

[Скачать](#)



MeCab выводит формат файла MeCab, который представляет собой табличный выходной файл на основе ASCII, описывающий результаты. Он имеет как табличные, так и текстовые строки, то есть английские строки с выравниванием по левому краю. Настройка MeCab: Терминология Структурная зависимость Преобразование кандзи-кана Теггер позиционных фраз Оценка параметров Другой Пример вывода MeCab: Смотрите также Список японских морфологических анализаторов использованная литература внешние ссылки Официальный сайт MeCab на GitHub Категория:Бесплатные морфологические анализаторы Категория:Программное обеспечение с лицензией BSD.

Дифференциальное представление N-ацетилгалактозамина гликолипидам в центральной нервной системе дрозофилы и развитии яичников. L-галактоза, L-фукоза и N-ацетилглюкозамин гликозилируются по-разному у дрозофилы, но функции образующихся гликолипидов в значительной степени неизвестны. Мы изучили распределение двух гликоконъюгатов, содержащих терминальный N-ацетилглюкозамин, глюкозил-N-ацетилгалактозамин и N-ацетилгликопротеин, в центральной нервной системе и гонадах самок дрозофилы. Гликопротеины синтезируются в фоторецепторах зрелого глаза, а также в личиночной ЦНС и в яичниках. Со всеми фракциями мембран связаны гликопротеины и гликолипиды различного углеводного состава и молекулярной массы. Однако N-ацетилгликопротеин специфически связан с кортикальными и перинуклеарными мембранами фоторецепторов и личиночной центральной нервной системы. Напротив, глюкозил-N-ацетилгалактозамин присутствует во всех мембранных фракциях, но его особенно много в ядерной мембране личиночной центральной нервной системы, а также в корковых и перинуклеарных мембранах фоторецепторов. Гликаны концентрируются на желточных оболочках эмбрионов до 20 ч после откладки яиц, как и гликаны, содержащие концевые сиаловые кислоты. Напротив, гликаны с концевым N-ацетилгалактозамином увеличиваются в нервных мышцах и мышцах стенки тела перед началом эмбрионального и личиночного миогенеза, и их наибольшее количество в мембранах развивающихся ооцитов. N-ацетилгалактозамин также в большом количестве присутствует в железе, секретирующей альфа-амилазу. Предполагается, что

#### What's New in the MeCab?

MeCab — это новая библиотека с открытым исходным кодом для морфологического анализа, которая была разработана как облегченная реализация ANSI/ISO C под лицензией GPLv3. Морфологический анализ японских письменных текстов (кана) выполняется с помощью ряда морфологических операций. Некоторые из них включают в себя: идентификацию границ морфем, сегментацию, генерацию признаков, классификацию, тегирование и т. д. Основное внимание в этом проекте уделяется объединению морфологического анализа в единый легкий исполняемый файл, который можно использовать в качестве инструмента командной строки. MeCab включает компоненты морфологического анализа, такие как сегментация и извлечение признаков. Сегментация работает с текстом кана и кандзи. Он может распознавать границы между морфемами и символами, что может быть полезно для контекстно-независимой сегментации, устранения неоднозначности и тегирования. MeCab также поддерживает подсчет символов, n-граммовую статистику и токенизацию. Извлечение признаков работает с категоризированными тегами морфем, поэтому оно может определить, что представлено в морфеме, не распознавая саму морфему. Он работает, подсчитывая сумму уникальных тегов морфем с их соответствующими позициями. Количество персонажей: Это компонент извлечения признаков для подсчета символов между двумя морфемами. Полученное число включает в себя количество символов, которые идут до и после двух морфем. Токенизация: Это компонент извлечения признаков для подсчета количества уникальных морфем в данном тексте. В подсчет включаются все морфемы в тексте, начиная с морфем, стоящих на первой позиции от начала текста. Разложение: Это компонент извлечения признаков для разложения заданного текста на морфемы и соответствующие им POS-теги. Подобно токенизации, это извлечение признаков на основе токенов, при котором вычисляется количество уникальных морфем в заданном тексте и тегах POS. Морфология: Это компонент извлечения признаков для маркировки морфем их POS-тегами. Экстракторы: Это компонент извлечения признаков для извлечения признаков. Он реализует все вышеперечисленные функции. Сегментация: Это компонент выделения признаков для распознавания границ сегментации. В этом пакете есть подробный файл README.md, а также примеры, поясняющие, как выполнять каждую из вышеупомянутых задач. Требования: Окно терминала/командной строки в среде GNU/Linux или Windows.

**System Requirements For MeCab:**

\* ОС: Windows XP/Vista/7/8 \* Процессор: Pentium 4 с тактовой частотой 2 ГГц \* Оперативная память: 256 МБ или больше  
\* Графика: Nvidia GeForce 6600, 7800, 8800 или ATI Radeon 9200. \* Жесткий диск: 1 Гб бесплатно \* CD-ROM: аркадные  
игры Microsoft Xbox Live, предварительно установленные на диске Xbox 360 \* Широкополосное подключение к  
Интернету \* Дистанционное воспроизведение: требуется широкополосное подключение к Интернету. \* USB-  
подключение: USB 2.0 или выше

Related links: