что нейронная-символическую система состоит из (I) переводов от логики к сети, (б) машинное обучение и рассуждение, (III) перевод от сети к логике.

Среди основных проблем в (I) находке пределы представления, в (II) представлениях находки, которые поддаются комплексным обучение и рассуждений, а в (III), производя адрес электронной ФФ извлечения естиве знаний от очень больших сетей.

Ниже приводится перечень научных вопросов, связанных с проблемами (I) до (III), каждый из которых может служить основой для научно-исследовательской программы в своем собственном праве:

- Совмещение первого порядка логики обучения и первого порядка логики рассуждения.
- Охватывая полуобучаемое и инкрементное обучение.
- Оценка и анализ масштабных завоеваний массового параллелизма.
- Реализация обучения-рассуждение действия циклов когнитивных агентов.
- Разработка представлений для обучения, которые учатся в составе ансамбля.
- Правило извлечения и интерпретируемости для сетей с тысячами нейронов.
- Применение фи принести на практике, и на самом деле обучения ФИ приносят функцию.
- Теоретически понимание ди Ф.Ф. еferences в поведении приложений между коннекционистскими и символическими методами.
- Разработка теории доказательств и теории типа для ИНСА.
- Разработка и аналогичная абдуктивный нейро-символика с целью автоматизации научного метода.
- Исследование нейронной-символическая модель фокус внимания и моделирование и контрастные эмоции и полезные функции.

В то время как некоторые из бывших недавно начали привлекать активное внимание researchers- примеры работы на системах нейронные-символической многоагентных, предложение по программе исследований, направленных на решение обучения, рассуждая действия цикла, или е и далее орт лучше теоретически понять теоретические основы применения ди еferences и след между нервной и логическими подходами, изложенными в разделе 9-наши знания об этих вопросах лишь ограниченные и многие вопросы все еще должны быть заданы и ответил.

Также с точки зрения подхода, несколько замечаний, кажется, на месте: в то время как видим его как весьма важное значение в отношении мотиваций и в качестве источника вдохновения, мы согласны с тем, что прямой подход снизу вверх (то есть, моделирование мозга) не является продуктивным, учитывая текущее состояние Изобразительное искусство. Кроме того, мы согласны с тем, что инженерный подход справедлив, но специфичны и не будет отвечать на важные вопросы, заданные, например, в более cognitively- или HLAI-ориентированных направлений исследований (в том числе некоторые из пунктов маркированного выше). Таким образом, основополагающий подход необходим комбинируя все направления различны ди ФФ, затронутых в обзорной статье, то есть, охватывающий (по крайней мере) ди Ф.Ф. различные элементы компьютерной науки (в частности, представления знаний и машинного обучения), когнитивной науки и когнитивной нейронауки.

Наконец, с точки зрения приложений, следующий можно пересчитать среди потенциальных «убийцы приложений» для нейронных-символической системы: ИНС были очень е и далее естиве при обработке изображений и выделения признаков, о чем свидетельствует, например, путем (Ji, Сю, Ян, & Yu, 2013)

и работа сообщалось в предыдущем разделе. Тем не менее, крупные символические системы были нормой для обработки текста (как большинство подходов, используемых в настоящее время полагаются на больших онтологий, являются и не FFI Cient обучать с нуля, и в значительной степени зависит от предварительной обработки данных). Даже если сети могут обучаться без предоставления исходного знания для выполнения, а также, например, WordNet, сеть будет очень ди FFI культа для поддержания и проверки. Neursymbolic система, способная сочетать сетевые модели для обработки изображений и устаревших символических систем для обработки текста, кажется, поэтому идеально подходит для мультимодальной обработки. Здесь такие понятия, как фи принести (наш работает, например, из раздела 2), как представляется важным, так что символические системы и сетевые модели могут быть интегрированы свободно на функциональном уровне. В этом процессе, несогласованности могут возникнуть, делая разрешение столкновений ключевым вопросом. Один из подходов, чтобы увидеть несоответствие в качестве триггера для обучения, с новой информацией в любой части комбинированной системы, служащей для рассмотрения в конфликт (по аналогии с механизмами, предусмотренных в рамках изучения-рассуждения действия, описанных в предыдущем разделе). Непосредственное воздействие этого приложения будет значительным во многих областях, в том числе в Интернет, интеллектуальные приложений и инструментов, а также безопасности.

11. Заключительные замечания

Это исследование суммирует начальные связи между полями информатики (и более конкретно ИИ), когнитивной науки, когнитивной неврологии и нейронную-символьных вычислений. Ссылки, установленные направлены на отвечая на давнишнее раздвоение в изучении человеческого и искусственный интеллект, а именно воспринятой дихотомия мозга и ума, характеризующийся в компьютерной науке символического подхода к автоматизированным рассуждению и статистическому подходом машинного обучения.

Мы рассмотрели основные характеристики и некоторые проблемы для нейронной-символической интеграции. В двух словах:

нейронные-символические системы = коннекционистский машин + логические абстракции

Потребность в богатые символах на основе представления знаний формализмов, которые будут включена в системы обучения явным образом Утверждались, по крайней мере с тех пор (Valiant, 1984) 'ы семенных бумаг, и поддержку для объединения логики и коннекционизма, или логика и обучение уже один из Тьюринга собственные исследовательские усилия (Pereira, 2012).

И символические и Коннекционистские парадигмы охватывают приближенный характер человеческого мышления, но когда взятые по отдельности имеют ди Ф.Ф. различны добродетель и от недостатков. Исследования в области интеграции двух имеет важные последствия, которые могут извлечь выгоду вычислений, когнитивной науки, и даже когнитивной нейронауки. Пределы е и след интеграции еctive могут осуществляться с помощью нейронных-символьных методов, следуя потребности приложений различны ди и след, где результаты принципиальных интеграций должны быть показанными выгодно на практике, по сравнению с чисто символическими или чисто коннекционистскими системами.

Задачи для нейронной-символической интеграции сегодня возникают от цели е и далее еctive интеграции, выразительного рассуждения и прочном обучения. Вычислительном, существует проблемы, связанные с более практическими аспектами применения нейронных-символическими систем в таких областях, как машиностроение, робототехника, семантическая сеть и т.д. Эти задачи включают в себя электронное и далее еctive вычисление логических моделей, е FFI Cient извлечение приемлемого знания и, в конечном счете, ударив из правильного баланса между сговорчивостью и выразительностью.

Таким образом, обращая внимание на события по обе стороны от разделения между символическим и суб-символической парадигмы, мы становимся ближе к единой теории, или, по крайней мере, содействие более быстрый и принципиального развития познавательных и компьютерных наук и искусственного интеллекта , Для нас это является конечной целью нейронной-символической интеграции вместе с соответствующим предоставлением нейронной-символической системы с выразительным рассуждением и мощными возможностями обучения.