

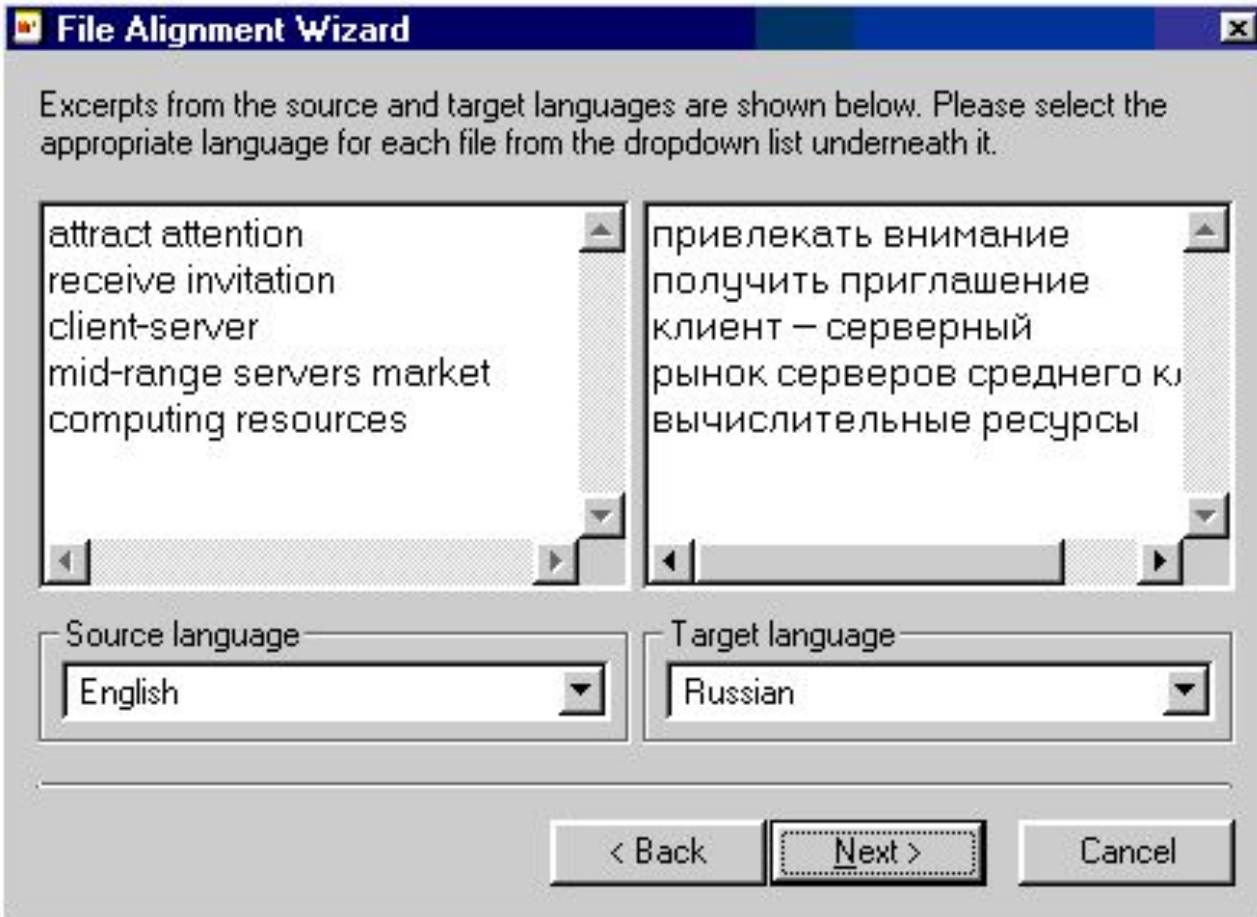
СИСТЕМЫ ПЕРЕВОДА И РАСПОЗНАВАНИЯ ТЕКСТОВ

