

БАЯЧОРОВА Б. Ж., ПАНКОВ П.С.

*КНУ им. Ж. Баласагына, Бишкек,
Институт математики Национальной академии
наук Кыргызской Республики, Бишкек*

BAYACHOROVA B. J., PANKOV P.S.

*J. Balasagyn KNU, Bishkek,
Institute of Mathematics of the National Academy
of Sciences of Kyrgyz Republic*

НЕЗАВИСИМОЕ КОМПЬЮТЕРНОЕ ПРЕДСТАВЛЕНИЕ ОСНОВ КЫРГЫЗСКОГО ЯЗЫКА

Кыргыз тилинин негиздерин компьютерде көз карандысыз чагылдыруу Independent computer presentation of base of Kyrgyz language

***Аннотация:** Одним из главных направлений современной информатики является построение интерактивных компьютерных представлений всех известных реальных и виртуальных объектов, чтобы дать пользователю возможность работать с ними эффективно и безопасно до реального использования. Нами введено понятие: независимое представление – такое, что для него не требуется знание других подобных объектов. Мы поставили задачу полностью независимого представления естественных языков, разработали элементы таких представлений, совместно с соавторами реализовали первую версию алгоритмического языка для этих целей. В этой статье описываем дальнейшее развитие такого подхода.*

***Аннотация:** Азыркы информатикадагы негизги багыттардын бири – бардык чыныгы жана эстетилген объекттердин компьютерде интерактивдүү чагылдырылышын түзүү жана пайдалануу үчүн алар менен иштөө эффективдүү жана коопсуз түрдө боло тургандай шартты камсыздоо. Биздин эмгектерде көз карандысыз чагылдырылыш түшүнүгү аныкталган, бул түшүнүк башка окшош объекттерди билүү зарылчылыгы жоктугу менен аныкталат. Биз табигый тилдерди толугу менен компьютерде көз карандысыз чагылдыруу маселесин коюп, мындай чагылдыруунун айрым элементтерин түзүп, авторлоштор менен бирдикте коюлган максатка ылайык алгоритмдик тилдин биринчи версиясын иштеп чыктык. Бул макалада аталган маселе боюнча изилдөөлөрдүн улантылышы баяндалды.*

***Annotation:** One of main tasks of present day informatics is developing of interactive computer presentations of all familiar real and virtual objects to offer the user the opportunity to master them safely and effectively before real treating. If such computer presentation does not depend on the user's knowledge and skills on similar objects, then we called it independent. We put the problem of completely independent presentations of natural languages. We have developed elements of such presentations, with coauthors have implemented a primary version of an algorithmic language for such purposes. In this paper we describe further development of this approach.*

***Ключевые слова:** кыргызский язык; компьютерное представление языка; независимое представление языка; интерактивное представление языка; алгоритм.*

***Урунттуу сөздөр:** кыргыз тили; тилди компьютерде чагылдыруу; тилди көз карандысыз чагылдыруу; тилди интерактивдүү чагылдыруу; алгоритм.*

***Key words:** Kyrgyz language; computer presentation of language; independent presentation of language; interactive presentation of language; algorithm.*

Введение

Одним из главных направлений современной информатики является построение компьютерных представлений всех известных реальных и виртуальных объектов. Известно, что более эффективным является интерактивное представление объектов, поскольку оно дает пользователю возможность самому выбирать способы изучения объекта.

Нами введено понятие: *независимое* представление – такое, что для него не требуется знание других подобных объектов (в частности, при изучении естественного языка - использование других языков). Тогда пользователь сразу начинает думать на изучаемом языке, без перевода.

Мы поставили проблему - разработать полностью независимое представление кыргызского языка.

Ранее, исследование и изучение живого языка осуществлялось с помощью (включая двуязычные словари и учебники) людей, которые достаточно хорошо знали этот язык; исследование мертвого языка осуществлялось посредством двуязычных текстов, которые сохранились, и имеющихся текстов с дополнительными неявными предположениями и заключениями.

