



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Σχεδίαση και ανάπτυξη ενός εικονικού περιβάλλοντος διαλογικής αλληλεπίδρασης για εκμάθηση των ελληνικών

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

ΠΑΠΑΔΟΠΟΥΛΟΥ ΔΗΜΗΤΡΙΟΥ

Επιβλέπων: Αλέξανδρος Ποταμιάνος
Αναπληρωτής Καθηγητής

Αθήνα, Νοέμβριος 2022



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

**Σχεδίαση και ανάπτυξη ενός εικονικού
περιβάλλοντος διαλογικής αλληλεπίδρασης για
εκμάθηση των ελληνικών**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

ΠΑΠΑΔΟΠΟΥΛΟΥ ΔΗΜΗΤΡΙΟΥ

Επιβλέπων: Αλέξανδρος Ποταμιάνος
Αναπληρωτής Καθηγητής

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 15 Νοεμβρίου 2022.

(Υπογραφή)

(Υπογραφή)

(Υπογραφή)

.....
Αλέξανδρος Ποταμιάνος
Αναπληρωτής Καθηγητής

.....
Αθανάσιος Κατσαμάνης
Κύριος Ερευνητής Ι.Ε.Λ, Ε.Κ. ΑΘΗΝΑ

.....
Στέφανος Κόλλιας
Καθηγητής

Αθήνα, Νοέμβριος 2022



Copyright © - All rights reserved. Με την επιφύλαξη παντός δικαιώματος.
Δημήτριος Παπαδόπουλος, 2022.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα.

Το περιεχόμενο αυτής της εργασίας δεν απηχεί απαραίτητα τις απόψεις του Τμήματος, του Επιβλέποντα, ή της επιτροπής που την ενέκρινε.

ΔΗΛΩΣΗ ΜΗ ΛΟΓΟΚΛΟΠΗΣ ΚΑΙ ΑΝΑΛΗΨΗΣ ΠΡΟΣΩΠΙΚΗΣ ΕΥΘΥΝΗΣ

Με πλήρη επίγνωση των συνεπειών του νόμου περί πνευματικών δικαιωμάτων, δηλώνω ενυπογράφως ότι είμαι αποκλειστικός συγγραφέας της παρούσας Πτυχιακής Εργασίας, για την ολοκλήρωση της οποίας κάθε βοήθεια είναι πλήρως αναγνωρισμένη και αναφέρεται λεπτομερώς στην εργασία αυτή. Έχω αναφέρει πλήρως και με σαφείς αναφορές, όλες τις πηγές χρήσης δεδομένων, απόψεων, θέσεων και προτάσεων, ιδεών και λεκτικών αναφορών, είτε κατά κυριολεξία είτε βάσει επιστημονικής παράφρασης. Αναλαμβάνω την προσωπική και ατομική ευθύνη ότι σε περίπτωση αποτυχίας στην υλοποίηση των ανωτέρω δηλωθέντων στοιχείων, είμαι υπόλογος έναντι λογοκλοπής, γεγονός που σημαίνει αποτυχία στην Πτυχιακή μου Εργασία και κατά συνέπεια αποτυχία απόκτησης του Τίτλου Σπουδών, πέραν των λοιπών συνεπειών του νόμου περί πνευματικών δικαιωμάτων. Δηλώνω, συνεπώς, ότι αυτή η Πτυχιακή Εργασία προετοιμάστηκε και ολοκληρώθηκε από εμένα προσωπικά και αποκλειστικά και ότι, αναλαμβάνω πλήρως όλες τις συνέπειες του νόμου στην περίπτωση κατά την οποία αποδειχθεί, διαχρονικά, ότι η εργασία αυτή ή τμήμα της δεν μου ανήκει διότι είναι προϊόν λογοκλοπής άλλης πνευματικής ιδιοκτησίας.

(Υπογραφή)

.....
Δημήτριος Παπαδόπουλος

15 Νοεμβρίου 2022

Περίληψη

Από τη μία, τα τελευταία χρόνια οι εφαρμογές εικονικής πραγματικότητας φαίνεται να κερδίζουν όλο και περισσότερο χώρο σε διάφορους τομείς της καθημερινότητάς μας. Ένας από αυτούς τους τομείς είναι και η εκπαίδευση και μάλιστα ο κλάδος της που αφορά την εκμάθηση γλωσσών. Μάλιστα το πεδίο των "virtual reality assisted language learning" εφαρμογών μετράει ήδη κάποια πρώτα βήματα και δείχνει ότι τραβάει όλο και περισσότερο το ενδιαφέρον με τις πιο δημοφιλείς από αυτές τις εφαρμογές να έχουν αρκετούς χιλιάδες χρήστες. Από την άλλη, μας ενδιαφέρουν οι διαλογικοί βοηθοί ή αλλιώς chatbots, τα οποία είναι ευφυή συστήματα που μπορούν να αναπτύξουν συζητήσεις σε φυσική γλώσσα και σε πραγματικό χρόνο. Από επιχειρήσεις μέχρι δημόσιες υπηρεσίες, από online πλατφόρμες εξυπηρέτησης πελατών μέχρι την αυτόματη τηλεφωνική υποστήριξη οι διαλογικοί βοηθοί εντοπίζονται σε πολλές εφαρμογές επικοινωνίας ανάμεσα στον άνθρωπο και τον υπολογιστή. Αντίστοιχα, chatbots έχουν χρησιμοποιηθεί και από κάποιες εφαρμογές εκμάθησης γλωσσών. Στα πλαίσια της παρούσας εργασίας, αφού έγιναν μια θεωρητική προσέγγιση των παραπάνω εννοιών και μια επισκόπηση των εφαρμογών σχετικών με την εκμάθηση γλωσσών, ορίστηκε το βασικό σημείο αναφοράς για μια καινοτόμα εφαρμογή εκμάθησης γλωσσών εικονικής πραγματικότητας με την παρουσία διαλογικών βοηθών. Στη συνέχεια, σχεδιάστηκε μια τέτοια εφαρμογή, αναπτύχθηκε μια πρωτότυπη έκδοση αυτής και μετά από μια διαδικασία αξιολόγησης προτάθηκαν μελλοντικές επεκτάσεις. Μέσω της εφαρμογής, ο χρήστης θα μπορεί να εξασκηθεί στην ελληνική γλώσσα συνομιλώντας με ανθρώπινους χαρακτήρες στα πλαίσια ρεαλιστικών σεναρίων εμπνευσμένα από την καθημερινότητα. Προς εξυπηρέτηση αυτού του σκοπού, εκτός από τις τεχνολογίες εικονικής πραγματικότητας και των διαλογικών βοηθών, επιστρατεύτηκαν και συνδυάστηκαν κι άλλες καινοτόμες τεχνολογίες, όπως είναι η αυτόματη αναγνώριση ομιλίας (ASR) και η σύνθεση φωνής (TTS).

Λέξεις Κλειδιά

εφαρμογή εκμάθησης γλωσσών, εικονική πραγματικότητα, διαλογικοί βοηθοί, αυτόματη αναγνώριση ομιλίας, σύνθεση φωνής, VRALL, ελληνική γλώσσα

Abstract

On the one hand, in recent years virtual reality applications seem to be gaining ground in various areas of our everyday life. One of these areas is education, and in particular the area of language learning. In fact, the field of "virtual reality assisted language learning" applications is already counting some first steps and seems to be attracting more and more interest with the most popular of these applications having several thousand users. On the other hand, we are interested in interactive assistants or chatbots, which are intelligent systems that can develop conversations in natural language and in real time. From businesses to public services, from online customer service platforms to automated telephone support, interactive assistants can be found in many human-computer communication applications. Similarly, chatbots have also been used by some language learning applications. In this diploma thesis, after a theoretical study of the above concepts and an overview of applications related to language learning, a baseline for an state-of-the-art virtual reality language learning application with the presence of interactive assistants was defined. Subsequently, such an application was designed, a prototype version of it was developed and after an evaluation process future extensions were proposed. Using the application, the user will be able to practice Greek language by communicating with human characters in the context of realistic scenarios inspired by everyday life. To serve this purpose, in addition to virtual reality and interactive assistants, other innovative technologies, such as automatic speech recognition (ASR) and speech synthesis (TTS), have been employed and combined.

Keywords

language learning application, virtual reality, conversational assistants, automatic speech recognition, speech synthesis, VRALL, Greek language

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον συμφοιτητή μου Παναγιώτη για τη σημαντική συνεισφορά που είχε στα πλαίσια της πρακτικής μας άσκησης, τους καθηγητές Νάσο και Ιωάννα για την έμπνευση και την καθοδήγηση που μου προσέφεραν και τον συνάδελφο Κοσμά για τη βοήθεια που μου παρείχε όποτε τη χρειάστηκα. Ευχαριστώ ακόμη όλα τα κοντινά μου πρόσωπα για την στήριξη που μου προσέφεραν, αλλά και την υπομονή που έδειξαν αυτούς τους μήνες. Τέλος, θα ήθελα να ευχαριστήσω όλους τους ανθρώπους που στη μέχρι τώρα πορεία μου φρόντιζαν με τον έναν ή τον άλλον τρόπο να μου υπενθυμίζουν πως η γνώση δεν πρέπει να είναι ούτε αυτοσκοπός, αλλά ούτε και μέσο προσωπικής επιτυχίας. Ας είναι ένα εργαλείο για να κάνουμε τον κόσμο μας καλύτερο.