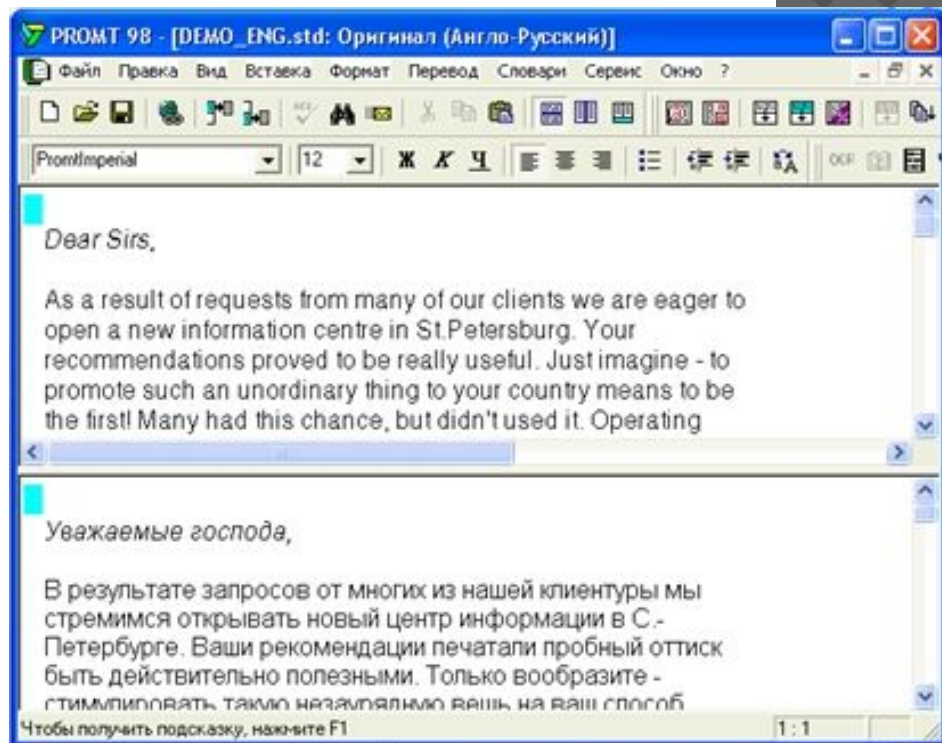
ОСНОВНЫЕ ВОПРОСЫ:

- ⦿ как работают программы-переводчики;
- ⦿ распознавание текста;
- ⦿ ввод в компьютер печатного и рукописного текста.

КАК РАБОТАЮТ ПРОГРАММЫ- ПЕРЕВОДЧИКИ

Чтобы найти перевод неизвестного иностранного слова, пользователю электронного словаря достаточно ввести это слово в строке поиска, и уже через несколько мгновений будет получен перевод. Современные текстовые процессоры имеют в своем составе словари, позволяющие производить орфографическую проверку правильности написания слов (на разных языках).





Рассмотрим простой пример. Переведем с помощью системы перевода на английский язык фразу:

Информатика — это наука об информации.

Результат перевода:

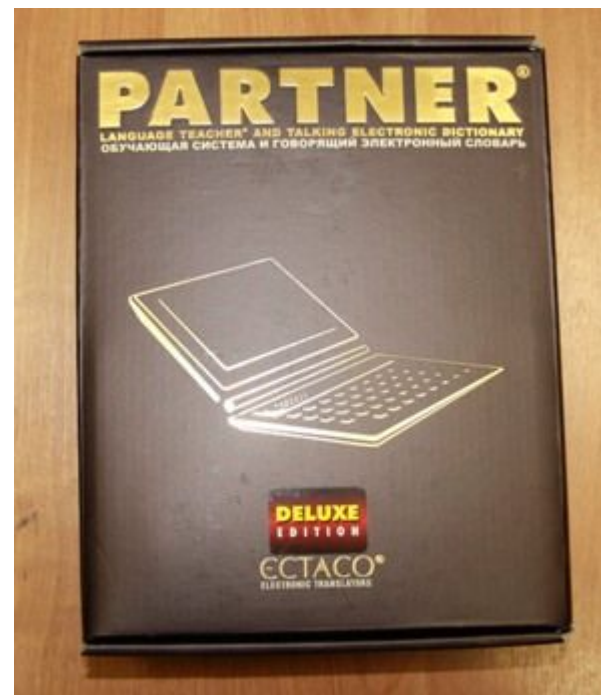
The computer science is an information science.