Изобретение записи звуков дало возможность зафиксировать объективно примеры устного языка. Изобретение звукового кино дало примеры связи фраз с ситуациями и действиями. Компьютерные игры дали пользователю возможность выбирать действия с соответствующими фразами. Существовали программные комплексы для изучения языка на основе родного языка, однако некоторые понятия были представлены в них независимо. Этот обзор демонстрирует, что до наших публикаций ранее не существовало полностью независимых представлений естественных языков.

Примечание. Для возможного общения с инопланетянами ранее был разработан язык Линкос. Он начинается рядом натуральных чисел. Затем вводятся знаки чисел и понятие «равняется». Каждый знак передается импульсом особой формы. После этого демонстрируются арифметические операции. Таким образом, вводятся понятия «больше», «меньше», «верно», «неверно», «возрастает», «убывает» и т. д.

Некоторые наши приемы аналогичны приемам этого языка.

Мы [1-10] внесли предложения и разработали элементы таких представлений. Ниже рассматриваем этот вопрос в целом. Мы будем базироваться на кыргызском языке, имея в виду также другие языки. Рассмотрим также обратную связь для того, чтобы проверить знание языка. По нашему мнению, случайная генерация задач и ситуаций [11] необходима для независимого представления понятий и объективной оценки знаний.

Также отметим, что программное обеспечение и оборудование (включая аппаратуру для распознавания речи) для того, чтобы учить и проверять произношение, известны для многих языков и широко используются.

Как известно, невозможно описывать язык средствами этого же языка (мы доказали, что можно изучать язык средствами этого же языка). Мы будем описывать изучение кыргызского языка при помощи русского языка.

1. Гипотезы и определения

Выготский Л.С. и Сахаров Л.С. [15] демонстрировали подобные вещи, имеющие различные признаки, называя их искусственным "языком" (только с существительными и прилагательными) для детей. Если ребенок называл вещи на этом "языке" должным образом, тогда его или ее спрашивали, почему он или она использовал эти слова. Виноград Т. [16] предложил давать команды роботу такими словами как "стол", "коробка", "блок", "пирамида", "шар", "схватить", "передвинуть", "отпустить".

Мы предложили следующее

Гипотеза 1. Подлинное понимание человеком текста на естественном языке может быть проверено посредством наблюдения действий человека в реальных ситуациях, соответствующих этому тексту.

Гипотеза 2. Современное компьютерное оборудование достаточно, чтобы моделировать ситуации, необходимые для обнаружения подлинного понимания многих понятий на естественных языках.

Определение 1. Пусть дано любое «понятие» (слово из языка). Если имеется алгоритм, реализованный на компьютере, который:

- демонстрирует пользователю (случайно генерируемые) достаточно большое количество ситуаций, покрывающих все существенные аспекты "понятия";
- дает команду, включающую это "понятие", в каждой ситуации;
- определяет действия пользователя и представляет их результаты явно на дисплее;

- проверяет, соответствует ли результат выполнения данной команде, то такой алгоритм называется компьютерным диалоговым представлением "понятия".

Примечание 1. Естественно, в данных командах должны присутствовать и другие слова. Но эти слова не должны давать никаких определений или объяснений "понятия".

Примечание 2. Таким способом могут быть представлены не только понятия, соответствующие реальным объектам, но и понятия, соответствующие воображаемым объектам; мы [12-14] продемонстрировали естественное диалоговое представление абстрактных пространств.

Мы предложили общую схему соответствующего программного обеспечения (оно было реализовано под именем *NotiLang*) из трех основных частей:

- "Формализованное подмножество естественного языка",
- "конструктор Задач" и "Игра для обучения и проверки" с этой целью.

В *NotiLang*, определение любого понятия включает некоторые Сущности. Предлагаем рассматривать минимальное число Сущностей как атрибут понятия. Как техническая основа для *NotiLang*, должно быть выбрано некоторое современное удобное в работе оборудование (персональные компьютеры). Математически, мы предложили

- прямоугольник (дисплей) *D*, каждая точка (пиксел) которого имеет цвет, для вывода;
- управление любой точкой на *D*, вместе с дискретным выбором способов (мышь с двумя кнопками) для ввода.

Примечание 4. Современные дисплеи формально дискретны, но они воспринимаются как непрерывные. Так что, мы можем осуществить непрерывное движение.

Примечание 5. Развитие компьютеров должно увеличить возможности коммуникации с ними, но это не повлияет существенно на предложенную классификацию понятий. Когда появятся непосредственные коммуникации между компьютером и мозгом, некоторые Аватарные глаголы станут Прямыми глаголами, и класс Принудительных глаголов расширится. *NotiLang* содержит:

- формализованное подмножество естественного языка;
- утверждения, описывающие окружающую среду: объекты; отношения между ними;
- утверждения, описывающие возможности пользователя;
- утверждения, описывающие возможности объектов (возможные действия);
- утверждения, описывающие условия, чтобы соответствовать высказыванию на естественном языке (во временно'м порядке).

Следующие Сущности необходимы.

Время (как последовательность необходимых действий); *Будущее*; *Прошлое*.

Пространство (виды *Пространства*: *Топологическое Пространство*; *Метрическое Пространство*; *Геометрическое Пространство*) (например, глагол ТАРТ подразумевает только топологию; глагол ТУРТ подразумевает геометрию; прилагательное АЛЫС подразумевает метрику).

(Единственный) *Курсор* (перемещаемый мышью).

Схватывающий Курсор (перемещаемый мышью с левой нажатой кнопкой).

Примечание 6. Согласно традиции программного обеспечения *Windows*, *Схватывающий Курсор* осуществляет параллельное перемещение *Вещи*; иногда это

осуществляется также с вращением; для других математических преобразований могут использоваться *Аватар* и *Инструмент*.

Аватар (объект, управляемый и идентифицирующийся с пользователем) (он также представляется как своего рода курсор).

Вещь (сама *Вещь* - объект, перемещаемый пользователем (со *Схватывающим Курсором*); *Вещь* с другими Объектами - объект, на который воздействует пользователь).

Специальная Вещь (будут описаны или изображены для каждого случая отдельно);

Движущаяся (сама собой) *Вещь*;

Часть (*Вещь* с отношением к другой *Вещи*);

Составная Вещь (составленная из других *Вещей*);

Инструмент (*Вещь*, воздействующая на другую *Вещь*);

Виртуальная Вещь, в том числе *Буквы* (соответствует виртуальному объекту, в отличие от

Вещи, которая имитирует реальный объект);

Место (подмножество дисплея D, или неподвижная *Вещь*);

Специальное Место (будут описаны или изображены для каждого случая отдельно); *Пустое Место* = *Место-для-Вещи* (связано с *Вещью*).

Движение (в форме *Движущейся Вещи* без действий пользователя).

Живое (*Живая Вещь*, которая может чувствовать; между прочим, некоторые устройства, связанные с компьютерами, являются *живыми* в современной речи).

Пара; *Множественность* (некоторые понятия означают пары; некоторые глаголы требуют некоторые однородные *Вещи*).

Тяготение (некоторые глаголы и предлоги требуют это).

Основные *Прилагательные*, типа *Цвета*, *Правый-Левый*, *Большой-Малый*.

Определяется функция *Имя* для всех объектов, кроме *Аватара* и *Курсоров*.

Чтобы пользователь мог догадываться, предложены три метода:

G1) единственность действия (или последовательности действий), которое естественно соответствует команде и ситуации;

G2) подобие (некоторые объекты имеют один и тот же атрибут, и этот атрибут упоминается дважды);

G3) чередование (новое Понятие и новое слово появляются вместе как альтернатива предыдущим Понятиям).

G4) показ Вещи вместе с Именем.

Определение 2. Понятия включают:

DE) Список Объектов;

DP) Список предыдущих Понятий (если это необходимо);

DI) Описание начальной Окружающей Среды (если это необходимо);

DC) Команда (написанная на подмножестве естественного языка);

DS) Последовательность Условий (во временно'м порядке) двух видов: $A \square B ?$ и $A \square B \square \square ?$ с логическими операциями И, ИЛИ, ИЛИ исключающее, НЕ.

Если все Условия выполнены, то выдается сообщение ООБА (Понятие было понято).

Если какое-нибудь Условие нарушено, то выдается сообщение ЖОК.

Кроме того, в режиме обучения, согласно традиции программного обеспечения, Windows:

- если пользователь пробует взять неправильную *Вещь*, то компьютер допускает только ее малый сдвиг и возвращает *Вещь* к ее стартовому положению;

- если пользователь пробует переместить соответствующую *Вещь* на неправильное место, компьютер также возвращает *Вещь* к ее стартовому положению

(число попыток не ограничено). Если дополнительные сущности (например, Тяготение, Притяжение, Отталкивание) задействованы, то компьютер может переместить Вещь в другое (не стартовое) положение.

В режиме тестирования, выдается только сообщение Нет.

В силу Определения 1, некоторые объекты (и иногда их положения) выбираются случайным образом при каждом выполнении Понятия; они будут обозначены (R). Правила грамматики NotiLang не будут детально описаны; они будут понятны в примерах ниже.

2. Примеры Понятий в NotiLang

Мы не будем описывать NotiLang формально. Один из способов осуществлять G1 – это строить такую окружающую среду, что единственное естественное действие является очевидным. Тогда изображения Специальных Вещей и Специальных Мест будут описаны для программиста неформально.

Примеры даются в соответствии с Определением 2 или неформально. Даются еще один из способов догадки (G1, G2 или G3) и изучаемые попутно грамматические формы.

2.1. Виртуальные Вещи - БУКВЫ (G4). Появляются и называются несколько БУКВ (R), после чего

2.2. Глагол КОЙ (G1). ТАБЫШ, БАРЫШ.

DE) Время; Топологическое Место; Курсор; (R) Вещь; (R) Место.

DP) Если используются Буквы, то - Имена этих букв.

DC) Имя (Вещь или Виртуальная Вещь) *НЫ + Имя (Место)*ГА +

"КОЙ!" DS) Схватывание в Вещи. Отпускание в Месте.

2.3. Глагол ТУЗ (G1). ТАБЫШ.

DE) Время; Геометрическое Место; Курсор; (R) Вещь(-без-части); (R) Часть; Пустое-местодля-части.

DC) Имя (Вещь)*НЫ + "ЖАСА!"

DS) Схватывание в Части. Отпускание в Пустом-месте.

2.4. Глагол ТИЙ (G1). ТАБЫШ.

DE) Время; Топологическое Место; Курсор; несколько (R) Вещь; Множественность.

DP) Имена Вещей.

DC) Имя (Вещь)*НЫ + "ТИЙ!"

DS) Курсор находится в Вещи.

2.5. "Начальные существительные": ЧАРЧЫ, УЧ БУРЧТУК, ТЕГЕРЕК вводятся косвенно как (R) Вещи при вводе глаголов ТУЗ или ТИЙ; существительные ЖАЙ, ТАКТА, КУТУ вводятся косвенно как (R) Места при вводе глагола КОЙ.

2.6. Глагол ЖАП (G1). ТАБЫШ.

DE) Время; Топологическое Место; Курсор; (R) Вещь1; (R) Вещь2 (большая, чем Вещь1 и Наложённая по отношению к Вещь1).

DP) Начальные существительные. Шаг 1.

DI) Курсор находится в Вещь2.

DC) Имя (Вещь1)*НЫ + "ЖАП!"

DS) Схватывание в Вещь2. Отпускание в Вещь1 [то есть Вещь1 покрыта Вещь2].

Шаг 2. (также вводится предлог МЕНЕН).

DE) Время; Топологическое Место; Курсор; (R) Вещь1; (R) Вещь2 (большой чем Вещь1); Множественный.

DC) Имя (Вещь1)*НЫ + Имя (Вещь2)+”МЕНЕН_ЖАП!” DS) Схватывание в Вещь2. Отпускание в Вещь1.

2.7. Глагол ТАП (G1). ТАБЫШ.

DE) Время; Топологическое Место; Курсор; (R) Вещи1; (R) Вещь2 (больше, чем Вещи1); Множественность.

DP) Начальные существительные; глаголы ЖАП и ТИЙ.