Αθήνα, Νοέμβριος 2022

Δημήτριος Παπαδόπουλος

Περιεχόμενα

| | |
|---|-----------|
| Περίληψη | 1 |
| Abstract | 3 |
| Ευχαριστίες | 5 |
| 1 Εισαγωγή | 11 |
| 2 Θεωρητικό Υπόβαθρο | 13 |
| 2.1 Μια σύντομη εισαγωγή στην Εικονική Πραγματικότητα | 13 |
| 2.2 Εικονική πραγματικότητα και εκμάθηση γλωσσών | 18 |
| 2.3 Διαλογικά Συστήματα | 23 |
| 2.4 Διαλογικά συστήματα και εκπαίδευση | 27 |
| 2.5 Σχετικές εφαρμογές | 33 |
| 3 Σχεδίαση | 37 |
| 3.1 Baseline μιας VRALL εφαρμογής και τι θέλουμε να βελτιώσουμε | 37 |
| 3.2 Χαρακτηριστικά της VRALL εφαρμογής | 40 |
| 3.2.1 Κύριες λειτουργίες | 40 |
| 3.2.2 Εκπαιδευτικές μέθοδοι | 41 |
| 3.2.3 Τεχνικά χαρακτηριστικά | 42 |
| 4 Υλοποίηση | 43 |
| 4.1 Επιλογή των εργαλείων ανάπτυξης | 43 |
| 4.2 Εικονικό περιβάλλον της εφαρμογής | 46 |
| 4.2.1 Δημιουργία του χώρου του σεναρίου | 46 |
| 4.2.2 Κίνηση του χρήστη στο χώρο | 47 |
| 4.2.3 Διεπαφές με το χρήστη | 49 |
| 4.3 Ανθρώπινα avatars | 51 |
| 4.3.1 Σχεδιασμός των κινήσεων στο πρόσωπο | 51 |
| 4.3.2 Τα animations στο Unity | 55 |
| 4.3.3 Υλοποίηση της κίνησης | 56 |
| 4.4 River - VA backend | 59 |
| 4.5 Διαλογικό σύστημα | 62 |
| 4.5.1 Εισαγωγή στο Rasa | 62 |
| 4.5.2 Υλοποίηση του διαλογικού μοντέλου | 66 |

| | |
|---|----------------|
| 4.6 Παραγωγή των hints | 70 |
| 4.7 Στήσιμο του client - συντονισμός του frontend | 75 |
| 5 Αξιολόγηση και Συμπεράσματα | 79 |
| 5.1 Φάσεις αξιολόγησης | 79 |
| 5.1.1 Διαμορφωτική αξιολόγηση | 79 |
| 5.1.2 Πιλοτική αξιολόγηση | 80 |
| 5.1.3 Αποτελέσματα | 82 |
| 5.2 Αποτίμηση και μελλοντικά βήματα | 86 |
| Παραρτήματα | 91 |
| Α΄ Ερωτηματολόγια | 93 |
| Α.1 Ερωτηματολόγιο χρήστη | 93 |
| Α.2 Ερωτηματολόγιο Υπευθύνου | 95 |
| Βιβλιογραφία | 100 |
| Συντομογραφίες - Αρκτικόλεξα - Ακρωνύμια | 101 |

Κατάλογος Σχημάτων

| | | |
|------|--|----|
| 2.1 | Το φάσμα του εικονικού συνεχούς [1] | 15 |
| 2.2 | Ιστορικές φωτογραφίες εφαρμογών εικονικής πραγματικότητας | 17 |
| 2.3 | Η τυπική αρχιτεκτονική ενός διαλογικού συστήματος | 26 |
| 2.4 | Framework σχεδίασης ενός εκπαιδευτικού διαλογικού βοηθού. | 32 |
| 4.1 | Οι χώροι της reception του ξενοδοχείου | 47 |
| 4.2 | Ο collider του OVRPlayerController | 48 |
| 4.3 | Τα παράθυρα διεπαφής με τον χρήστη. | 50 |
| 4.4 | Μερικές εκφράσεις προσώπου από το MakeHuman. | 51 |
| 4.5 | Μερικές εκφράσεις προσώπου που αντιστοιχούν σε φωνήματα. | 52 |
| 4.6 | Προσαρμογή των πλεγμάτων στο πρόσωπο μέσω Βλενδερ. | 52 |
| 4.7 | Κινήσεις φωνημάτων μέσω των βλενδοσηπες στο Unity. | 53 |
| 4.8 | Η αρχιτεκτονική του TCN μοντέλου για την αναγνώριση φωνημάτων. | 54 |
| 4.9 | Η ροή λειτουργίας του συστήματος Mecanim. | 55 |
| 4.10 | Τα μέρη του humanoid avatar στο Unity. | 56 |
| 4.11 | Animator Controllers και mask. | 58 |
| 4.12 | Η αρχιτεκτονική της εφαρμογής. | 59 |
| 4.13 | Η επικοινωνία στο εσωτερικό του River. | 60 |
| 4.14 | Η επέκταση της επικοινωνίας στο River για τα hints. | 61 |
| 4.15 | Το pipeline του Rasa. | 64 |
| 4.16 | Παράδειγμα ροής ενός pipeline. | 65 |
| 4.17 | Το pipeline που υλοποιήθηκε. | 69 |
| 4.18 | Επικοινωνία meta_channel και rasa server για την παραγωγή των hints. | 74 |
| 5.1 | Αποτελέσματα σχετικά με τη VR εμπειρία | 82 |
| 5.2 | Στοιχεία UI σχετικά με την εκμάθηση γλωσσών | 83 |
| 5.3 | Αποτελέσματα σχετικά με το διαλογικό σύστημα | 84 |
| 5.4 | