FAR Glass

CS 361: Mid-term Assignment

Andrea Tongsak
Ariel Meshorer
Cole Theodore
David Korotky
Iain Moncrief
Joshua Minyard
Robert Vincent

Due: Saturday, November 6, 2021 at 6:00 p.m., on Canvas

> Those who can imagine anything,
> can create the impossible.
>
> *Alan Turing*

Contents

1 General Description

FAR (Field Augmented Reality) Glass is a **knowledge transfer system** which serves as an **intelligent cognitive assistant**.

Users of FAR are professionals—mechanics, medics, and other specialists—who are expected to maintain and pull from an ever-expanding skill set of complex tasks and procedures in order to do their jobs. In most cases, the basic training for these professionals can be as short as a few months to a few weeks. The remainder of their skills is acquired through years of apprenticeship under those with more experience. FAR is designed to bridge this *experience gap*, i.e. the gap between a worker's fundamental skill set the wide range of unexpected possibilities they have to face every day in the real world.

When focusing on the knowledge transfer aspect of the system, users become—in addition to the professionals themselves—the trainers, supervisors, and managers of those professionals. In other words, users are not only those who seek to benefit from the knowledge *in* the FAR system but also those who seek to pass on knowledge to others *through* the FAR system. In the following, we shall refer to the former as **workers** and the latter as **managers**.

The knowledge transfer system's main functions are

(1) extracting **task instructions** from checklists, illustrated manuals, and training videos;

(2) translating task instructions into computer-interpretable **task graphs**; and

(3) providing content for **just-in-time guidance**.

Workers interface with FAR through a wearable optical display ergonomically as similar as possible to an ordinary pair of eyeglasses. The display functions as a low-interference augmented reality headset which seamlessly presents instructions and guidance to the worker in an intuitive and organic manner, effectively transforming the viewer's entire work space into a platform for digital information.

Managers, on the other hand, experience a *far* simpler interface. In short, FAR's knowledge transfer system is almost entirely automated. Managers simply connect the FAR glasses to their computers via USB and upload their instructional materials of choice. The moment that the files have been completely transferred to the device's hard drive, FAR begins its work of translating the instructional materials into its own language, the *task graph*, a complete data representation of all the logical steps that are involved in the correct performance of a task. Thus, the input to FAR comes from the field itself: the reams of educational and training material both already available and to be created in service of a profession. The output of FAR, the task graph, is a kind of digital *checklist* under the aegis of which the system monitors and guides the worker as they traverse the graph from start to finish.

2 Usage Scenarios

2.1 Chef

A professional chef wants to teach his sous-chefs how to prepare a dish. The chef has previously recorded a video of himself walking through the preparation process for this particular dish.

The chef uploads the video to each sous-chef's FAR glass via USB. Like any other USB device, the chef sees FAR as just another USB mounted drive. The moment the file finishes transferring to the hard drive, FAR begins to transform the video into a task graph. FAR is able to map each step of the video to a node in a task graph, so that every step of the preparation of the meal is encoded into a computer-interpretable format.

The chef now gives the FAR glasses to his sous-chefs, and FAR begins to provide instructions on the preparation of the meal. The sous-chefs can now prepare the meal without the professional chef's direct supervision.

2.2 Semiconductor Company

Kim, the CEO of an IT semiconductor company, wants to ensure that there is higher accuracy in the assembly of their products. Since there are multiple steps involved in building an integrated circuit, Kim decides that FAR would help her production team make less mistakes.

Kim connects each individual FAR glass using a USB port to her computer where she has multiple checklists stored in PDF format. The checklists outline the steps in the manufacture and assembly of the circuits. Kim transfers multiple PDF documents from her computer into the FAR database. The moment the files finish transferring to the FAR hard drive, FAR begins to transform the checklists PDFs into task graphs.

Within FAR, the task graphs are then indexed according to the connections between them. Each stage of the circuit manufacturing process, although registered as a separate task, is linked to the next stage's task graph. Thus, if Kim wanted to walk through the entire manufacturing process from start to finish, she can choose to traverse the entire linked task graph from start to finish.