DI) Краткое описание. Все Вещи1 покрыты копиями Вещь2.

DC) Имя (Вещь1)*НЫ + "ТААП, АНЫ ТИЙ!"

DS) Курсор в Вещь1 [после сдвига Вещи2 с некоторых или всех Вещи1].

2.8. Прилагательные КЫЗЫЛ, ЖАШЫЛ (G2).

DE) Курсор; (R) Красная Вещь; Зеленая Вещь (R); Зеленое Место (R), КОЙ.

DP) Начальные существительные.

DC) "Жашыл»+Имя (Вещь)*НЫ+ "Жашыл» Имя (Место)*ГА+”

КОЙ!" DS) Схватывание в Зеленой Вещи. Отпускание в Месте.

2.9. Глагол ТҮРТ (G1).

DE) Время; Аватар; (R) Вещь; (R) Место.

DP) Начальные существительные.

DI) [Если Аватар касается Вещи, то Вещь движется в противоположном направлении]. DC) Имя (Вещь)*НЫ + Имя(Место)*ГА + "ТҮРТ!" DS) центр Вещи находится внутри Места.

2.10. Глагол СИММЕТРИЯЛА (G1). Сущность трудно охарактеризовать.

DI) Немного асимметричный объект.

DC) " СИММЕТРИЯЛА!"

DS) пользователь должен сделать объект симметрическим.

3. Классификация Понятий в *NotiLang*

С предложенным подходом набор "понятий" полуупорядочивается по предшествованию: из каждых двух "понятий" или одно должно быть введено прежде другого или они могут быть представлены одновременно, или они могут быть представлены независимо.

Первичными будем называть такие слова, что для их понимания не требуется знание других слов изучаемого языка. Они представляются самым естественным или единственно возможным действием в данной ситуации. По нашему опыту, к таким относятся глаголы КОЙ, ЖАП, АЧ, ТҮЗ, КЕЛ, КИР, ЧЫК, КӨРСӨТ, ЧАЛ.

Вторичными будем называть такие слова, что для их понимания требуется знание других слов изучаемого языка или соответствующих действий. Например, глагол ТАП, послелог МЕНЕН.

Мы предлагаем следующую классификацию глаголов (известно, что некоторые глаголы являются полисемантическими и могут попасть в различные разделы). Кроме того, некоторые глаголы могут быть введены с минимальным числом Сущностей, но могут также использоваться с дополнительными Объектами.

Например: ТАШТЫ ТҮРТ!; ТАШТЫ ТАЯК МЕНЕН ТҮРТ!

- Прямые глаголы: с *Курсором* (ЖЫЛДЫР, АЛ, КОЙ, ТАП, КАТ).

- Инструментальные глаголы: с *Инструментом* или другой *Вещью* (БЫЧАК МЕНЕН КЕС, БАЛТА МЕНЕН ЖАР).

- Аватарные глаголы с внешними действиями: (БАР, ЧЫК, КИР, КЕЛ, МИН).

- Аватарные глаголы с «внутренними» действиями: (САНА, ОЙЛО).

Из-за порядка предшествования понятий, переходные глаголы могут быть подразделены на "независимые" от прямых дополнений (ЖЫЛДЫР, АЛ, КОЙ), и "зависимые" (ЖАЗ, БОЁО).

Заклучение

Отметим, что по предложенной нами методике возникает несколько иная классификация наук:

- действия с объектами, как с твердыми телами (механика) (ЖЫЛДЫР, АЛ, КОЙ);
- изменение отношения к объекту (КАРА, АЧ, УК);
- геометрическое изменение объекта (БҮКТӨ);
- топологическое изменение объекта (КЕС, БЕЛ);
- существенное преобразование объекта (здесь объединяются некоторые разделы физики и химия) (ӨРТТӨ).

По данной методике нами разработаны основы независимого электронного учебника и первая версия комплексного экзамена по кыргызскому языку. Построенный комплексный экзамен одобрен Национальной комиссией по государственному языку при Президенте Кыргызской Республики и используется в учебных заведениях.

Надеемся, что предлагаемая нами методика применима для любых существующих языков, т.е. она универсальна, но для ее реализации необходима разработка соответствующих программ. В этом плане возникает множество исследовательских задач для совместного действия программистов с лингвистами.

Список использованных источников

1. Pankov P.S., Aidaraliyeva J.Sh., Lopatkin V.S. Active English on computer // Conference "Improving Content and Approach in the Teaching of English Language in the Context of Educational Reform". - Bishkek, Kyrgyz State Pedagogical University, 1996. – Pp. 25-27.
2. Pankov P.S., Alimbay E. Virtual Environment for Interactive Learning Languages // Human Language Technologies as a Challenge for Computer Science and Linguistics: Proceedings of 2nd Language and Technology Conference. - Poznan, Poland, 2005. – Pp. 357-360.
3. Панков П.С., Баячорова Б.Ж. Компьютерные подмножества естественных языков // Вестник КНУ им. Ж. Баласагына: серия б. Естественно-технические науки. - Выпуск 3. Информационные технологии. - 2005. - С. 5-10.
4. Pankov P.S., Alaeva S.A., Kutsenko V.A. Algorithmic Interactive Presentation of Notions // "Speech and Computer" SPECOM'2006: Proceedings of the 11th International Conference. - Anatolya Publishers, St. Petersburg, 2006. – Pp. 478-480.
5. Pankov P.S., Dolmatova P.S. Algorithmical Language for Computer-Based Presentation of Notions // 4th International Conference on Electronics and Computer. – Almaty: Suleyman Demirel University, 2007. – Pp. 274-279.
6. Bayachorova B.J., Pankov P. Independent Computer Presentation of a Natural Language // Varia Informatica. – Lublin: Polish Information Processing Society, 2009. – Pp. 73-84.
7. Pankov P.S., Bayachorova B.J., Juraev M. Mathematical Models for Independent Computer Presentation of Turkic Languages // TWMS Journal of Pure and Applied Mathematics, Volume 3, No.1, 2012. – Pp. 92-102.
8. Панков П.С., Баячорова Б.Ж. Кыргыз тилин компьютерде көрсөтүү каражаттары // Ж. Баласагын атындагы КУУ жарчысы: Серия 3. Чыгарылышы 4(8), 2009. 9. Панков П., Баячорова Б., Жураев М. Кыргыз тилин компьютерде чагылдыруу. – Бишкек: Турар, 2010. – 172 б.
10. Bayachorova B., Pankov P. Independent Computer Presentation of Abstract Notions of Natural Languages // Actual Problems of Computer Science. - Lublin: ECCS Foundation, No. 1(3), 2013. – Pp. 73-84.
11. Панков П.С., Джаналиева Ж.П. Опыт и перспективы использования комплекса UNIQTEST уникальных тестовых заданий в учебном процессе // Тез. докл. научно-практической конф. «Образование и наука в новом геополитическом пространстве». - Бишкек, 1995.
12. Pankov P.S., Bayachorova B.J. Using computers to perform non-Euclidean topological spaces // The 6-th conference and exhibition on computer graphics and visualization "Graphicon-96", 1996, Vol. 2. - Saint-Petersburg.
13. Pankov P. S., Bayachorova B. J., Terehin A. V. Computer Presentation of Four-Dimensional Spaces // The 8-th International Conference on Computer Graphics and Visualization "GraphiCon-98". - Moscow.
14. Vorubaev A.A., Pankov P.S., Chekeev A.A. Spaces Uniformed by Coverings. - Hungarian-Kyrgyz Friendship Society, Budapest, 2003 (Chapter 4. Constructive and computer presentations of uniform spaces).
15. Выготский Л.С., Сахаров Л.С. Исследование формирования понятий: методология двойного стимулирования // Хрестоматия по общей психологии. Психология мышления. – Москва: изд. МГУ, 1981.
16. Winograd T. Understanding Natural Language. - Massachusetts Institute of Technology, New York, 1972.

Рецензент: Кененбаева Г.М. – доктор физико-математических наук, г.н.с. лаборатории вычислительной математики ИМ НАН КР