

Компьютерная японская поэзия по-башкирски

И компьютерная техника, и компьютерное творчество развиваются давно. Энтузиасты-учёные и программисты уже лет сорок назад учили компьютер писать музыку, рисовать картины, писать стихи и романы. Эстетические достоинства таких экспериментов довольно сомнительны. Интереснее всегда было другое: наблюдать, как причудливо соединяются звуки, линии и смыслы в творениях бездушной машины; пытаться научить компьютер тому, на что, как до XX века считалось, способен только человек, и таким образом самим понять, как искусство создаётся людьми.

В последние десятилетия, когда компьютеры стремительно подешевели и перестали быть принадлежностью исключительно научных институтов, в такого рода экспериментах был сделан внушительный шаг вперёд. Среди бесчисленных примеров компьютерного творчества можно назвать один достаточно экстравагантный и даже имеющий коммерческий успех: норвежский писатель Аудун Мортенсен в 2010 году опубликовал роман под незамысловатым названием «Роман», суть которого в том, что все предложения этого текста взяты из «Лолиты» В.В. Набокова и выстроены в обратном порядке, а также изменены имена собственные оригинала¹. Разумеется, основная работа по созданию этого произведения была выполнена специальной компьютерной программой. Книга имела приличную цену и бодро раскупалась читателями. Не счесть и слухов про того или иного современного российского писателя, будто бы все его тексты на самом деле создаются компьютером.

Компьютеры уже научились кое-что понимать в человеческом тексте, реферировать его², производить тематический анализ, то есть делать вывод «про что здесь написано», составлять в целом похожие на человеческие, хотя порой и не очень осмысленные тексты³. В целом это не так уж и мало.

Но есть одно «но».

Хотя и существуют некоторые универсальные принципы построения такого рода систем, всё-таки каждый раз решения следует создавать для каждого конкретного языка отдельно. Для русского, английского и других языков международного общения таких сервисов, программ и других компьютерных инструментов создано уже порядочное количество, а вот до башкирского компьютерная лингвистика пока не добралась.

Поэтому мы, сотрудники лаборатории компьютерной филологии Башкирского государственного университета, решили сделать какую-то систему, которая умела бы говорить по-башкирски. Во-первых, чтобы привлечь внимание к этой простаивающей, но интереснейшей области: компьютерный синтез речи на башкирском языке. Во-вторых, чтобы показать, что такие системы в принципе возможны, хотя никто до сих пор не пытался их создавать. Конечно, мы хотели и немного развлечься и повеселить других, ведь понятно, что созданные компьютером тексты – это, как минимум, забавно.

Для первой попытки мы избрали жанр японских поэтических трёхстиший, создаваемых по определённым правилам. Первая строка такого стихотворения должна состоять из 5-ти слогов, вторая из 7-ми, а третья снова из 5-ти. Конечно, между хокку и башкирской поэзией большая культурная дистанция. Однако есть один текст, который эту дистанцию сокращает. В одном рассказе уфимского писателя Юрия Горюхина девушка Мауляна волею судеб попадает из башкирского аула в Японию. Её путь сопровождают разнообразные литературные события, вроде встречи с А. И. Гончаровым и А. П. Чеховым, поэтому самый известный японский поэтический жанр не мог пройти мимо истории Мауляны.

¹ <http://www.aftenposten.no/meninger/debatt/article3859167.ece>

² Пример автоматически создаваемых рефератов – сервис Яндекс.Новости: <http://news.yandex.ru/>

³ См. сервис Яндекс.Рефераты: <http://referats.yandex.ru/>

Главный критерий успешности проекта для нас состоял в грамматической правильности порождаемого текста, и именно её нам добиться удалось. Несколько иначе дело обстояло с осмысленностью, но здесь мы и не пытались ничего усложнять. Ведь речь идёт о поэзии, про которую Б. Пастернак писал: «И чем случайней, тем вернее / Слагаются стихи навзрыд». То есть из нерациональных, случайно соединённых слов, по нашей мысли, как раз и должна родиться настоящая поэзия.

Свою программу мы поместили на сайт нашей лаборатории в Интернете. Любой желающий может зайти на страницу⁴ и посмотреть, как это работает. Абсолютно новый текст появляется при каждой перезагрузке и вероятность его повторения крайне мала.

