

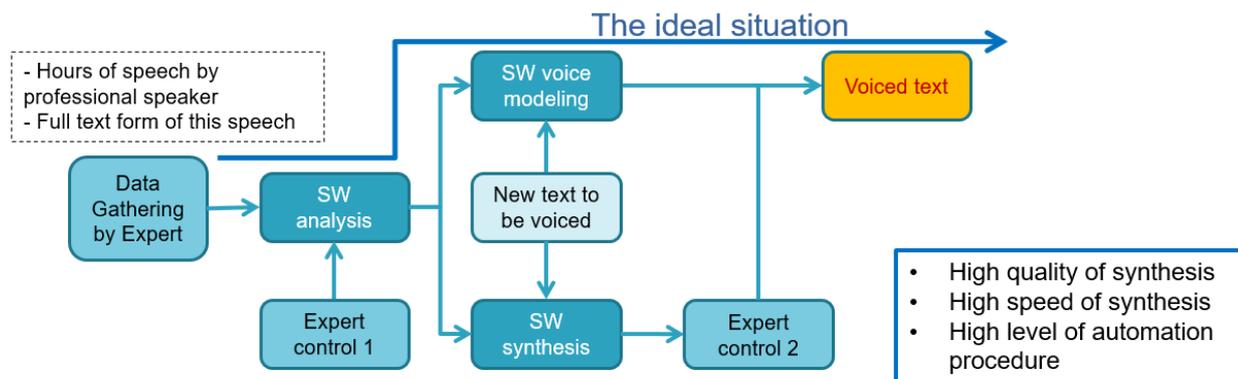
Кассу А-Р.М.

Пример из проектной практики при разработке нового продукта по имитации голоса

(Кейс 2 основной презентации по защите)

Дата разработки 2015-2016 год

1. **Ситуация:** У заказчика криминалистической лаборатории по анализу биометрических голосовых образцов есть потребность во внедрении продукции нового класса, которые расширяют ее делают ее исследовательские возможности уникальными, а именно по синтезу речи.
2. **Краткое описание проблемной ситуации:** Технология должна синтезировать новую речь по голосу диктора с различными на слух параметрами качества от естественной и настоящей записи, но классическая технология синтеза речи на основе построения голосовой модели диктора предъявляет большие требования по качеству исходного аудиоматериала, которые Заказчик не может предоставить.



3. **Ключевая задача:** Найти альтернативный алгоритм/технология синтеза речи по исходным некачественным образцам речи.
4. **Способ решения:** По алгоритму решения изобретательских задач АРИЗ-85-В:

Шаг 1.1. Записаны условия мини-задачи. Описание ТС, ТП-1 и ТП-2.

Техническая система:

(Назначение) Для синтеза речи

(основные части системы) исходный речевой материал, Голосовая модель

Техническое противоречие (ТП 1):

Если обеспечивать высокое качество исходного речевого материала диктора (объем, ОСШ, уровень реверберации и др.) то это позволит использовать доступную технологию автоматического синтеза речи, построив голосовую модель для воспроизведения новой речи (синтезированной), но это приведет к значительному росту привлеченных ресурсов для обеспечения задачи сбора речевого материала высокого качества.

Техническое противоречие (ТП 2):

Если качество исходного речевого материала диктора низкое, это позволит снизить требования к ресурсам по сбору исходного речевого

материала, но не позволит использовать доступную технологию автоматического синтеза для синтеза новой речи.

Шаг 1.2. Выделена и записана конфликтующая пара элементов: Изделие: речь диктора, Инструмент: Технология синтеза

Шаг 1.3. Изображение графических схем ТП-1 и ТП-2 для работы с Заказчиком (опущено)

Шаг 1.4. Выбрана формулировка ТП-2 в которой исключается возможность использования классической технологии синтеза, основанной на наличие высококачественных исходных аудио материалах. Главным производственным процессом является способ обработки исходного аудиоматериала.

