Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования  
Пермский государственный национальный исследовательский университет

Кафедра математического обеспечения вычислительных систем

Алгоритмы синхронизации агентов в системе моделирования «Better»

Выпускная квалификационная работа магистра

Работу выполнил студент группы «ПМИ‑1,2‑2011НМ» 2-го курса магистратуры механико-математического факультета  
\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Митраков А.А.

Научный руководитель:  
Доцент, кандидат физико-математических наук  
\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Замятина Е.Б.

“\_\_\_\_”\_\_\_\_\_\_\_\_\_\_\_20\_\_\_г.

Пермь, 2013

Аннотация

Данная работа посвящена алгоритмам синхронизации логических процессов для систем распределённого агентного моделирования. В области имитационного моделирования существуют два обширных класса алгоритмов: консервативные и оптимистические, причём в основном выбор делается в пользу последних. Однако оптимистические алгоритмы не лишены недостатков, которые специалисты и эксперты и по сей день пытаются нивелировать.

В работе предложен новейший подход к оптимизации данного класса алгоритмов, основанный на знаниях об исследуемой модели. В тексте представлены описания существующих и предлагаемых алгоритмов, оценка сложности, область их применения, практические результаты. Также описана архитектура агентной платформы «Better», на которой проводились эксперименты.

Оглавление

[Введение 4](#_Toc357088347)

[Глава 1. Преимущества агентного имитационного моделирования 6](#_Toc357088348)

[Имитационное моделирование 6](#_Toc357088349)

[Мультиагентные системы 8](#_Toc357088350)

[Агентное моделирование 10](#_Toc357088351)

[Глава 2. Постановка задачи. Проблема синхронизации логических процессов 12](#_Toc357088352)

[Распределённое имитационное моделирование 12](#_Toc357088353)

[Консервативные алгоритмы синхронизации 14](#_Toc357088354)

[Оптимистические алгоритмы синхронизации 17](#_Toc357088355)

[Сравнение классов алгоритмов 19](#_Toc357088356)

[Глава 3. Оптимизация алгоритмов синхронизации логических процессов 22](#_Toc357088357)

[Модификации консервативных алгоритмов 22](#_Toc357088358)

[Модификации оптимистических алгоритмов 23](#_Toc357088359)

[Методы, направленные на устранение побочных эффектов оптимистических алгоритмов 25](#_Toc357088360)

[Глава 4. Разработка платформы. Архитектура, технологии и инструментальные средства 26](#_Toc357088361)

[Глава 5. Алгоритмы синхронизации агентов, основанные на знаниях о модели 27](#_Toc357088362)

[Глава 6. Извлечение знаний о модели в динамике методами Process Mining 28](#_Toc357088363)

[Глава 7. Эксперименты. Практические результаты 29](#_Toc357088364)

[Заключение 30](#_Toc357088365)

[Библиографический список 31](#_Toc357088366)

[Глоссарий 32](#_Toc357088367)

[Приложение 33](#_Toc357088368)

Введение

Данная работа посвящена разработке системы распределённого агентного моделирования и исследованию алгоритмов синхронизации агентов. Агентное моделирование является сравнительно новым и весьма перспективным направлением в области имитационного моделирования, поскольку данный подход способен решать очень сложные и трудноформализуемые задачи, такие как моделирование социальных, экономических и политических процессов, сложные динамические и стохастические задачи в эпидемиологии, экологии, производстве, логистике, транспорте, военно-стратегическом планировании и т.д.

Основными предпосылками к развитию агентного моделирования являются развитие методов инженерии знаний и искусственного интеллекта, а также совершенствование мощностей аппаратного обеспечения. Первое послужило причиной перехода от классических реактивных агентов к интеллектуальным – способным функционировать на основе заложенных в него знаний, принимать решения в нестандартных ситуациях и самообучаться в динамике. Это позволило моделировать такие сложные сущности, как человек, фирма, социальная группа и государство.

Развитие аппаратуры играет не менее важную роль – в силу своей интеллектуальности, агенты могут занимать большие объёмы вычислений. Для традиционных методов моделирования, например, системной динамики, не имеет значения масштаб моделируемого объекта – покупатели в супермаркете или совокупность жителей мегалополиса. Для агентного подхода, описывающего поведение «снизу вверх», это является кричным показателем.

В настоящее время сложные вычислительные задачи решаются с помощью распределённых систем и специального класса алгоритмов. Агентное моделирование хорошо подходит для распараллеливания (как на уровне данных, так и на уровне задач), поэтому распределённость – одно из главных требований к платформе. За основу модели имитации взята идея распределённого дискретно-событийного моделирования – множество логических процессов продвигают время, обрабатывают события и обмениваются друг с другом сообщениями.