Конечно, мы не ожидали такой реакции от республиканских СМИ. Ведь наша лаборатория сделала много другого интересного: прототип башкирского поэтического корпуса на стихах Мустая Карима, профессиональное электронное издание текстов башкирского фольклорного фонда, до этого хранившегося в шкафу на кафедре, обзор всего башкирского интернета... Тем не менее, всё это журналистов не интересовало. Зато, чтобы узнать и сообщить публике побольше о «компьютерных хокку», нам звонили с радио, писали в газетах, несколько телеканалов рассказали про нашу весёлую разработку в своих репортажах. А в «Комсомольской правде» со свойственными журналистам этой газеты энтузиазмом и ставкой на сенсацию написали, что мы «решили удивить мир». Конечно, никого мы удивлять не собирались, но очевидно, что интерес наша разработка вызвала, так что нужно рассказать о ней подробнее.

Алгоритм случайным образом выбирает слова из словаря программы объёмом примерно в 11500 единиц и встраивает их в один из ритмико-грамматических сценариев, то есть таких способов организации текста, при котором все правила грамматики были бы соблюдены, а строки по количеству слогов соответствовали классической схеме хокку: 5-7-5. Выяснилось, что для башкирского языка таких сценариев не менее 90. Скажем, чтобы все слова ровно уместились в отведённых им ритмом местах, это по одному из сценариев должны быть трёхсложное подлежащее, двухсложное определение, односложное дополнение в основном падеже в первой строке, двусложный послелог, двусложное сказуемое, односложное подлежащее второго предложения, двусложное дополнение во второй строке, двусложный послелог и трёхсложное сказуемое в третьей строке. В таком поэтическом произведении могут использоваться различные падежные формы имён и временные формы глаголов. Причём с неравномерным вероятностным распределением. Это значит, что более естественная для хокку форма настоящего времени будет встречаться чаще, чем прошедшее время, а единственное число чаще множественного.

Какого рода тексты получаются из работы такого алгоритма? Случайность на то и случайность, что невозможно заранее угадать, что увидишь на экране. Например, часто приходится сталкиваться с унылой бессмыслицей, которую трудно даже назвать забавной:

*Тарамыштар йор
шәл һымак итә. Сүл шәл
һымак гәрләшмәй.*

(Сухожилие как быстрая шаль. Пустыня, как шаль, не гремит)

Это необходимый, но, разумеется, далеко не конечный этап того пути, который башкирская компьютерная лингвистика (а с ней и башкирская компьютерная литература) просто пока не успела пройти. Чтобы у компьютера стали получаться более осмысленные и красивые стихи, действительно напоминающие поэзию, нужно ещё многое изучить в версификации настоящих мастеров слова. Нужно многое посчитать (ведь родной для компьютера – язык цифр), выявить сочетаемостные тенденции, подвести лексическую, грамматическую, стиховедческую статистику, и только на этой основе затем писать программы, которые будут сочинять, например, сонеты на башкирском языке.

Пока это не сделано, можно вчитаться в случайные тексты, создаваемые машиной уже сейчас.

⁴ <http://lcp.bashedu.ru/cgi-bin/haiku.html>

*Совинформбюро
уттан алып күмә. Һүз
һын өсөн өркә.*

(Совинформбюро берет из огня и закапывает. Слово для статуи боится)

Анализировать такой текст тем способом, которым это обычно делают литературоведы, было бы странно: ведь филологов интересует «авторское художественное задание» или культурные влияния, которых здесь нет и быть не может по определению. Но если притвориться, будто мы не знаем, что это хокку написано компьютером, то в этом странном и даже немного мистическом тексте можно увидеть намёк на распространённый литературный сюжет об оживающей статуе, реализующий себя, например, в легенде о Дон Жуане. Ушедшее в историю (но и оставшееся в истории) Совинформбюро ассоциируется со сводками с фронта, с голосом войны, а значит, находится в том же понятийном поле, что и огонь (сражения), и закапывание (павших). «Слово» (һүз) – основной инструмент любого информангентства, появляясь во втором предложении, оно скрепляет стихотворение, не даёт ему распасться на две не связанные фразы. Перед нами явно не худший образец компьютерной поэзии.