А теперь с помощью той же программы переведем эту фразу на русский язык. Получим:

Информатика — информатика.

ЭЛЕКТРОННЫЕ УСТРОЙСТВА: ECTACO PARTNER ER 900 DELUX

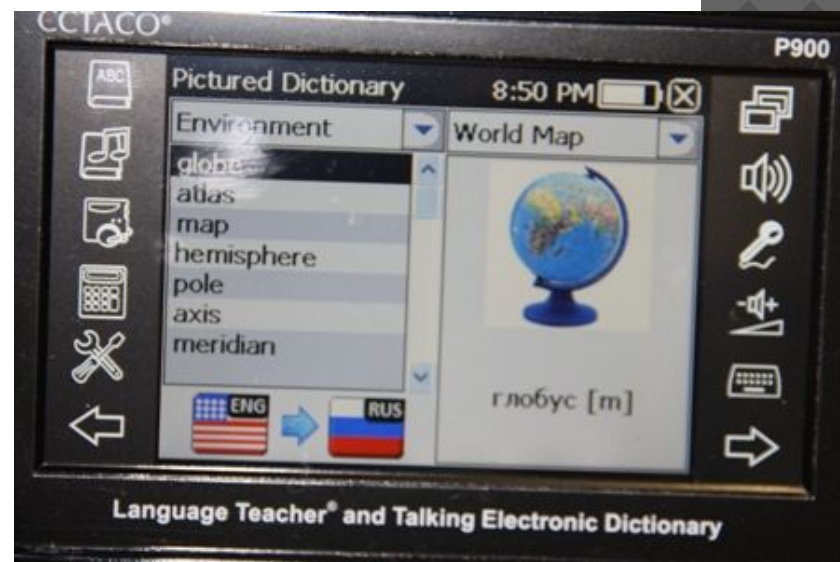
может не только переводить слова и выражения с русского на английский и с английского на русский языки, но и распознавать вашу речь, проигрывать музыку и даже обучать английскому языку! Переводчик оснащён цветным сенсорным дисплеем, может воспроизводить музыку с карты памяти.



КОРПУС И КОМПЛЕКТУЮЩИЕ:



КРУПНЫЕ КНОПКИ, ЦВЕТНОЙ СЕНСОРНЫЙ ЭКРАН



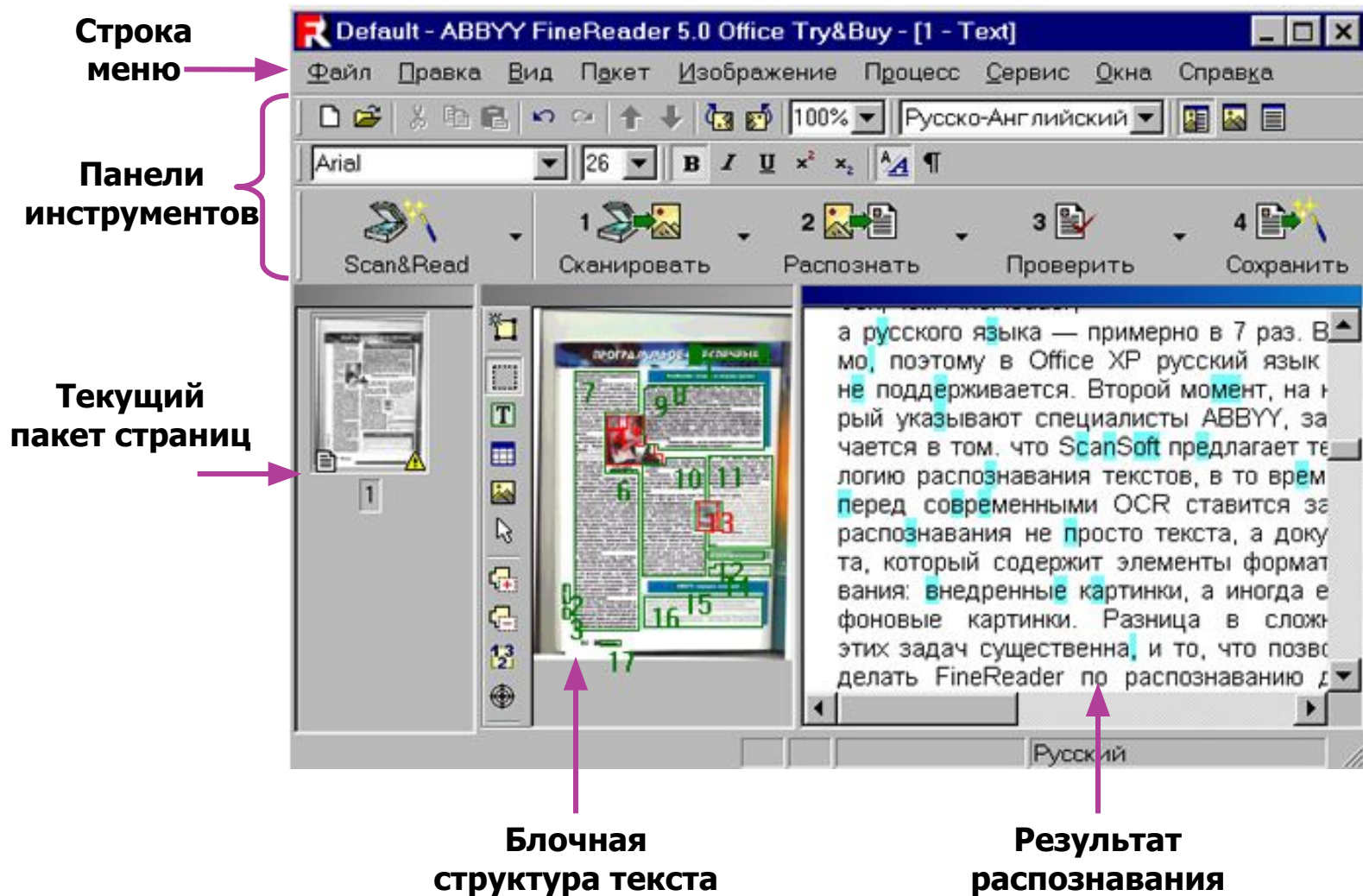
ПРОГРАММЫ ОПТИЧЕСКОГО РАСПОЗНАВАНИЯ ТЕКСТА

Например ABBY Fine Reader

P



ОКНО ПРОГРАММЫ FINEREADER



ПРОЦЕСС ОБРАБОТКИ FINEREADER

- **Сканирование** (сканер, цифровой фотоаппарат, цифровая видеокамера).
- **Сегментация** - выделение блоков на изображении.
- **Распознавание** - неоднозначно опознанные символы выделяются цветом.
- **Проверка ошибок**- можно провести проверку грамматики.
- **Сохранение** результатов в виде отформатированного или неотформатированного документа, или прямой передачи в другое приложение - WORD, Excel в буфер обмена Windows.

СИСТЕМА ОПТИЧЕСКОГО РАСПОЗНАВАНИЯ СИМВОЛОВ

- Если исходный документ имеет **типографическое** качество (достаточно крупный шрифт, отсутствие плохо напечатанных символов или исправлений), то задача распознавания решается методом сравнения с **растровым шаблоном**.

А В В Ф Я

.И...Н

- При распознавании документов с **низким** качеством печати (машинописный текст, факс и т.д.) используется метод распознавания структурных элементов (отрезков, колец, дуг и др.) символов. В искаженном символьном изображении выделяются характерные детали и сравниваются со **структурными шаблонами** символов.

СИСТЕМА ОПТИЧЕСКОГО РАСПОЗНАВАНИЯ ФОРМ

Единый государственный экзамен - 2005
Бланк регистрации

Заполнить гелевой или капиллярной ручкой ЧЕРНЫМИ чернилами ЗАГЛАВНЫМИ ПЕЧАТНЫМИ БУКВАМИ по следующему образцу:
А В В Г Д Е Ж З И Й К Л М Н О П Р С Т У Ф Х Ц Ч Ш Щ Ъ Ы Ь Э Ю Я І І 2 3 4 5 6 7 8 9 0 X V I L

ВНИМАНИЕ! Данный бланк использовать только совместно с двумя другими бланками из данного пакета

Сведения об участнике единого государственного экзамена

Фамилия _____
Имя _____
Отчество _____

Документ Серiya _____ Номер _____ Пол Ж М

Резерв - 1 _____ Резерв - 2 _____ Резерв - 3 _____ Факт выхода из аудитории во время экзамена

ЗАМЕЧАНИЯ участника ЕГЭ по процедуре проведения ЕГЭ.

