

# Калькуляция эмоций

## Музыка и искусственный интеллект

Алексей Устинов

Трудно представить что-либо иное, кроме компьютеров и электронных сетей, что могло бы стать столь же эффективным средством, объединяющим интеллектуальные усилия миллионов и обеспечивающим функционирование по сути коллективного мирового разума.

### Автомат, интеллект, граница

Мы почти не заметили, как вступили в эпоху роботов. Пусть еще не в ту, о которой писали фантасты, но «умных» электронных автоматов и программ с каждым днем становится все больше. Еще недавно калькулятор, мгновенно извлекающий квадратный корень, вызывал восторг, а теперь нас не удивляет программа для анимации, выполняющая громадное число операций за доли секунды. Действительно, производительность машин растет, превосходя самые оптимистичные прогнозы, но главное то, что компьютерные системы приобретают явные интеллектуальные характеристики и все более «очеловечиваются». Исследования по «искусственному интеллекту» (ИИ) с первых шагов включили в сферу своих интересов изучение интеллектуальных алгоритмов, апробированных природой и присутствующих мозгу человека. Это не могло не сказаться на развитии самих средств вычислительной техники: изучая собственную природу, человек «учит» компьютерные системы видеть, слышать, воспроизводить, запоминать и даже «думать» как человек.

Примечательно, что к «умным» устройствам и программам сформировалось отношение, чем-то напоминающее отношение к детям или к животным — братьям нашим меньшим. Как иначе объяснить то, что машина наделяется качествами живого существа, ее «поведение» оценивается как интеллектуальное, а также то, что о ней говорят «она думает, решила, не может найти» или даже «она не понимает»? Но где лежит граница между интеллектуальным и неинтеллектуальным поведением, интеллектом и автоматом? С одной стороны, машина очень многое делает быстро и эффективно, значительно превосходя возможности человека, с другой — она не может решать тривиальные для человеческого мозга задачи. Может, поэтому мы более склонны рассматривать победу машины над человеком в шахматах — занятии, бесспорно, интеллектуальном — не как проявление интеллекта, творческой способности, а как итог перебора множества комбинаций и результат исключительно вычислительный. Сильный шахматный ход, сделанный программой, может быть прокомментирован словами «найден», «выбран», а аналогичный ход, произведенный шахматистом, характеризуется как «гениальный», «озарение».

Может быть, именно в творческой составляющей поведения лежит грань между интеллектом и автоматизмом? Но, пожалуй, понятие «творчество» еще менее однозначно, чем понятие «интеллект». Если ученик музыкальной школы самостоятельно сочиняет мелодию, без задания, без изучения приемов композиции, то это оценивается не иначе как «творческий» результат, ведь ученик сделал то, чему его не учили. Но когда такую же мелодию сочиняет студент консерватории, то здесь, скорее всего, будет отказано в признании результата «творческим». Оценивается, в сущности, следующее — сделал или нет человек что-либо новое с точки зрения своего потенциального уровня.

Мнения специалистов по ИИ в определении критерия интеллектуальности системы или алгоритма расходятся даже тогда, когда спор лежит в пределах одной предметной сферы. Нередко таким критерием считается способность системы к самообучению, к самостоятельной генерации алгоритма решения задачи. Но систем с такими функциями еще крайне мало среди тех, что традиционно относятся к интеллектуальным. Действительно, вполне интеллектуальное «поведение» персонажей компьютерных игр задается исходными, вложенными в программу стохастическими и переборными алгоритмами. Впрочем, и генерация музыкального аккомпанемента на основе произвольных входных данных и исходных шаблонов как немусыкантами, так и музыкантами, нередко оценивается как интеллектуальный результат. Одним словом, мы готовы признавать интеллектуальными те задачи, которые сами не можем решать так же быстро и качественно, как машина.

### Обманывайте, но так, чтобы было приятно!

Пользователь может не обладать хорошими профессиональными навыками, но при этом неплохо решать задачи в выбранной сфере, если использует интеллектуальную программу, включающую необходимые технологии и приемы и автоматизирующую рутинные операции. Весьма характерную тенденцию иллюстрирует простой, но показательный пример. Автоматические фотокамеры, превратившие всех желающих в почти профессиональных фотографов, получили (что очень несправедливо) название «дурачок». Владелец интеллектуального устройства может ничего не знать о технологии фотографии, быть дилетантом и делать неплохие снимки — теперь машина знает, как лучше снимать.

Некоторые штрихи к вопросу об интеллектуальности, пусть и не столь существенные, добавляет анализ эмоциональных реакций человека на действия маши-

ны. «Самостоятельные» решения аппарата и программы — автоматические настройки, подсказки, автозапуски подпрограмм и пр. — далеко не всегда вызывают положительные эмоции, даже в случаях, когда эти «инициативы» обусловлены нашими непосредственными действиями. Крайне болезненны для человека ошибки в собственном распознавании, обусловленные схожестью некоторых признаков объектов. Вспомним о типичной реакции на звонок речевого робота АТС, сообщающего о задолженности. Сам факт задолженности — мелочь, но то, что мы по инерции слушаем, отвечаем, полагая, что говорим с человеком, и лишь через мгновение понимая, что на другом конце провода робот, — ничего, кроме досады, не вызывает. Разрабатывая тему отношений «Человек и робот», фантасты — так уж сложилось, — перебрали, наверное, все возможные варианты. В книге или на экране — это одно дело, взгляд со стороны, но представьте, к каким потрясениям приводило бы чувство, возникшее у простого смертного к «искусственной половине», то есть к роботу. Так что не совсем ясно — будет ли человек действительно рад существенной интеллектуализации автоматов и особенно их человекоподобным реализациям.

### «Странная конструкция» (мальчик о девочке, фольклор)

Как известно, элементом «процессора» человека, точнее биологической нейронной сети (БНС), является нейрон. Взаимодействие нейронов осуществляется с помощью серий частотно-модулированных импульсов, причем частота модуляции лежит в диапазоне десятков герц, что более чем в 1 млн. раз (!) ниже быстродействия электронных схем. И несмотря на то что количество передаваемой нейроном информации при таком быстродействии не превышает нескольких десятков бит в секунду, всей нейронной сети требуется всего несколько миллисекунд для решения таких сложных задач, как, например, распознавание зрительных образов. При решении подобных задач «процессор» человека выполняет параллельные вычисления — анализируемая информация не передается по цепи, а распределяется в межнейронных связях. И хотя отдельные элементы «процессора» человека существенно «медленнее» обычных чипов, БНС в целом решает такие интеллектуальные задачи, которые пока не под силу современным компьютерам. Машина фон Неймана и биологическая нейронная система — настолько разные вещи, что сравнение компьютера с мозгом человека, мягко говоря, не совсем корректно. В компьютере, как правило, один сложный и быстродействующий процессор, в БНС «процессоров» очень много, но они простые и медленные. «Память» у человека распределенная, с адресацией по содержанию и интегрирована в «процессор»; у компьютера память отделена от процессора, а адресация не зависит от содержания. Вычисления в компьютере последователь-

ные, выполняются хранимыми программами и касаются численных и символьных операций, а «вычисления» в БНС параллельные, распределенные, базируются на самообучающихся алгоритмах и связаны с восприятием, образами.

### К матери-природе. Возвращение булева сына

Одно из важных мест в исследованиях по ИИ занимают процедуры анализа данных, алгоритмы распознавания и классификации. В интеллектуальных системах такие алгоритмы служат инструментами для принятия решений. Но пока компьютеру еще очень сложно соревноваться с БНС в задачах распознавания. Известно, что во многих сферах внедрение автоматических процессов столкнулось с серьезными трудностями. И вовсе не из-за того, что производительность компьютера недостаточно высока, а потому, что пока мало действительно *интеллектуальных* алгоритмов, обеспечивающих эффективное взаимодействие системы с внешним миром. Обычные компьютерные программы функционируют только в тех пределах, в которых задача была осмыслена и формализована разработчиком. Но время ставит качественно новые задачи, и теперь машина — мощное и удобное орудие — должна *сама* думать, должна стать интеллектуальным партнером, помощником, таким, который, получив команду «подумай и сделай», *поймет* задачу, *сгенерирует* алгоритм решения и *решил* ее. Похоже, что другого пути нет: необходимо учить компьютеры распознавать, запоминать, самообучаться и генерировать алгоритмы, копируя живую природу. Как быстро будет достигнут новый качественный уровень компьютерных технологий на пути заимствования принципов организации и функционирования БНС — покажет будущее. А надежду, что результаты будут, вселяет очередной подъем активности в исследованиях по ИИ. Стремительный рост производительности машин, совершенствование средств ввода и хранения данных позволяют строить модели нейронных сетей даже на обычных машинах, а это, несомненно, стимулирует дальнейшие исследования. Но особые перспективы обещает развитие таких технологий, как VLSI (сверхвысокий уровень интеграции) и оптических аппаратных средств, делающих создание искусственных нейронных систем и моделирование биологических функций вполне реальным делом.

### Все не так уж плохо, а будет еще...

Но все-таки современный компьютер, праправнук машины фон Неймана, — полезная и мощная вещь. Развитие компьютерных технологий привело к глобальной интеграции информации и знаний. Все полезное, пригодное для упорядочения и формализации, обращается в алгоритмы и файлы, становится доступным миллионам людей, развивая и интегрируя сами биологические нейронные системы. Трудно предста-

вить что-либо иное, кроме компьютеров и электронных сетей, что могло бы стать столь же эффективным средством, объединяющим интеллектуальные усилия миллионов и обеспечивающим функционирование по сути коллективного мирового разума.

Уникальность компьютера как орудия труда — что особенно важно для развития человека, — состоит в том, что он позволил делать многое, не оперируя чем-то материальным. Мы не заметили, как были вовлечены в мир математических моделей. Если вне компьютерных технологий пока сохраняется обоняние, вкусовые и тактильные ощущения, то воспринимаемые нами и вполне достоверные зрительные и звуковые образы уже неплохо «живут» в чипах. Мы можем не только оперировать моделями материальных объектов иной раз эффективнее, чем их прототипами, мы создаем образы и ситуации, которые не существуют в природе как реальность. Не всем дана возможность спроектировать и построить реальное здание, расписать своды храма, дирижировать оркестром или воплотить в жизнь финансовую программу компании. Но смоделировать все это, посмотреть и оценить результат, учиться и развиваться почти в реальной обстановке может каждый. И средство для достижения этого — компьютер, который действительно является интегратором информации, знаний и технологий.

### Это не музыка, а это не интеллект!

Развитие электронной музыки очень тесно связано с компьютерными технологиями, а корни взаимодействия этих разных областей уходят в далекое прошлое — в начала математики и музыкальной акустики. Немало параллелей между нотной партитурой, описывающей временной процесс, и текстом программы — блоки и циклы, условия и метки. Кажется, что даже первое гуманитарное применение ЭВМ было музыкальным — инженеры, равнодушные к музыке, заставляли первые машины играть мелодии.

В музыкальных компьютерных технологиях, как и в других областях (например, в компьютерной графике), развивается два подхода. Один связан с управлением параметрической моделью звука, партии, произведения (MIDI-технология), другой — с оперированием аналогом реального объекта (аудиотехнология). Оба подхода имеют свои достоинства и недостатки, постоянно развиваются и нередко дополняют друг друга. Но несмотря на то что в последнее десятилетие в аудиотехнологии сделаны заметные новации, ставшие возможными благодаря росту быстродействия и объема памяти, параметрические модели звуковых объектов остаются более удобными для управления и трансформации. Кажется, что главным недостатком моделирования должно быть несовершенство моделей — неестественность, ненатуральность звуковых объектов. Но достижение правдоподобия в моделировании определяется тем, насколько подробно

модель описывает реальный объект. Такие особенности восприятия, как *пороги достоверности* и *восстановление образов*, играющие существенную роль в процессах распознавания, делают создание правдоподобных моделей вполне достижимым. Подтверждение этому — эксперименты в психологии восприятия, а также бесспорные факты, в частности, такой, как неспособность сегодня не только любителя, но и профессионала отличить смоделированное звучание рояля от настоящего.

В музыкальной электронике применяется уже немало интеллектуальных алгоритмов. В ней порой раньше, чем в других областях, создаются и приживаются новейшие технологии — предмет обязывает. В последних моделях гитарных MIDI-контроллеров используются нейропроцессоры для решения непростой задачи определения высоты тона и приема извлечения, алгоритмы распознавания и обучения используются в считке нотного текста и переводе его в MIDI-файл, а также и в решении обратной задачи — преобразовании MIDI-данных реального исполнения в нотную форму. Методы ИИ находят применение в синтезе звука и в интерактивных системах синхронного оркестрового сопровождения исполнителя на акустическом инструменте. Отмеченные задачи лишь на первый взгляд могут казаться типовыми; в каждой из них своя специфика и свои методы. И нередко эти задачи очень сложны. Если говорить об автоматической нотации, то есть о преобразовании звучащего произведения в нотную форму, то в сравнении с задачей распознавания слитной речи эта задача, для случая оркестрового исполнения, оказывается значительно более сложной. Она не решена даже в первом приближении, и это затрудняет какие-либо прогнозы.

### Кто здесь главный и в чем секрет... атомной бомбы?

Музыка — многогранное и уникальное явление. Как средство эмоциональной коммуникации она является чрезвычайно благоприятной областью для изучения природы человека, его интеллекта и психики. Только музыка касается тонкой эмоциональной сферы, подсознания, используя столь специфические знаковые системы и крайне неточную семантику. Может, поэтому музыка представляет особый интерес для исследований в области ИИ. Но мы не всегда помним о значении музыки в функционировании социума, хотя ее роль вряд ли уменьшилась со времен Древней Греции, когда секвенции имели более определенный смысл и предназначение и когда философы считали инструментальное (без пения) музицирование растлевающим дух и противным богам. Почему музыка — нечто эфемерное — так действительна? Почему современному подростку, когда, как кажется, «весь мир против тебя», так дорог мотив в исполнении боготворимого певца, а его деду — песня времен второй мировой? Почему нам может быть неинтересно тщатель-

Как идеологу метода Performance Modeling мне кажется нескромным добавлять что-то еще к тому, что изложено в статье. Возможно, что сторонние оценки работы коллектива окажутся более интересными для читателя.

«...Теперь мы знаем, что это действительно значительный, мирового класса продукт...»

*Крис Джеймс,  
директор по новым продуктам,  
Microsoft Co.*

«...Я взволнован вашей работой...»

*Грег Хендершотт,  
основатель Twelve Tone Systems*

«... Все, кто смотрел ваш продукт, были так же увлечены, как и я...»

*Мортен Саезер,  
управляющий по продукту,  
Twelve Tone Systems*

«...Это революция в электронной музыке!»

*Давид Барон, инженер,  
Accent Software International*

«...Фантастика! Это действительно путь добиться «реального чувства» в исполнении...»

*Джейсон Силзл, продюсер,  
старший инженер, Dove Digital Media*

«...Я использую секвенсоры 12 лет, но никогда не видел ничего подобного, столь творческого и потрясающего...»

*Стив Орич, композитор,  
тестировщик музыкальных программ*

«...В последние 10 лет ваше программное обеспечение явно открывает новый и захватывающий путь...»

*Крис Муллен, обозреватель Keyboard Player*

«...Я слушал все ваши MIDI-примеры. Они ошеломляющие!»

*Джон Бартелт, композитор,  
тестировщик музыкальных программ*

но спланированное маркетинговое мероприятие, шикарная обстановка, если организаторы «завели» совсем не ту музыку? Почему так легко можно испортить отношения, если слишком откровенно высказаться о чем-то любимом исполнителе?... Для многих музыка — это духовное, эмоциональное «убежище» в нашем далеком от идеала мире. А мелодия, ритмический орнамент, характер исполнения — некоторый «эмоциональный код», поэтому цепочка здесь очень

коротка: исполнитель — музыка — эмоции — душа — своеобразный внемерный переход от материального к идеальному...

Если в театре главное действующее лицо — актер, то в музыке это — исполнитель. Видимо, поэтому к фигуре исполнителя привлечено основное внимание исследователей. Из сложных интеллектуальных и психофизиологических процессов, протекающих при живом исполнении, в компьютерные технологии переносятся пока лишь самые простые. Неудивительно, что музыкальные процессы оказались очень сложными для моделирования. Музыка — особый язык, непередаваемый на язык речи. И сущность ее далеко не в формальной основе, не в нотных символах, укладываемых в частотную и временную сетки; она — в *музыкальной выразительности*.

