Представление знаний является основной темой в ИИ, и во многих исследованиях предпринимаются попытки представить сущности и отношения базы знаний в непрерывном векторном пространстве. Среди этих попыток методы, основанные на проекции, строящие вектора сущностей и отношений, минимизируя потери при переводе из головного объекта в хвостовой. Несмотря на успех этих методов, они страдают от чрезмерно упрощенной метрики потерь и недостаточно конкурентоспособны для моделирования разнообразных и сложных объектов / отношений в базах знаний. Чтобы решить эту проблему, мы предлагаем transA, адаптивный метрический подход эмбеддинга, использующий идеи обучения метрикам для обеспечения более гибкого метода эмбеддинга. Эксперименты проводятся с эталонными наборами данных, и предлагаемый нами метод обеспечивает значительные и последовательные улучшения по сравнению с современными базовыми показателями.

Графы знаний, такие как Wordnet (Miller 1995) и Freebase (Bollacker et al. 2008), играют важную роль в поиске и приложениях ИИ. Недавние исследования, такие как расширение запросов, предпочитают использовать графы знаний (Bao et al., 2014), в то время как некоторые промышленные приложения, такие как роботы, отвечающие на вопросы, также основаны на графах знаний (Fader, Zettlemoyer и Etzioni, 2014). Однако графы знаний являются символическими и логическими, где вряд ли могут быть применены численные методы машинного обучения. Этот недостаток является одной из наиболее важных проблем при использовании графа знаний. Чтобы обеспечить общую парадигму для поддержки вычислений на графе знаний, были предложены различные методы встраивания графа знаний, такие как TransE (Bordes et al. 2013), Trash (Wang et al. 2014) и Trans (Lin et al. 2015).

Эмбеддинг - это новый подход к решению проблемы представления и рассуждения для графа знаний. Он преобразует сущности и отношения в непрерывные векторные пространства, где можно выполнить завершение графа знаний и классификацию знаний. Чаще всего граф знаний состоит из троек (h, r, t), где представлены головной объект h, отношение r и хвостовой объект t. Среди всех предложенных подходов к встраиванию методы, основанные на геометрии, являются важной отраслью, обеспечивающей самые современные прогностические характеристики. Более конкретно, методы вложения на основе геометрии представляют объект или отношение в виде k-мерного вектора, затем определяют функцию оценки fr (h, t) для измерения правдоподобия тройки (h, r, t). Такие подходы почти следуют одному и тому же геометрическому принципу h + r ≈ t и применяют одну и ту же метрику потерь || h + r − t||2 2, но отличаются пространством отношений, в котором головной объект h соединяется с конечным объектом t.

Однако показатель потерь в моделях, основанных на переводе, чрезмерно упрощен. Этот недостаток делает существующие методы встраивания некомпетентными для моделирования различных и сложных объектов / отношений в базе знаний.

Во-первых, из-за негибкости метрики потерь в современных методах, основанных на переводе, применяются сферические эквипотенциальные гиперповерхности с различной вероятностью, где чем ближе к центру, тем более вероятна тройка. Как показано на рис.1, в (а) применены сферические эквипотенциальные гиперповерхности, поэтому трудно отличить совпадающие хвостовые объекты от несогласованных. Как правило, в графе знаний сложные отношения, такие как отношения "один ко многим", "многие ко одному" и "многие ко многим", всегда приводят к сложным топологиям эмуляции. Хотя сложная ситуация с внедрением является неотложной задачей, сферические эквипотенциальные гиперповерхности недостаточно гибки, чтобы характеризовать топологии, что делает современные методы, основанные на трансляции, некомпетентными для этой задачи.

Во-вторых, из-за чрезмерно упрощенной метрики потерь современные методы, основанные на переводе, обрабатывают каждое измерение идентично. Это наблюдение приводит к недостатку, проиллюстрированному на рис.2. Поскольку каждое измерение обрабатывается одинаково в (a) 1 (Пунктирные линии указывают компонент потерь по оси x (hx + rx - tx) и компонент потерь по оси y (hy + ry − ty)), неправильные объекты сопоставляются, потому что они ближе, чем правильные значения, измеряемые изотропным евклидовым расстоянием. Следовательно, у нас есть веские основания предполагать, что на связь могут влиять только несколько конкретных измерений, в то время как другие несвязанные измерения будут зашумленными. Идентичная обработка всех измерений приводит к появлению большого количества шумов и ухудшению производительности.

Мотивированные этими двумя проблемами, в этой статье мы предлагаем transA, метод встраивания с использованием адаптивной и гибкой метрики. Во-первых, transA применяет эллиптические поверхности вместо сферических поверхностей. Таким образом, можно было бы лучше представить сложные топологии встраивания, вызванные сложными отношениями. Затем, как проанализировано в “Адаптивном метрическом подходе”, transA можно было бы рассматривать как взвешивание преобразованных измерений объектов. Таким образом, шум от несвязанных измерений подавляется. Мы демонстрируем наши идеи на рис.1 (б) и рис.2 (б).

Подводя итог, можно сказать, что transA использует идеи адаптивных метрик для лучшего представления знаний. Наш метод эффективно моделирует различные и сложные объекты / отношения в базе знаний и превосходит все современные базовые модели со значительными улучшениями в экспериментах.

Остальная часть статьи организована следующим образом: мы рассматриваем соответствующие исследования, а затем представляем наш подход вместе с теоретическим анализом. Далее представлены эксперименты, и в заключительной части мы подводим итоги нашей работы.

Мы классифицируем предыдущие исследования по двум направлениям: одно - методы встраивания на основе перевода, а другое включает в себя множество других методов встраивания.

Все методы, основанные на переводе, имеют общий принцип h + r ≈ t, но отличаются определением пространства, связанного с отношениями, в котором головной объект h соединяется с хвостовым объектом t. Этот принцип указывает на то, что t должен быть ближайшим соседом (h + r). Следовательно, все методы, основанные на переводе, имеют одинаковую форму функции оценки, которая применяет евклидово расстояние для измерения потерь следующим образом: fr(h, t) = ||hr + r − tr||2 2, где hr, tr - векторы вложения объектов, спроецированные в пространство, специфичное для отношений. Обратите внимание, что эта ветвь методов сохраняет самую современную производительность.

1. TransE (Bordes et al., 2013) помещает объекты в исходное пространство, скажем, hr = h, tr = t.
2. TransH (Wang et al. 2014) проецирует сущности в гиперплоскость для решения проблемы сложных отношений, скажем, hr = h − wrT h wr, tr = t − wrT t wr.
3. TransR (Lin et al. 2015) преобразует сущности по одной и той же матрице, чтобы также решить проблему внедрения сложных взаимосвязей, например: hr = Mrh, tr = Mrt.

Проецирование объектов на разные гиперплоскости или преобразование объектов с помощью разных матриц позволяет объектам играть разные роли в разных ситуациях внедрения. Однако, как утверждается во “Введении”, эти методы не подходят для хорошего моделирования сложных графов знаний и, в частности, неудовлетворительно работают в различных и сложных ситуациях с объектами / отношениями из-за чрезмерно упрощенной метрики.

TransM (Fan et al., 2014) предварительно вычисляет отдельный вес для каждой тренировочной тройки, чтобы добиться лучших результатов.

В этом разделе мы хотели бы представить адаптивный метрический подход transA и представить теоретический анализ с двух точек зрения.