Заполнение НЕ ОБЯЗАТЕЛЬНО.

Отметьте замечания по проведению экзамена:

<input type="checkbox"/> Организованная доставка участников в пункт проведения ЕГЭ при самостоятельном времени в пути более 1 часа	<input type="checkbox"/> Присутствие в аудитории преподавателей общеобразовательного предмета, по которому проводится ЕГЭ
<input type="checkbox"/> Вскрытие доставочного пакета осуществлялось НЕ в присутствии участника ЕГЭ	<input type="checkbox"/> Наличие нарушений дисциплины в аудитории

FineReader Forms

Единый государственный экзамен - 2005
Бланк ответов № 1

Заполнить гелевой или капиллярной ручкой ЧЕРНЫМИ чернилами ЗАГЛАВНЫМИ ПЕЧАТНЫМИ БУКВАМИ по следующему образцу:
А В В Г Д Е Ж З И Й К Л М Н О П Р С Т У Ф Х Ц Ч Ш Щ Ъ Ы Ь Э Ю Я І І 2 3 4 5 6 7 8 9 0 X V I L

Регистр: Прописные Смешанный

Название предмета _____

Номер варианта _____

Служебная отметка _____

Итого баллов: **102**

ВНИМАНИЕ! Данный бланк использовать только совместно с двумя другими бланками из данного пакета

Номера заданий типа А с выделенными вариантами ответов и вариантами ответов в строке ответов

Обязательный признак [X] Видеть варианты ответов можно только в строке ответов

1	A	B	C	D	1
2	A	B	C	D	2
3	A	B	C	D	3
4	A	B	C	D	4
5	A	B	C	D	5
6	A	B	C	D	6
7	A	B	C	D	7
8	A	B	C	D	8
9	A	B	C	D	9
10	A	B	C	D	10
11	A	B	C	D	11
12	A	B	C	D	12
13	A	B	C	D	13
14	A	B	C	D	14
15	A	B	C	D	15
16	A	B	C	D	16
17	A	B	C	D	17
18	A	B	C	D	18
19	A	B	C	D	19
20	A	B	C	D	20

Задания на свободный ответ по заданию типа В

Резерв - 4	1	2	3	4	5
Резерв - 5	1	2	3	4	5

Результаты выполнения заданий типа В с ответами в краткой форме

B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12

Символы отличных ответов по заданию типа В

- **Бланком** называется стандартный лист бумаги, на котором размещается постоянная информация и отведено место для переменной.
- Сложность состоит в том, что необходимо распознать написанные от руки символы, довольно сильно различающиеся у разных людей.
- Кроме того система должна определить, к какому полю относится распознаваемый текст.

СИСТЕМА ОПТИЧЕСКОГО РАСПОЗНАВАНИЯ ФОРМ

- Для обработки бланков предназначено специальное приложение **FineReader Forms**.
- Для распознавания содержимого бланка необходимо предварительно создать шаблон формы.

Сервис/ Шаблоны

- Шаблон используют на этапе сегментации. Сегментация в данном случае состоит в наложении шаблона.
- Положение шаблона корректируется в соответствии с тем, насколько ровно был размещён бланк при сканировании.
- Заключительный этап состоит в распознавании содержимого бланка.

СИСТЕМА РАСПОЗНАВАНИЯ РУКОПИСНОГО ТЕКСТА

- С появлением первого карманного компьютера Newton фирмы Apple в 1990 году начали создаваться системы распознавания рукописного текста. Такие системы преобразуют текст, написанный на экране карманного компьютера специальной ручкой, в текстовый компьютерный документ.

РАСПОЗНАВАНИЕ РУКОПИСНОГО ТЕКСТА