С формальной точки зрения, выразительное исполнение — это определенное сочетание тонких изменений акустических параметров — частоты, амплитуды, спектра во времени. Но на вопрос, почему они именно такие, нет ответа ни в музыкознании, ни в акустике, ни в психологии. В человеко-машинной системе (исполнитель — инструмент) исполнитель, субъект берет на себя «организующую, управляющую» функцию — от осмысливания идеи композитора, эмоционального проживания и развертывания музыкальной партитуры-программы до выполнения механических манипуляций над инструментом, порождающих саму музыку. Исполнитель как субъективное звено в схеме эмоциональной коммуникации привносит в музыку очень многое. В процессе интерпретации объем «информации», закодированной в партитуре, разворачивается в несоизмеримо больший содержательный объем реального звучания. Именно исполнительским фактором обусловлены и многозначность, «размытость» музыкальных знаков, многоплановость музыкального языка. Исполнитель, образно говоря, «неимоверно усложняет» грамматику нотного текста, «привнося в нее в процессе живого интонирования *грамматику выразительности*» (АМентюков). Своеобразие музыки как языка не только в том, что этот язык включает множество различных грамматик, а в том, что эти грамматики постоянно изменяются. Конечно, это норма для любого развивающегося искусства, но, пожалуй, только в музыке и знаки, и их сочетания столь абстрактны.

Однако *исполнительский фактор* является атрибутом только классической схемы музыкальной коммуникации: *композитор — партитура — исполнитель — звучащее произведение — слушатель*. В компьютерных технологиях создания и воспроизведения музыки исполнитель до сих пор отсутствовал. И действительно, создавая произведение, композитор вводит в машину партитуру, которая затем воспроизводится синтезатором. Где здесь исполнительская интерпретация, где главная фигура, субъект, обеспечивающий выразительность музыки? По словам Д.Оп-

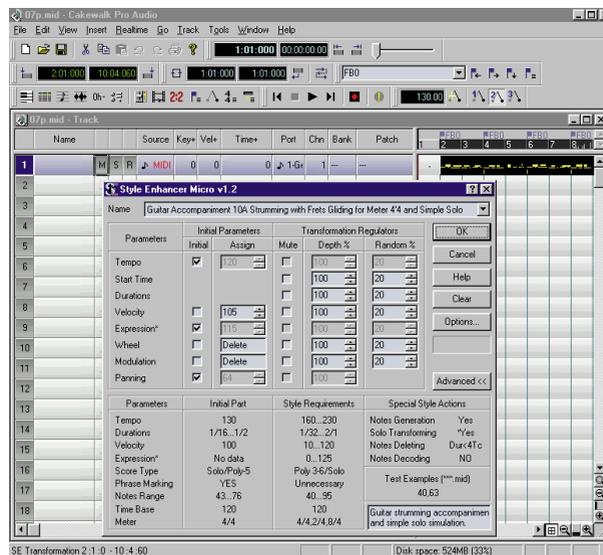
пенгейма, «...компьютер сделает то, что я вею, — ни больше ни меньше. Компьютеры пропускают такие важные аспекты музыки, как выразительность и нюансы; то, что делает одно и то же произведение разным у каждого исполнителя. Мы должны выяснить, как привнести это в компьютерную музыку». Таким образом, теперь, когда в музыкальной электронике в основном решены задачи синтеза и воспроизведения партитур, когда стали обиходными интеллектуальные программы автоаранжировки, генерации мелодий и целых композиций, исследователи направили свое внимание на более сложную проблему — на моделирование исполнения.

Конечно, композитор, создавая произведение с помощью компьютера, вводит в программу определенные исполнительские нюансы. Но его возможности при этом неизмеримо меньше тех, что подвластны живому музыканту. Никто не в состоянии разъять интуитивный, тонко организованный процесс живого интонирования на составляющие, годные для программирования исполнения. Поэтому достижение здесь даже среднего результата требует огромных усилий. Эти обстоятельства привели к поиску нетрадиционных выразительных средств и разработке новых композиторских техник. Произведения, созданные в результате таких поисков, несомненно, изменили взгляды на музыку. Социальное функционирование электронной музыки привело к формированию новых эстетических норм, и она справедливо получила определение «механистичной», «космической», «неживой» — иными словами, бездушной.

## От слов к делу

Но можно ли смоделировать процесс исполнения, используя методы ИИ, избавив тем самым компьютерную музыку от механистичности? Компьютерное моделирование исполнения не может опираться на нестрогие сведения и информацию описательного характера, на то, что накоплено традиционным музыкальным опытом. С какими бы упрощениями ни ставилась задача моделирования, ее решение предполагает формальное описание процесса исполнения и разработку адекватной модели. Иными словами, нужен *формальный критерий выразительности*. Попытки определения такого критерия сделаны в исследовании закономерностей в образцах живого исполнения на основе анализа акустических параметров (М.Клайнс), при разработке модели восприятия и самообучающихся программ исполнения музыки (Г.Видмер). Однако задача моделирования исполнения оказалась настолько сложной, что ни один из частных теоретических подходов не привел пока к заметным результатам. Кажется, что критерий выразительности не может быть определен в разумные сроки, даже в случае когда речь идет о конкретной исполнительской манере и одном музыкальном инструменте.

Выходом из сложившейся ситуации является использование алгоритмов исполнения, задаваемых экспертом, то есть метода «обучения с учителем». Необходима человеко-машинная система, в которой музыкант-эксперт может задавать параметры исполнения и обучать программу в процессе отладки модели. Особенность данного подхода заключается в том, что здесь не нужен критерий выразительности в виде набора логических условий, коэффициентов и акустических параметров. Отлаживая алгоритм, эксперт добивается правдоподобного исполнения, не зная и не интересуясь тем, сколько миллисекунд составляет пауза на взятие дыхания или перенос смычка. Но формальный критерий выразительности *появляет-*



ся в системе, когда эксперт заканчивает отладку модели исполнения. Отлаженная модель — алгоритм исполнения — далее может использоваться композитором, аранжировщиком для создания композиций, «оживления» и «очеловечивания» MIDI-файлов.

Именно такой подход, получивший название Performance Modeling, в течение ряда лет разрабатывается новосибирской фирмой NTONYX ([www.ntonyx.com](http://www.ntonyx.com)). «Последователи Термена из Сибири», как названы специалисты NTONYX в издании Leonardo Music Journal (USA), стали первой исследовательской группой, добившейся практических результатов в моделировании музыкального исполнения (см. также КомпьютерПресс №5'97).

Вслед за первой пользовательской программой моделирования исполнения — Style Enhancer v1.0, выпущенной в 1996 году, на основе Performance Modeling было создано еще несколько продуктов. Последние из них — Style Enhancer Micro 1.2 (SEM1.2) и Style Enhancer Micro 1.28 (SEM1.28) — MIDI Plug-In'ы, разработанные в контакте с Twelve Tone Systems и предназначенные для работы с секвенсорами Cakewalk Pro Audio 6.0, 7.0 и 8.0. Новые продукты — дальнейшее развитие уникаль-

ной технологии NTONYX, — позволяющие придавать MIDI-последовательностям характер звучания живой музыки. С помощью интеллектуальных алгоритмов SEM1.2(8) анализируют MIDI-данные, внося в них, а также изменяя команды контроллеров — Velocity, Volume, Pitch Wheel, Expression и других. Результатом преобразования в одних случаях является имитация игры музыканта на реальном инструменте, а в других — достижение нереального, но интересного характера исполнения. Сравнительный анализ MIDI-файлов, полученных с помощью Style Enhancer, показал, что метод Performance Modeling дает на практике более убедительный музыкальный результат, чем использование гитарных и духовых MIDI-контроллеров или ручное редактирование данных в секвенсоре.

SEM1.2 и SEM1.28 — это простые в использовании, быстрые и эффективные программы. Оба Plug-in'a включают библиотеку из 100 интеллектуальных сти-



**Алексей Устинов** — доцент кафедры компьютеризации Новосибирской консерватории, руководитель NTONYX Computer Laboratory.

лей-алгоритмов преобразования MIDI-данных. В процессе создания библиотеки особое внимание было уделено «трудным» для имитации инструментам, в первую очередь

струнным и духовым. Стили охватывают такие инструменты, как гитара (акустическая, блюзовая, overdriven, distortion), акустический, безладовый, slap-бас, семейство духовых (саксофон, флейта и др.), смычковых, голосовых тембров, ударные и др., а также включают специальные алгоритмы, такие как, например, преобразование Notes into Pitch changes. Теперь пользователю доступна самая полная коллекция стилей для имитации как реальных инструментов, так и нереальных манер. Технология Performance Modeling, интегрированная в MIDI-секцию мощного профессионального секвенсора Cakewalk, дает композитору и аранжировщику — как профессионалу, так и любителю, недоступные ранее уникальные возможности музыкального творчества. ■