Получая новые хокку, можно снова набрести на что-нибудь ещё поэтически любопытное:

*Баз кар төргәккә
табан касмай. Кылыу ас
тиклем һыуыны.*

(Подпол к снежному свертку не бежит. Деяние до низу охладилось)

Здесь вполне в духе японских хокку нарисована зимняя сельская картина застывшей жизни. «Снежный» и «охладилась» соединились в этом стихотворении как будто бы нарочно, так же, как ассоциативно связанными между собой оказались и другие ключевые слова: «не бежит», «деяние <...> охладилось», что можно понять в том, смысле, что всё остановилось, замерло, всякое действие (деяние) промёрзло до низу, как полностью, до самого дна, превращается в лёд озеро при сильном морозе. Свёрток – это либо то, что хранится в подполе (из-за «подпола», где, кстати, обычно холодно, ясно, что речь идёт о деревне: в городе на этом месте был бы «холодильник»), либо принявший форму свёртка снежный сугроб.

Совершенно непреднамеренно (и тем удивительнее и «вернее») в одном хокку порой соединяются слова одной тематики:

*Гранатомёттар
тыу кеүек калкынмай. Шаң
шарт менән юймай.*

(Гранатометы не вздымаются, как знамена. Звук не стирает от случая к случаю)

«Гранатомёты», это грузное и непоэтичное слово, конечно, было бы смешным в любом стихотворении, если это не стихотворение о войне. Но здесь именно про войну, важным атрибутом которой являются знамёна. Понять «искусственный интеллект» можно в том смысле, что у войны есть парадная сторона, метонимически отражённая в знамёнах, а есть прозаическая, и они в разных аспектах не равны друг другу. Вторую часть истолковать сложнее, но очевидно, что это продолжение той же вечной для литературы военной темы: «стирание» – это один из поэтических синонимов уничтожения, а звук – постоянный спутник мощного оружия.

В следующем тексте нетрудно увидеть романтический сюжет в духе немецких баллад середины XIX века:

*Өң һәнәрсегә
карай кайыра. Таж сүс
менән арбаны.*

(Нора к охотнику поворачивает. Корона волокном заколдовала)

Охотник, типический персонаж фольклорных историй, попадает в сказочный лес, в котором объекты вдруг начинают двигаться в соответствии со своей волшебной логикой (как оживает в русских сказках избушка, поставленная на «курьи ножки»), а в это время в королевском замке корона неведомым способом захватывает власть над человеком, как в эпопее Толкиена (разумеется, вслед за длинной мифологической традицией германцев) таким колдовским предметом оказывается кольцо повелителя тёмных сил.

Ещё одна зарисовка, на этот раз из жизни в дороге:

*Карауанһарай
айһыз айланмалеккә
табан көрмәлдө.*

(Караван-сарай в безлунную вращательность заплелся)

Безлунная ночь, оживлённая (вращающаяся, «клубящаяся») автомобильная трасса и «вплетённая» в этот пейзаж придорожная гостиница. Соседство суеты и умиротворения, движения и отдыха, спаянность и слитость противоположностей. «Диалектика», как сказали бы философы.

Посмотрим на ещё один случайно порождённый текст:

*Мең кабелселар
күн һымак элә. Нур аһ!
кау саклы бакмай.*

(Тысяча кабельщиков вешает, как кожу. Луч, ах! как трут, не подчиняется)

С одной стороны, здесь слышится недоговоренность. Что и куда вешают кабельщики? Чему не подчиняется луч? Почему его сравнивают с трутом?

С другой стороны, чем это нелогичнее такого аутентичного японского трёхстишия?

*Вишни в весеннем расцвете.
Но я – о горе! – бессилён открыть
Мешок, где спрятаны песни.*

Здесь тоже возникнет масса вопросов, задайся мы целью рационально истолковать стихи. В то же время, если не занудствовать и отдаться поэтическому восприятию, то всё становится на свои места: весна и песня – вполне приемлемое поэтическое соседство. А мешок с песнями – прекрасная метафора.

Если абстрагироваться от знания о том, что лежащие перед нами строки написаны бездушной и пока ещё довольно глупой машиной, которая, в сущности, ничего и не «имела в виду», то стоящий за текстом лирический сюжет в голове у читателя возникнет практически сам собой. Кабель действительно может напоминать кожу – сброшенную кожу змеи, гнущуюся и податливую. Во второй фразе речь идёт о луче света, во всём коже-кабелю противопоставленному: он и светлый (кабель в руках кабельщиков почему-то представляется исключительно в тёмной обмотке), он и прямой, негнувшийся – не подчиняющийся никакому воздействию. Да и сравнение с трутом не выглядит таким уж бессмысленным: этот древний инструмент ещё первобытного человека должен быть достаточно упруг и прям, чтобы быть похожим на луч. Выходит, что компьютер, хотя и случайно, но освоил древнейший композиционный приём – приём противопоставления.

А вот тоже совсем по-японски:

*Берләштереүсе
экскаваторсы һымак
абстракциялай*

(Объединяющий абстрагируется, как экскаваторщик)

Дело в том, что поэтика японских стихов предполагает появление в «кадре» простых работяг, вроде ремесленников или сборщиков риса:

*Сборщик риса вниз трясёт
с высушенной сельди пыль.*

Так и здесь: элемент уже современного российского, а не средневекового японского пейзажа, экскаваторщик, по сути, выполняет ту же функцию, что и сборщик риса у поэта Бонтё. Стоит вспомнить и «кабельщиков» из предыдущего трёхстишия.

А иногда алгоритм изрекает самые настоящие житейские мудрости. Как и любые житейские мудрости, они банальны, но ничего похожего мы от электронно-вычислительного механизма ожидать не могли:

*Интеллигентлык
ишле белдереусегә
күрә бөкрайә.*

(Интеллигентность из-за «знатоков» кривится)

Действительно, разве не сталкивались мы все с особым интеллигентским (далеко не красящим его носителей) снобизмом, направленным на учёность или стремление показать себя «знатоком» в какой-либо области?

Или такая максима, требующая, правда, некоторой расшифровки:

*Яңылык сыйга
хәтле узмай. Сим сарга
табан зураймай.*

(Новость до визга не доходит. Оправа до точильного камня не увеличивается)

Один из остроумных коллег, которому мы показывали эту разработку, даже предложил использовать генератор хокку в качестве оракула – предсказателя событий, изъясняющегося традиционно витиевато, как и должен изъясняться тот, кто видит туманное будущее. Если задаться такой целью, то можно было бы написать и программу, которая бы толковала такие пророчества. Кстати, основной поэтической формой, в которой знаменитый Нострадамус доносил до читателя свои таинственные предсказания, было (как и хокку) особого вида трёхстишие – терцина.

Этот текст, к примеру, очень ясно изложенное предсказание ареста:

*Бәйләусе мишәр
кеүек этә. Тупкен йорт
өсөн яһана.*

(Вяжущий как мишарь толкает. Засов для дома делается)

Мишаря оставим на совести электроники, а вот остальное понятно: «вяжущий» – тот, кто связывает руки арестанту, конвоир, который «толкает» заключённого, таким образом приказывая тому идти вперёд; в этом контексте «засов для дома» просто невозможно прочесть иначе, нежели «тюремный замок». «Делается» означает, что всё это только свершится в будущем, а сейчас просто происходят необходимые мрачные приготовления.

Иногда наш компьютер просто-таки последователь Руссо, будто «забывший», что без технологий, которые он здесь ругает, не было бы его самого:

*Йогошлолоктар
балтасылыктан бирле
күренгеләйзәр.*

(Заразность, начиная с плотничанья появляться начала)

В этом хокку виртуальный автор явно сетует на развитие цивилизации и скучает по первобытным временам, когда жизнь была естественнее, здоровее и ближе к природе.

Играть в эту игру – её можно было бы назвать «пойми непонимаемое» – можно очень долго, ведь немало пройдёт времени до того момента, как хокку станут повторяться.

Конечно, было бы самонадеянно утверждать, будто мы внесли какой-то особый вклад в башкирскую литературу. Мы даже не рассчитываем, что кто-то из профессиональных поэтов воспримет нашу разработку всерьёз (хотя литераторы могли бы обращаться к ней в поисках свежих творческих идей). И это по-своему хорошо: в сущности, мир был бы гораздо лучше, если бы всё настолько же серьёзное и важное, как башкирская компьютерная лингвистика, двигалось к своим целям через такие лёгкие и полшуточные этапы, как наши «автоматические хокку».