Одним из ключевых алгоритмов для поддержки распределённого дискретно-событийного моделирования является алгоритм синхронизации времени. Проблема заключается в том, что увеличение количества вычислительных узлов в N раз в идеале должна привести к соответствующему приросту в скорости вычислений в N раз, однако на практике это далеко не так. Приходится платить за коммуникацию агентов, неравномерность производительности и, разумеется, алгоритмы синхронизации.

Синхронизации времени являются самым узким местом в распределённом моделировании. Существует немало исследований, преимущественно зарубежных, о разработке новых методов синхронизации и совершенствованию уже существующих. Однако до сих пор *не найдено идеального решения*, полностью устраняющего ожидания в консервативных алгоритмах, либо устраняющего откаты в оптимистических. Данные алгоритмы порождают и другие проблемы – нехватку памяти, переполнение сети служебными сообщениями и т.д.

Целью работы является разработка агентной платформы и реализация нового оптимистического алгоритма синхронизации агентов на базе данной платформы. Практика показала, что наиболее прогрессивными являются те алгоритмы, которые используют знания о модели (lookahead, lookback, знания протокола сообщений, оценка следующего события). Предложенный алгоритм не только использует явно специфицированные знания о модели, но и выводит новые знания в ходе имитационного прогона.

Теоретическая значимость работы заключается в повышении эффективности параллелизма работы распределённых агентных платформ за счёт применения специальных алгоритмов синхронизации. Практическая ценность работы заключается в сокращении накладных расходов при синхронизации агентов, что приводит к снижению затрат по памяти, по процессорному времени и по нагруженности коммуникационной среды.

1. Преимущества агентного имитационного моделирования

В данной главе будет представлен краткий экскурс в область агентного моделирования: предпосылки возникновения, актуальность, сфера применения. Также будут показаны общие и отличительные признаки агентного моделирования по сравнению с традиционными методами машинной имитации.

Имитационное моделирование

Моделирование – один из способов познания окружающего мира, при использовании которого исследуемая система заменяется более простым объектом, описывающим реальную систему с заданной степенью точности.

Моделирование применяется в случаях, когда проведение экспериментов над реальной системой невозможно, затруднительно, дорого, нецелесообразно или сопровождается риском для аппаратуры или жизни и здоровья человека. Также моделирование возможно использовать для ускорения получения статистических данных о состоянии реального объекта за длительный период времени.

Различают физическое и математическое моделирование. Физической моделью может служить уменьшенная копия исследуемого объекта (например, модель башенного крана). При использовании математического моделирования поведение системы описывается с помощью формул. Из математического моделирования отделилась новая форма исследования – имитационное моделирование.

Имитационная модель – это программа, которая описывает структуру и воспроизводит поведение реальной системы во времени. В основном, имитационное моделирование применяют для исследования особым образом организованных сложных динамических систем с элементами случайности и множеством входных параметров за некоторый промежуток времени. В ходе моделирования собирается статистика о различных аспектах функционирования системы. Затем процесс имитации анализируется и повторяется для выявления зависимостей от входных параметров.

Зачастую наряду с понятием «имитационное моделирование» применяют термин «машинная имитация».

Область применимости имитационного моделирования

Ярким примером задачи, решаемой методами машинной имитации, является моделирование систем массового обслуживания. Хотя задача и может показаться специализированной, схожие проблемы возникают во многих областях науки, техники, производства, логистики, культуры, туризма и сервиса.

Оплата времени квалифицированного работника и времени использования аппаратуры составляет немалую долю расходов компаний. Определение оптимального графика использования ресурсов, позволяющего системе эффективно выполнять поставленные задачи, позволяет снизить расходы, учесть риски и повысить надёжность реальных систем.

Но также имитационное моделирование может применяться и для других задач: исследование биологических, социальных, экономических явлений, прогнозирование поведения сложных динамических систем, моделирование техногенных аварий и катастроф, выявление закономерностей и тенденций в различных временных процессах.

Преимущества имитационного моделирования

Применение имитационных моделей дает множество преимуществ по сравнению с выполнением экспериментов над реальной системой и использованием других методов.

*Стоимость*. Зачастую построить имитационную модель гораздо дешевле, чем нанимать персонал, закупать реальное оборудование, привлекать службы экспертов и специалистов. Модель можно сколь угодно много раз прогнать заново, изменяя и регулируя входные параметры.

*Время*. В реальности оценить эффективность нововведений можно лишь через месяцы или даже годы. Имитационная модель позволяет определить оптимальность таких изменений за считанные минуты, необходимые для проведения эксперимента.

*Воспроизводимость*. Современная жизнь требует от организаций быстрой реакции на изменение ситуации на рынке. Любые малейшие изменения потребуют и быстрого изменения в структуре производства. С помощью имитационной модели можно провести неограниченное количество экспериментов с разными параметрами, чтобы получить наилучший результат.

*Точность*. Традиционные расчетные математические методы требуют применения высокой степени абстракции и не учитывают мелкие детали. Имитационное моделирование позволяет описать структуру системы и её процессы в естественном виде, не прибегая к использованию формул и строгих математических зависимостей.

*Наглядность*. Имитационная модель обладает возможностями визуализации процесса работы системы во времени, схематичного задания её структуры и выдачи результатов в графическом виде. Это позволяет наглядно представить полученное решение и донести заложенные в него идеи до клиента или заказчика.

*Универсальность*. Имитационное моделирование позволяет решать задачи из широкого круга областей: от исследования покупателей в супермаркете до имитации сложных экономических и политических процессов, где стандартные приёмы исследования далеко не всегда дают адекватный результат.

Мультиагентные системы

С развитием информационных технологий, ростом информационных ресурсов, совершенствованием сетевой инфраструктуры и увеличением производительности вычислительных машин появляются всё новые и новые задачи. Более сложные, более ресурсозатратные и менее формализуемые. Новые задачи требуют поиска новых подходов и методов решения. Порою задачи бывают такими, что использование традиционных подходов и алгоритмов неприменимо в силу специфики области применения: недетерминированности, постоянной изменчивости, слабоформализуемости и т.д.

Возникла необходимость в более мощных и гибких интеллектуальных программных системах, способных непрерывно приобретать новые знания и изменять свою структуру и функции, развиваясь и адаптируясь к решаемым задачам и условиям внешней среды.

Одним из путей решения подобных задач можно назвать применение мультиагентных систем (МАС) – особой формы области искусственного интеллекта, которая базируется на знаниях и эвристических алгоритмах кооперативного поиска решения задачи.

Ключевым элементом этих систем становится программный агент, способный воспринимать ситуацию, принимать решения и взаимодействовать с другими агентами. В основу успешного коллективного решения задачи положено 3 фундаментальных принципа: кооперация, координация и коммуникация агентов.

В упрощённом виде можно представить агента как некоторую сущность, обладающую памятью и собственной базой знаний, умеющей находить решение некоторой узкой специфической задачи, взаимодействовать с другими агентами и менять правила поведения в динамике. Следует особо оговорить, что ни один агент в МАС не способен решить общую задачу самостоятельно.

Агента можно создать таким образом, что он будет иметь определенное отношение к принятию рискованных решений в условиях неопределенности. Агенты функционируют в едином виртуальном мире. В ходе переговоров агентов формируется текущее решение проблемы, которое гибко меняется в соответствии с динамикой среды.

Существует множество определений агента. Но обычно агентов определяют через свойства, которыми те должны обладать, а именно:

* автономность – способность действовать без внешнего управляющего воздействия и осуществлять контроль собственных действий и внутреннего состояния;
* активность – способность ставить цели и выполнять заданные действия для их достижения;
* реактивность – адекватное восприятие состояния среды и реакция на его изменение;
* коммуникативность – взаимодействие с другими агентами;
* целенаправленность – предполагает наличие собственных источников мотивации;
* открытость – любой агент представляет собой открытую систему;
* базовые знания – знания агента о себе, других агентах, окружающей среде.

Для достижения целей интеллектуальные агенты взаимодействуют друг с другом, устанавливают связь между собой и выполняют заданные действия или операции в соответствии с имеющимися целями и намерениями. Более того, агент, выполняя заданные действия, изменяет состояние среды и учитывает возможность возникновения нестандартных состояний (тупиков, отсутствие ресурса и др.).

В общем случае среда, где функционирует агент, имеет определенное состояние и поведение, которые могут быть известны полностью или частично. Агент познаёт среду, приспосабливается к ней и изменяет её своими действиями.

В отличие от традиционных систем, в которых решение ищется с помощью централизованных, последовательных и детерминированных алгоритмов, в мультиагентных системах решение достигается в результате распределённого взаимодействия множества агентов, ориентированных на поиск не столько оптимального, сколько наилучшего из возможных решений на текущий момент времени.

Эти новые принципы радикально отличают МАС от классических алгоритмов и позволяют применять агентов для решения крайне сложных задач.

Агентное моделирование

Классическое имитационное моделирование, объединяясь с мультиагентными системами, породило принципиально новую ветвь научного исследования – агентное моделирование.

В литературе можно найти множество различных определений агентного моделирования. В целом агентное моделирование можно определить как метод имитационного моделирования, исследующий поведение децентрализованных агентов и то, как это поведение определяет поведение всей системы в целом. При разработке агентной модели, инженер вводит параметры агентов, определяет их поведение, помещает их в некую окружающую среду, устанавливает возможные связи, после чего запускает моделирование. Индивидуальное поведение каждого агента воплощает глобальное поведение моделируемой системы.

Традиционные подходы имитационного моделирования рассматривают объекты как некоторые пассивные заявки (транзакты) в процессе. Например, модели системной динамики полны предположений о глобальных законах распределения, пуассоновских потоках клиентов и т.д. В процессном моделировании (дискретно-событийное моделирование) имитация рассматривается как совокупность процессов, объединяющих набор дискретных событий. Эти методы превосходят “аналитическое моделирование” в возможности учитывать случайность, динамику и нелинейность, но они проигрывают в том факте, что сами объекты и их взаимосвязи могут быть абсолютно различными.

В действительности, в классике поток клиентов рассматривают как некоторый марковский процесс. Однако люди могут быть с различными доходами и интересами, они могут иметь разную производительность труда; могут взаимодействовать и конкурировать, могут зависеть один от другого; могут консультироваться с друзьями или прохожими, могут принимать ошибочные решения.

Итак, мы переходим к агентному моделированию – абсолютно новому подходу в исследовании сложных динамических систем, интегрирующему в себе методы имитационного моделирования и мультиагентных систем.

Агентное моделирование избавлено от описанных выше ограничений, поскольку оно предполагает сосредоточение непосредственно на отдельных объектах, их поведении и коммуникации. Агентная модель – это ряд взаимодействующих активных объектов, которые отражают объекты и отношения в реальном мире. Таким образом, агентное моделирование делает шаг вперед в понимании и управлении совокупностью сложных социальных и бизнес процессов.

Фундаментальное отличие агентного моделирования состоит в том, что модель строится не «сверху вниз», как в системной динамике или процессном моделировании, а «снизу вверх». Таким образом, исследователю не требуется знать глобальных законов распределения. Вместо этого сами агенты за счёт своего персонального поведения формируют целостную картину об исследуемом процессе или явлении.

Область применения агентного моделирования

Хороший пример использования агентного моделирования – потребительский рынок. В очень динамичной, конкурентной и сложной среде рынка выбор покупателя зачастую зависит от индивидуальных особенностей, врожденной активности потребителя, сети контактов, а также внешних влияний, которые лучше всего описываются с помощью агентного моделирования.

Другой пример – моделирование мирового экономического кризиса, связанного с политикой ведения хозяйства различных стран. Все государства преследуют свои собственные цели, они ведут переговоры, налаживают торговые связи, вступают в конфликты, заключают торговые сделки. Вряд ли политику государства, чувствительную к любым изменениям как в мире, так и внутри страны, можно описать с помощью математических формул.

Однако не следует думать, что агентное моделирование применимо только для решения задач коммуникативного характера. Задачи, связанные с логистикой, производством, цепочками поставок или бизнес-процессами, также решаются с помощью агентного моделирования. Например, организация производства на предприятии с управлением ресурсами, движением сырья и планированием бюджета может быть эффективно исследована с помощью агентов.

Ещё одна отрасль – исследование социальных процессов. Например, моделирование распространения инфекции в случае эпидемий, выявление автомобильных заторов на городских улицах, исследование эвакуации людей с места техногенной аварии могут оказаться под силу лишь агентному подходу к моделированию.

1. Постановка задачи. Проблема синхронизации логических процессов

В данной главе мы рассмотрим переход от последовательного моделирования к распределённому, выявим проблемы, возникающие в процессе данного перехода и покажем, какие существуют алгоритмы (классы алгоритмов) для их решения. В конце главы будет показано, какими недостатками обладают эти алгоритмы, и как их можно было бы преодолеть.

Распределённое имитационное моделирование

Первоначально имитационное моделирование, зародившись в 50-х гг. прошлого столетия, представляло собой последовательный симулятор (все вычисления проводились на одном компьютере). Появлялись новые формы функционирования имитационного моделирования (системная динамика, непрерывное моделирование, дискретно-событийное моделирование), но модель вычислений оставалась прежней – последовательной.

Рассмотрим подробнее схему последовательного симулятора на примере дискретно-событийного моделирования (SDES, от англ. Sequential Discrete Event Simulation).

[рис]

В системе имеется совокупность объектов моделирования. Это могут быть потоки заявок (транзактов) для процессо-ориентированных систем, агрегаты и вентили для системной динамики, агенты для агентного моделирования и т.д. Отдельная компонента – механизм продвижения времени – выполняет планировку (диспетчеризацию) событий: у отдельных объектов просматриваются локальные списки событий (календарь событий) и выбирается событие с наименьшей временнóй меткой. Далее это событие обрабатывается, возможно, порождая новые события, и далее этот алгоритм зацикливается.

[рис.]

Остановом моделирования могут служить либо пустые списки событий (все события обработаны), либо специальные условия (окончание времени моделирования, наступление «особого» события и т.п.).

Какие можно предложить оптимизации и усовершенствования данных симуляторов? Первая и, пожалуй, единственная практически значимая модификация – совершенствование структур данных, применяемых в списках событий. Очевидно, соответствующие структуры данных должны учитывать 2 важных свойства:

* события должны быть упорядочены по времени (или, по крайней мере, должен быть реализован эффективный поиск минимума, как, например, в Binary-Search-Tree);
* должна быть реализована эффективная вставка в конец (поскольку новые события будут иметь временнýю метку больше, чем обработанные).

На рис. приведён пример структуры данных для списка событий (хорошее, но не самое эффективное решение)

[рис.]

На этом, видимо, перечень модификаций заканчивается. Следующим шагом послужило появление распределённого моделирования.

Специфика распределённого моделирования

Параллельное имитационное моделирования зародилось в конце 70-х гг прошлого столетия. Впредь мы будем рассматривать распределённое имитационное моделирование на примере дискретно-событийного моделирования (PDES, от англ. Parallel Discrete Event Simulation).

Общая идея достаточно проста и очевидна: теперь моделируемые сущности выполняются не на одном вычислительном узле, а на 2-х и более. Это могут быть системы с общей памятью (SMP-архитектуры), машины с массовым параллелизмом (MPP-архитектуры), несколько узлов в составе кластера, NUMA-платформы или суперкомпьютеры.

Каждый моделируемый объект также сохраняет локальный календарь событий. Схема функционирования не изменяется – выбирается событие с наименьшей временнóй меткой, далее оно обрабатывается, порождая новые события. На практике число имеющихся вычислительных узлов бывает много меньше числа моделируемых сущностей, в связи с чем последние объединяются внутри одного логического процесса.

Определение. *Логический процесс* – независимая обособленная функциональная единица распределённого дискретно-событийного моделирования, выполняющаяся на 1-м вычислительном узле и связанная с др. процессами в единую сеть посредством передачи сообщений.

К настоящему времени специалисты выделяют два подхода к организации параллельных средств машинной имитации:

* монолитные системы, в которых моделирование сводится к взаимодействию совокупности логических процессов;
* программные компоненты, которые позволяют объединить существующие средства имитации в одну общую сеть для реализации параллельного моделирования.

К последним относят технологию HLA (High Level Architecture), которая де-факто объявлена стандартом распределённого моделирования в США. К агентным системам, поддерживающим HLA, относят “HLA-AGENT” и “SIMAGENT TOOLKIT”.

Проблема локальной каузальности

Основной фундаментальной проблемой распределённого моделирования, к которой данная работа имеет непосредственное отношение, является нарушение локальной каузальности.

Определение. Под *локальной каузальностью* понимается гарантия того, что все события в ходе имитационного прогона будут выполнены по возрастанию (или, строго говоря, неубыванию) временных меток.

Иными словами, обеспечение локальной каузальности гарантирует, что все события будут обработаны в хронологическом порядке. В последовательных машинах таковой проблемы не было: симулятор всегда выбирал минимальное событие. В параллельных системах все логические процессы независимы, а значит, возможна ситуация обработки события «из прошлого».

Пример нарушения локальной каузальности проиллюстрирован на рис.

[рис.]

Если локальные часы логического процесса имеют время t1, а в это время поступает сообщение с временнóй меткой t2, то при t2≤t1, очевидно, нарушается локальная каузальность (поступает событие «из прошлого»).

На помощь нам приходят 2 обширных класса алгоритмов синхронизации: консервативные и оптимистические.

Консервативные алгоритмы синхронизации

Консервативные алгоритмы появились исторически раньше своих оптимистических аналогов. Впервые подобный алгоритм был описан Чанди, Мисрой и Брайнтом в 1977-1979 гг.

Общая идея данного семейства алгоритмов – не позволять логическому процессу продвигать время вперёд, пока он не убедится, что все остальные процессы достигли этого времени. Другими словами, процесс блокируется до тех пор, пока он точно не будет знать о том, что выполнение следующего события безопасно.

Пример работы консервативного алгоритма приведён на рис.



Из рисунка видно, что логический процесс не будет выполнять событие с t=5, поскольку теоретически возможна такая ситуация, что LP3 пришлёт событие с меткой t=4. Как только процесс получит сообщение с меткой t≥5, он увеличивает LBTS, а значит, сможет продолжить работу.

Определение. *LBTS* (от англ. «Lower Bound of TimeStamps» – нижняя граница временных меток) – время, равное минимуму временных меток по всем входным очередям сообщений.

Формальная цель данного класса алгоритмов – исключить теоретическую возможность нарушения локальной каузальности.

Проблемы, связанные с реализацией консервативных алгоритмов

Основной проблемой консервативных алгоритмов является теоретическая возможность взаимных блокировок (deadlocks). В связи с этим, консервативный алгоритм без разрешения данного конфликта вообще не может считаться алгоритмом (в силу определения последнего).

Пример возникновения тупика изображён на рис.

[рис]

Простейший способ разрешения тупиковых ситуаций был предложен Чанди, Брайнтом и Мисрой [] в работе «Консервативный алгоритм с нулевыми сообщениями». В области распределённого имитационного моделирования данный метод является классикой.