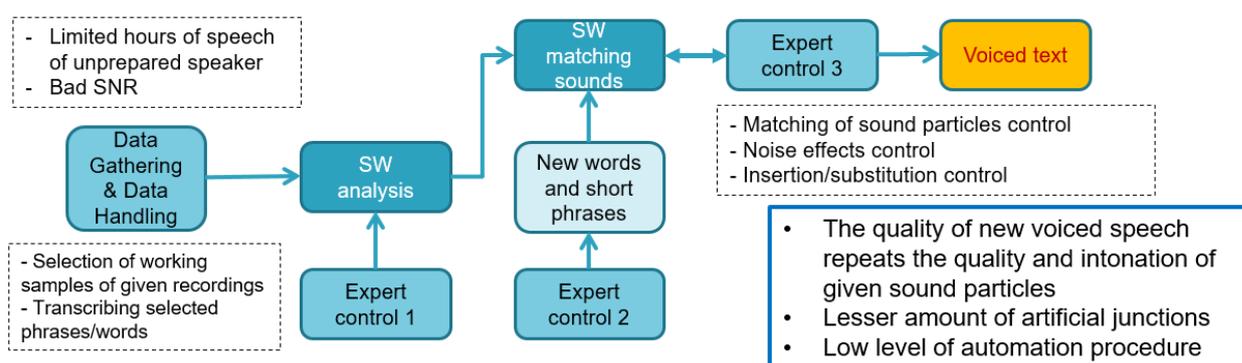
Шаг 1.5. Усилен конфликт: Система синтеза должна позволять генерировать речь любого качества по произвольному тексту диктора исходя из полученного материала с любыми параметрами качества.

Шаг 1.6. Записана формулировка задачи: конфликтующая пара, усиленная формулировка, что должен сделать икс-элемент, чтобы обеспечить разрешение конфликта, например, улучшив качество исходного материала или устранив требования к обеспечению высокого качества.

Шаг 1.7. Проверка возможности получения решения модели задачи применяя системы стандартов. Решение найдено с помощью следующих стандартов:

- *Разделение противоречия в пространстве:* Новая система не использует голосовую модель, являющейся частью пространства БД инструмента.
- *Системный переход:* переход к системе, работающей на микроуровне (Разбиение исходного речевого материала на исходные частицы речи – малые комбинации фонем и слогов).

5. **Результат решения:** Таким образом, получен метод «компилятивного синтеза» с последующей настройкой акустических параметров.



«Если задача не решена, перейти ко второй части АРИЗ. **Если задача решена, можно перейти к седьмой части АРИЗ, хотя и в этом случае рекомендуется продолжить анализ со второй части.**» <https://altshuller.ru/triz/ariz85v-1.asp#7>

Выбираем путь рекомендации.

Шаг 2.1. Определить оперативную зону (ОЗ)

Оперативная зона в данной задаче представляет собой информационное пространство – база данных голосовой модели.

Шаг 2.2. Определить оперативное время (ОВ)

Оперативное время это режим работы (онлайн/оффлайн) для автоматической или экспертной обработки данных

Конфликтное время T1 – время полной автоматической обработки в онлайн (кратковременный)

Время до конфликта T2 – время предобработки данных экспертом в оффлайн режиме

Шаг 2.3. Определить вещественно-полевые ресурсы (ВПР)

1. Внутрисистемные
 - А) ВПР инструмента (Технология синтеза): Автоматический синтез на основе построения голосовой модели, Автоматизированный компилятивный синтез с привлечением эксперта.
 - Б) ВПР изделия (Речь диктора): Объем речевого материала, канал записи речи, средства шумочистки аудиоматериала, средства редактирования аудиоматериала, средства улучшения и адаптации акустических условий
2. Внешнесистемные
 - А) ВПР среды: Средства хранения и организации БД речевых и звуковых частиц, Программный продукт экспертной обработки аудио и речевых данных, Средства визуализации данных, Средства проигрывания и прослушивания, Фильтры обработки шумов и др.
 - Б) ВПР общие: Экспертные навыки, методические материалы, Носители речевой информации, файловые базы и др.
3. Надсистемные: Инструктора-методисты, Процедуры идентификации диктора по голосу.

