

**РЕЦЕНЗИЯ на выпускную квалификационную работу
обучающегося СПбГУ Тюляндина Ивана Владимировича
по теме: «Исследование применимости специализации алгоритма
Витерби скрытой марковской моделью»**


В работе студента рассматривается научная область частичных вычислений (специализации): подхода к оптимизации программ на основе предвычисления ее статически заданной части; более точно, задается вопрос, может ли быть алгоритм Витерби (алгоритм определения наиболее вероятного пути в скрытой марковской модели (СММ), приводящего к максимизации апостериорной вероятности наблюдаемых событий) быть частично вычислен для заданной СММ с получением выигрыша в производительности относительно неспециализированной версии. Для этого была вручную разработана специализированная версия алгоритма Витерби, выраженного в матричной форме, основанная на предвычислении произведений матриц, зафиксированных для заданной СММ, после чего проведены экспериментальные сравнительные исследования.

Работа хорошо структурирована, написана понятным языком и легко читается. Результаты работы показывают целесообразность применения предложенного специализированного алгоритма Витерби по сравнению с неспециализированной версией для задач биоинформатики, однако существующие state-of-the-art решения (CUDAMPF) выигрывают у специализированной версии. После прочтения работы у рецензента возникли следующие вопросы.

1. Помимо биоинформатики, алгоритм Витерби может применяться и во многих других областях (декодирование сигналов, частеречная разметка, и так далее). Однако как при рассмотрении аналогов, так и при экспериментальном исследовании мы фокусируемся исключительно на специфичных для биоинформатики подходах и данных. Каким образом в таком случае мы можем **экстраполировать**, применим ли предложенный специализированный алгоритм Витерби в других областях?
2. Почему была выбрана схема **ручной**, а не автоматизированной специализации?
3. Как специализированная версия сравнивается с CUDAMPF в случае, если **не учитывать** время, затрачиваемое на специализацию?
4. Почему при переносе реализации на GPU мы получили **падение** производительности по сравнению с решением на CPU?
5. Можно ли объяснить более детально причину **скачка** времени обработки на GPU СММ с количеством состояний больше 2000? Возможно ли каким-либо образом устранить данный скачок?
6. Почему время, показанное CUDAMPF на одинаковых данных, в экспериментах для SuiteSparse и cuASR **отличается** примерно на 20%?

Считаю, что при условии успешного ответа на поставленные вопросы студент **достоин** присвоения степени магистра по соответствующему направлению и заслуживает оценки «хорошо».

« 27 » 05 2021 г.


Подпись


ФИО