Основные ограничения алгоритма:

* Топология сети фиксирована и известна каждому логическому процессу
* Каждый логический процесс высылает сообщения с неубывающими временными метками
* Коммуникационная среда гарантирует, что все сообщения придут строго в порядке их отправления

Суть данного алгоритма заключается в том, что логические процессы после каждой обработки события высылают всем соседям сообщение специального типа – нулевые сообщения (Null Messages). Их временнáя метка t=a0 свидетельствует о том, что процесс гарантированно в будущем не пришлёт сообщений с меткой t<a0.

Определение. *Нулевое сообщение* – сообщение, несущее в себе только информацию о временнóю метке.

При получении нулевого сообщения логический процесс увеличивает LBTS, что делает все локальные события с t≤LBTS безопасными. Далее он обрабатывает безопасное событие (ели таковое имеется) и вновь высылает всем соседям нулевые сообщения.

Пример работы алгоритма изображён на рис.



Алгоритм оставляет открытым один вопрос: как оценить временнýю метку нулевого сообщения? Одно из возможных решений – применение lookahead («забегание вперёд») []. В этом случае после обработки события с меткой t=t0 логический процесс отправит нулевые сообщения с меткой t=t0+lookahead. В качестве значения lookahead могут использоваться оценки физических задержек устройств, знание нижнего предела в потоке заявок, дискретизация времени и т.д.

Возможные пути оптимизации консервативных алгоритмов будут сформулированы в главе 3.

Оптимистические алгоритмы синхронизации

Оптимистические алгоритмы являются вторым классом алгоритмов синхронизации. Первый алгоритм под названием «Time Warp» был предложен Джефферсоном[] в 1985г.

Оптимистические алгоритмы, в отличие от консервативных, не блокируют управление логического процесса, а просто продвигают время вперёд. В случае нарушения локальной каузальности специальные механизмы выполняют откат текущего состояния модели на прежнее согласованное состояние.

Так, например, в случае отката, вызванным событием с меткой t=t0, алгоритм должен вернуть состояние модели ***x*** на время:

где *Х* – множество состояний модели

*t(x)* – временнáя метка состояния *х*

Иными словами, выбирается состояние, предшествующее по времени сообщению с меткой t0.

Сообщение «из прошлого» в зарубежной литературе называется «Straggler Message» (в отечественной литературе термин официально не утверждён). Общая схема работы алгоритма изображена на рис.

[рис]

В отличие от консервативного алгоритма с нулевыми сообщениями, Time Warp не накладывает ограничений ни на структуру и топологию сети, ни на качество коммуникационной среды (доставка должна быть гарантирована, но порядок получения сообщений из канала не обязан совпадать с порядком отправления).

Проблемы, связанные с реализацией оптимистических алгоритмов

Одной из проблем такого алгоритма является необходимость восстановления любого предыдущего состояния модели. Основными методами решения могут служить следующие:

* хранение моментальных снимков модели;
* обратные вычисления.

Очевидно, в первом случае алгоритм приводит к значительным издержкам по памяти, а во втором – по процессорному времени. Более того, обратные вычисления, в свою очередь, обладают ещё двумя недостатками:

* требуются наличие специальных алгоритмов и процедур, способных для каждого f вычислить f-1;
* существуют такие функции f, что вычисление f-1 трудозатратно или невозможно (так называемые «односторонние» функции).

Другой проблемой алгоритма являются сообщения, отправленные непосредственно перед процедурой отката. Пример продемонстрирован на рис.

[рис]

Действительно, если логический процесс отправил сообщение с временнóй меткой t=10 и обработал событие с t=12, а затем получил сообщение «из прошлого» с меткой t=8, то он обязан откатиться назад на время t≤8. Что делать с отправленным сообщением? Джефферсон в своей работе предложил идею применения антисообщений (Anti-Messages).

Определение. *Антисообщение* – сообщение, отменяющее ранее высланное сообщение.

Итак, если процесс получает антисообщение, то возникает одна из следующих ситуаций:

1. Если временнáя метка антисообщения больше текущего времени процесса, то можно заключить, что «позитивное» сообщение ещё находится во входящей очереди. В этом случае оба сообщения взаимно уничтожаются (а сам симулятор даже не узнает об их существовании).
2. Если временнáя метка антисообщения меньше или равна текущему времени процесса, то, согласно общему правилу, возникает откат.

Следует учесть, что в последнем случае откат, вызванный антисообщением, может породить новые антисообщения, которые могут вызвать откаты на других процессах, которые могут привести к новым антисообщениям и т.д.

Определение. *Каскадный откат* – откат, порождённый принятым от другого логического процесса антисообщением. Количество таких цепочек откатов называют *порядком каскадного отката*.

Сравнение классов алгоритмов

Каковы же критерии выбора того или иного алгоритма? Как правило, выбор зависит от конкретной задачи, но в целом оптимистические алгоритмы показали бóльшую эффективность (под эффективностью будем понимать время вычисления имитационного прогона).