After her production team uses the FAR system for training, Kim notices a higher accuracy and faster times when they are working on the assembly product. Thus, FAR helped her company reduce overhead costs from defective circuits.

2.3 Home Improvement Company

The CEO of a home improvement company needs to update the illustrated manuals in the company's FAR system. The company already uses FAR to train its employees on how to repair appliances, but all of their appliances have been re-designed since the last instructions were uploaded to FAR. The appliances can no longer be repaired with the old instructions, so the company decides to upload new illustrated manuals.

The CEO connects the FAR glasses via USB to his computer and pulls up the computer's file browser. He opens the drive to see a list of files. These files are the old illustrated manuals which the CEO had previously uploaded into FAR. The CEO deletes all of these files, since they are deprecated. The moment he does this, the corresponding task graph stored in FAR is also deleted and no longer able to be referenced.

The CEO now drags the new, up-to-date illustrated manuals into the FAR folder. The moment the manuals have completed transferring to the hard drive, FAR begins to transform the new illustrated manuals into task graphs.

The FAR system now has completely new task graphs for the illustrated manuals for the re-designed appliances. The next day, the employees of the company come to work and receive instructions pertaining to the new appliances.

2.4 Commercial Kitchen

A commercial kitchen is responsible for preparing the dishes for dozens of menus. Recently, the kitchen has adopted the FAR system in order to train and assist their cooks. As a result, they have uploaded and converted hundreds of recipes into task graphs.

Giuseppe, a newly hired cook, has been tasked with making a curry-based dish. As part of the recipe, Giuseppe must caramelize onions. However, Giuseppe, fresh out of culinary school, does not remember how to caramelize onions, and the FAR assistant has so far only mentioned the words "caramelize onions" as an instruction.

Giuseppe wants clarification on caramelizing onions. Giuseppe says, "FAR, expand caramelizing onions". As part of the data uploaded to FAR, the task database already contains a separate sub-recipe for caramelized onions. This recipe has already been indexed so as to be referenced by every recipe that calls for caramelized onions. Thus, FAR is able to easily extend the task graph for the main dish without interrupting the progress made in the main dish. (Note: this interaction with FAR is not under the scope of our focus; we merely include it in order to demonstrate the underlying knowledge transfer process with the database).

FAR begins to instruct Giuseppe in the preparation of caramelized onions. Giuseppe follows the instructions given to him faithfully and completes the caramelized onions. FAR then switches back to the main recipe and continues assisting from where it left off.

The usage scenario above demonstrates the interlinked nature of the task database. Tasks are organized in a manner so that they are available for lookup both by name and also by need. Tasks that are duplicated are removed so that multiple recipes which utilize caramelized onions can merely reference a single caramelized onion task graph which is then called whenever clarification is needed or the recipe is requested alone.

2.5 Factory

A factory prides itself on manufacturing hand-made goods. Recently, there has been an increase in refunds to their customers due to defects in their products. The company decides to adopt FAR in order to help their employees improve the quality of their products.

Julia is in charge of quality control at the factory. She is knowledgeable about all aspects of the manufacturing process and the areas where quality can be improved. She writes a new set of written procedures for the hand-assembly process with certain keywords highlighted in red. FAR has been designed so that, once these new procedures are uploaded to this system, the red highlighted keywords will be marked as important. The FAR system now has the information to, later on, be able to warn the employees whenever a particular red-highlighted step is being performed.

Julia uploads the new procedures to FAR via USB. The moment the files finish transferring to the hard drive, FAR begins to transform the documents into task graphs. FAR recognizes that the new procedures are very similar to the old procedures but merely contain new red highlights and several new steps for quality control at the end. FAR is able to recognize that this is a new version of the original procedure and replaces the old task graph with the new one.

The next day, the employees, now equipped with FAR glass, receive the new instructions Julia uploaded the previous evening and no longer make the mistakes that were causing the defects earlier on. After the latest batch of products ship, Julia notes an 90% reduction in refunds to customers.

3 User Stories

1. As a user, I need to edit existing task graphs by uploading new versions/deleting old versions of human-readable instructions so that the system can stay up to date.