Допускаем использование микро уровневой структуры изделия (речь диктора) – кластеризацией фонем и частиц.

Шаг 3.1. Записать формулировку ИКР-1:

Икс-элемент, абсолютно не усложняя систему и не вызывая вредных явлений, устраняет в течение ОВ

Сбор частиц аудиоматериала и фонем новой речи производится вызовом из базы данных, их компоновка и полуавтоматическая настройка (с привлечением эксперта) для устранения следов микро монтажа.

Шаг 3.2. Усилить формулировку ИКР-1 дополнительным требованием:

Использовать частицы и фонемы только исходного речевого материала.

Шаг 3.3. Записать формулировку физического противоречия на макроуровне:

Извлекая частицы и фонемы из соответствующей БД (ОЗ) в течении T1 (ОВ) должна осуществляться их склейка, чтобы формировать новообразованную речь, но не должны склеиваться, т.к. остаются следы возникновения дефектов микро монтажа.

Шаг 3.4. Записать формулировку физического противоречия на микроуровне: В местах склейки частиц и фонем (ОЗ) производится фильтрация и балансировка акустических условий в автоматическом режиме T1 (ОВ), чтобы устранить эффект расхождения акустических условий, но не такая обработка не должна быть, т.к. может привести к повреждению самих частиц и фонем (повреждение АЧХ речевого сигнала).

Шаг 3.5. Записать формулировку ИКР-2:

Склейка частиц и фонем с обеспечением фильтрации и балансировки должна сама обеспечивать сохранность от повреждений АЧХ речевого сигнала.

Шаг 3.6. Проверить возможность применения системы стандартов к решению физической задачи, сформулированной в виде ИКР-2.

- *Разделение противоречия во времени:* Новая система может работать в оффлайн режиме. Т.е. допускается экспертная предобработка синтезированной речи для обеспечения должного уровня качества нового материала, при которой эксперт проводит прослушивание и тюнинг узких мест вовремя склейки, имея доступный широкий спектр визуальных и инструментальных средств программного обеспечения для эффективной работы.

На этом, решение задачи найдено (удовлетворяет требования заказчиков) и можно перейти к седьмой части АРИЗ.

Шаг 7.1. Контроль ответа. Можно ввести уже имеющийся набор инструментарий и программных средств экспертной шумоочистки, настройки интонаций и эмоций к речевому материалу.

Шаг 7.2. Провести предварительную оценку полученного решения.

а) Обеспечивает ли полученное решение выполнение главного требования ИКР-1 ("Элемент сам...")?

- Да, Программное продукт позволяет эксперту в заданное пользователем время производить речь из доступного речевого материала.

б) Какое физическое противоречие устранено (и устранено ли) полученным решением?

- Устранено физическое противоречие, что система может синтезировать речь из исходного материал плохого качества, благодаря разделения процесса обработки во времени между автоматическими средствами и ресурсами эксперта, а также благодаря разделения в пространстве, где вместо новые данные формируются не из базы данных голосовой модели а из БД частиц и фонем.

в) Содержит ли полученная система хотя бы один хорошо управляемый элемент? Какой именно? Как осуществлять управление?

- Да. Управляемый элемент: качество склейки частиц, управление автоматизировано.

г) Годится ли решение, найденное для "одноцикловой" модели задачи в реальных условиях со многими циклами?

Да. Апробация проводилась на производстве речи разной длины и содержания.

Шаг 7.3. Проверка на новизну проводилась анализом решений в данной области среди продукции компаний конкурентов.

Шаг 7.4. Возникшие задачи в результате данного решения: 1) Необходимость формирования методического материала для обучения экспертов навыкам работы в режиме обработки новой речи. 2) Защита синтезированной речи для задач обеспечения биометрической идентификации диктора по голосу.

Данные задачи не являются сложными и можно было устранить доступными протоколами и ресурсами. (Шаги 8 и 9 опущены).