Консервативные алгоритмы тратят много времени на ожидание – по факту, система представляют собой «псевдопараллельный» симулятор. Тем не менее, на практике таковые до сих пор применяются. В основном, это связано с многочисленными недостатками оптимистических алгоритмов, которые в силу тех или иных причин оказываются критичными в конкретных системах.

С другой стороны, если исследуемая модель располагает весьма эффективными оценками lookback, то, возможно, консервативный алгоритм будет наилучшим решением. Также отметим, что в плане реализации консервативные алгоритмы зачастую проще в разработке, отладке и сопровождении.

Прежде чем переходить к сравнительной характеристике, рассмотрим ещё несколько недостатков алгоритмов синхронизации, которые являются следствием уже перечисленных недостатков.

Побочные эффекты от применения алгоритмов синхронизации

В предыдущих параграфах были отражены лишь общие проблемы консервативных и оптимистических алгоритмов. В данном разделе рассмотрим ещё некоторые неочевидные проблемы.

Во-первых, в консервативных алгоритмах возможна перегрузка сети нулевыми сообщениями. В классическом варианте логический процесс должен после каждого обработанного события вычислить оценку LBTS и отправить нулевые сообщения всем соседям. В зависимости от эффективности выбранного lookahead, нулевых сообщений может быть на порядок больше информационных

Во-вторых, выбор самого lookahead зачастую становится проблемой. В крайних случаях, исследователи просто полагают его равным 1 ед. модельного времени (или некоторому ∆t).

Что касается оптимистических алгоритмов, то следует отметить, что механизм мгновенных снимков требует огромное количество памяти, поскольку логический процесс заведомо *не располагает* информацией о том, будет совершён откат или нет. Это приводит к тому, что требуется сохранять снимок модели после *каждого* обработанного события. Как следствие, память следует каким-то образом освобождать. Для этих целей авторы придумали понятие *GVT*.

Определение. *GVT* (от англ. «Global Virtual Time» – глобальное виртуальное время) – нижняя граница временных меток всех событий на всех логических процессах и всех сообщений в системе с учётом всех рекурсивных откатов.

Таким образом, если локальное время процесса t=60, а GVT=15, то все моментальные снимки модели для t<15 гарантированно не будут востребованы, а значит, их можно удалить. Следует подчеркнуть, что GVT также служит надёжным средством фиксации операций ввода-вывода, которые, очевидно, невозможно «откатить» (так называемые *Irrevocable operations*).

Обычно логические процессы считают собственную оценку GVT (которая, как легко заметить, будет ниже или равна реальному значению). Вычислив оценку GVT, процесс фиксирует операции ввода-вывода и освобождает память. Данный процесс в западной литературе получил название *Fossil Collection* (букв. – «сбор ископаемых»).

Напоследок отметим, что глубокие рекурсивные откаты представляют наибольшую угрозу производительности вычислительного процесса, ведь в этом случае алгоритм проигрывает не только по памяти и процессорному времени локального узла, но и становится причиной излишнего расхода памяти и процессорного времени на других узлах. Если процессы на взаимные откаты и антисообщения тратят времени больше, чем на реальные вычисления, то это несомненно ставит под угрозу общую эффективность моделирования.

Общая сравнительная характеристика алгоритмов синхронизации

Ниже в таблица 1 представлена сравнительная характеристика консервативных и оптимистических алгоритмов синхронизации. Обращаем внимание: сравниваются два классических алгоритма (Null Messages Algorithm и Time Warp) без каких-либо модификаций.

Таблица . Сравнительная характеристика алгоритмов синхронизации

| **Признак сравнения** | **Null Messages Algorithm** | **Time Warp Algorithm** |
| --- | --- | --- |
| Требует пересылки сообщений по неубыванию временных меток | Да | Нет |
| Требует фиксированную топологию сети | Да | Нет |
| Требует, чтобы коммуникационная среда гарантировала доставку сообщений в порядке их отправления | Да | Нет |
| Требует механизм разрешения тупиков | Да | Нет |
| Требует наличия служебных сообщений | Да (нулевые сообщения) | Да (Антисообщения) |
| Дополнительные вычисления | LBTS | GVT |
| Требует механизм откатов | Нет | Да |
| Уровень параллелизма | Низкий | Высокий |
| Требования к памяти | Расходы памяти несущественны | Расходы памяти существенны |
| Расход времени процессора | Время может тратиться на ожидание (простой процессора) | Время может тратиться на ненужные вычисления (ввиду откатов) |
| Эффективность вычислений | Зависит от времени ожидания | Зависит от количества откатов |
| Сложность реализации | Относительно проста | Весьма не тривиальна |

Поскольку оптимистические алгоритмы всё же являются более перспективными в плане параллелизма, большинство специалистов в области имитационного моделирования ставят объектом исследования именно их. Была предложена масса улучшений и модификаций существующего Time Warp, направленных на устранение его многочисленных недостатков.