2. As a user, I need to index task graphs so that when they are needed (beyond knowledge transfer) I can freely navigate through them.

3. As a user, I need to upload an illustrated instruction manual to the task database so that the system can give my workers relevant assistance.

4. As a user, I need to upload a training video to the task database so that the system can provide my colleagues relevant assistance.

5. As a user, I need to upload a checklist to the task database so that the system can provide me relevant assistance.

4 Logical Architecture

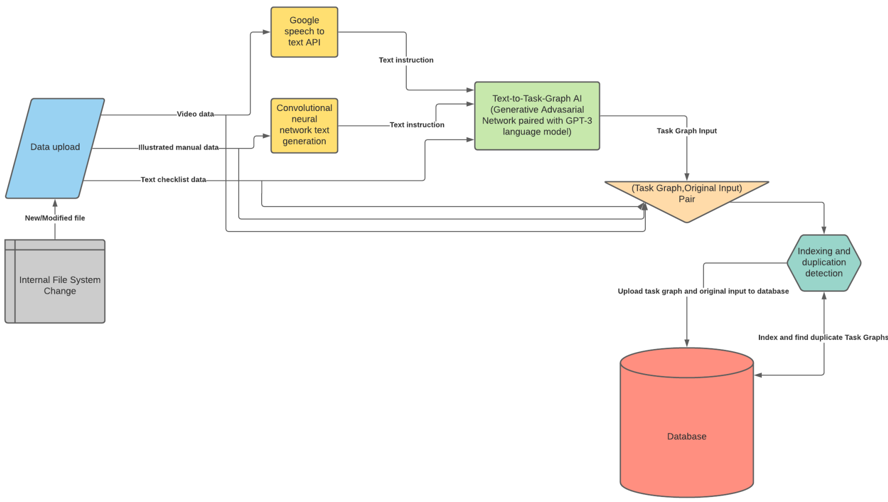

Figure 1: Logical Architecture diagram

4.1 Top level software components

4.1.1 Data Upload

This component uploads human instructions in the form of checklists, illustrated manuals, and videos to the database. Checklists will be read from .pdf, .txt, and .csv files. Illustrated manuals will have one or more picture instructions in a single .pdf file, and all uploaded videos will be converted to mp4 using pre-existing conversion software like UniConverter. Uploaded data will be immediately processed by Speech-to-Text AI and Image processing AI.

4.1.2 Speech-to-Text AI

This AI processes the audio of videos uploaded to the database to create a text script of instructions. It will be developed using one of the Speech-to-Text client libraries provided by Google Cloud.

4.1.3 Image Processing AI

This AI recognizes objects in images and creates text to describe the images. It would use Vision AI from Google Cloud to recognize objects in pictures, and it would use an algorithm derived from Microsoft's Seeing AI to describe them.

4.1.4 Text-to-Task-Graph (T2G) AI

This AI will receive the processed instructions from the Speech-to-Text AI and the Image Processing AI. It uses IBM Watson to order the text and parse it into incremental steps, then uses a GAN (Generative Adversarial Network) to generate the task graph. It then structures these steps into a task graph. This task graph is stored in the database where it will be accessed by the Real-Time Guidance AI.

4.1.5 Indexing & Duplication Detection

This algorithm would receive any new, updated, or deleted task graph and original input pair. It would then search through the database and remove or merge any duplicate information that's added. It would compute the hamming distance between any existing pair and the one to be added, then if they are within a certain range, the new information would not be inserted.

4.1.6 Database

The raw instruction data and the system-created task graph are stored in the database. The user can explore the database to see all of its contents in files organized to their instruction set. The database is accessed by the Data Upload and Text-to-Task-Graph AI.

5 Physical Architecture

5.1 Identification

FAR Glasses

5.2 Purpose

The purpose of FAR Glasses are the visual and interactive interface between the AR and the user in completing everyday physical tasks. For this reason, FAR Glasses not only need to be lightweight, affordable, and attractive, but simple and intuitive to use.

5.3 Received Software Component

The FAR Glasses will receive all software components, as it is the primary hardware component for the system. The Data Upload component will pass the uploaded data from the user's computer system into RAM. The CPU calls the Google Speech to Text AI and the Image Processing AI to process the instructions in the RAM. The CPU then calls the Text-to-Task-Graph (T2G) AI, which processes the text output from the Speech to Text and Image Processing AI. The T2G sends the task-graph output to the Indexing & Duplication Detection software. The Indexing & Duplication Detection software will send the new, if modified or removed, task graph to the storage. When the user activates the intelligent assistant, the task-graph is moved to the RAM.

5.4 Technical Characteristics

5.4.1 Commercial Elements

- Size and Weight
  - Weight: 22 g
  - Size: 120mm x 60mm x 25mm
- Glass/Prescription Options
  - Thin glass
  - Options to customize the color of frames and prescription lens
  - Display embedded in lenses (640x360p sRGB)
- Water Resistant/Dust Resistant and Drop Resistant
  - IP67

5.4.2 Technical Elements

- Chip/CPU
  - Qualcomm Snapdragon XR1
  - ARM architecture
  - AI accelerator
- RAM
  - 16 GB RAM LPDDR5
- Battery
  - USB-C Charging Port
  - Polymer Li Battery in the form of a battery pack
- User Curation
  - Microphone and tap-responsive
- Connectivity
  - Networking card and antenna for wireless connectivity
  - USB connection for direct access to storage
- Storage

- 512GB eMMC flash storage for software and user entered data

- Camera
  - 5 megapixel
- Speaker
  - 2x2 Watt speakers

6 Technical Justification

The FAR system fulfills the foreseen requirements for implementing Knowledge Transfer. Technical choices include uploading instruction data via USB, converting human instructions into computer interpretable task graphs with AI components, and merging hardware and software to support our proposed solution for Knowledge Transfer. Justification for each choice is described below.

6.1 Uploading data from a computer via USB

Uploading data into the FAR Glasses is a vital feature that provides the system with enough context to guide the human user during tasks. By uploading via USB directly to the FAR glasses, the system prevents potential security risks that come with using a cloud.

6.2 Using AI to process data

Since three different forms of human instruction will be input to the system, the system must have multiple AI components to convert those instructions into text to create task graphs. The T2G AI takes text as input and generates a task graph representing the input. It uses IBM Watson to order the text and parse it into incremental steps, then uses a GAN (Generative Adversarial Network) to generate the task graph.

The Image Processing AI converts illustrative manuals into text. It will use Vision AI from Google Cloud to recognize objects in pictures, and an algorithm derived from Microsoft's Seeing AI to describe them in text.

The Speech-to-Text AI will use Speech-to-Text client libraries provided by Google Cloud. By using both the Image Processing AI and the Speech-to-Text AI, the system can build instructions from audio and video sources. Data in the form of a checklist is already in text, so the T2G AI handles them directly. The CPU supports these AI components with an AI accelerator.

6.3 Storing user data in FAR

The system uses the hardware components of RAM and eMMC flash storage as the physical location for the database. This allows for rapid retrieval of data for real-time guidance.

6.4 Software Engineering Practices

The project uses reuse-oriented engineering by applying pre-existing AIs and APIs such as IBM Watson, Microsoft's Seeing AI, and machine learning libraries from Google Cloud. The project applies the Scrum methodology; requirements can be analyzed at the beginning of each sprint, and each user story is implementable in a single sprint.

7 Requirements Traceability

		Architecture						
		Data upload	Internal file system	Google speech to text API	Convolutional neural network engine	Text-to-task graph AI	Indexing and deduplication detection	Task Database
User Requirements	Modify existing task graphs by uploading new versions/deleting old versions of human-readable instructions.	x	x					x
	Index task graphs for navigation.						x	x
	Upload an illustrated instruction manual to the task database.	x	x		x			x
	Upload a training video to the task database so that the system.	x	x	x				x
	Upload a checklist to the task database so that the system.	x	x			x		x

Figure 2: Shows the traceability between the architecture and the user requirements.