Данная работа также ставит целью разработать оптимистический алгоритм, который был бы эффективнее алгоритма Джефферсона и многих его модификаций.

1. Оптимизация алгоритмов синхронизации логических процессов

В данной главе будут предложены улучшения, оптимизации и модификации стандартных алгоритмов синхронизации логических процессов. Что касается оптимистических методов, то помимо улучшения непосредственно самих алгоритмов, авторы пытаются устранить побочные эффекты «оптимизма».

Модификации консервативных алгоритмов

Lookahead

Первой модификацией алгоритма было использование уже описанного ранее lookahead. Изобретателями данного приёма являются Nicol и Rilay (1988г.).

Данный механизм является одним из самых эффективных, поскольку он использует знания о модели.

Carrier-Null Message

Алгоритм Carrier-Null Message был изобретён британскими исследователями Вентоном Каем и Стифаном Тёрнером[] в 1990г. и позже обобщён Кеннетом Вудом в 1992г.

[…]

Control flow graph

Авторами данного подхода являются Cota и Sargent (1990г.)

[…]

Lin, Lazowska и Baer в 1990г. опубликовали статью, где был предложен класс систем, для которых теоретически невозможно предсказать значение lookahead.

Deadlock detection and recovery algorithm

Авторами данного алгоритма по прежнему выступают Чанди и Мисра.

[…]

Bounded lag simulation protocol

Иное название данного подхода – алгоритм консервативного временного окна (Conservative Time Window Algorithm). Автор протокола – Лубачевский (1989г.)[]

Conditional Events algorithm

Алгоритм условных событий был разработан Шерманом в 1990г [].

Dag Consistent Parallel Simulation

Авторы – Лим, Лоу и Тёрнер

Прочие алгоритмы

Кумар (1989г.)

Невисон (1990г.)

Врайз (1990г.)

Модификации оптимистических алгоритмов

Lazy Cancellation Algorithm

Алгоритм Джефферсона фактически использует механизм «принудительной отмены» (Aggressive Cancellation). Анат Гафни в 1988г. предложил механизм так называемой «отложенной отмены» (Lazy Cancellation).

Adaptive Cancellation Algorithm

Автор – Уилси.

Jump Forward Algorithm

Метод также носит название «Алгоритм отложенного отката» (Lazy reevaluation или Lazy rollback algorithm). Автор – Фуджимото (1990г.) []

Wolf Calls Algorithm

Мадиссети, Волранд и Мессершмидт в 1988г. [] предложили алгоритм «Волчьего воя».

Moving Time Windows

Алгоритм, предложенный Соколом, Бриско и Вайлендом в 1988г., до сих пор остаётся самой популярной идеей борьбы с чрезмерным «оптимизмом» [].

Filtered Rollback Algorithm

Идея, предложенная Шварцом и Вейтсом в 1989г. [], фактически является объединением оптимистического Time Warp и консервативного Bounded Lag алгоритмов.

Direct Cancellation Algorithm

Механизм «Прямой Отмены» впервые ввёл Фуджимото (1989г.)[] для архитектур с общей памятью.

Space-Time Simulation

Чанди и Шерман в 1992г. [] предложили довольно необычную, но весьма эффективную идею…

Filtering Algorithm

Пракаш и Субраманиан (1992г.) []

Breathing Time Buckets

Предложенный Стейнманом в 1993г. [] алгоритм «Breathing Time Buckets» и его модификация «Breathing Time Warp» (комбинация «Breathing Time Buckets» и «Time Warp») получили весьма широкое распространение за счёт полного отказа от анти-сообщений.

Rollback Relaxation Algorithm

(Palaniswamy, Aji)

Genetic Algorithm for Moving Time Windows

Wang, Tropper

Probabilistic Time Window Adaptation

Luthi

Synchronization based on FIPA-Protocol

Pawlaszczyk

Прочие алгоритмы

Алгоритм, использующий LookBack (Chen, Szymanski)

Алгоритм, использующий оценку GVT (сообщение не отправляется, пока GVT не вырастет до его метки)

Методы, направленные на устранение побочных эффектов оптимистических алгоритмов

Алгоритм Pruneback (Preiss & Loucks)

метод освобождения памяти

Алгоритм Adaptive Checkpointing (Fleischmann)

cost function to balance cost of saving states VS. cost of coasting forward over unsaved states

Алгоритм Reverse Computation (Carothers)

не все операции обратно вычислимы, спец. компилятор, спец random-генераторы

Алгоритмы эффективного вычисления GVT

1. Разработка платформы. Архитектура, технологии и инструментальные средства
2. Алгоритмы синхронизации агентов, основанные на знаниях о модели
3. Извлечение знаний о модели в динамике методами Process Mining
4. Эксперименты. Практические результаты

Заключение

Библиографический список

Глоссарий

Приложение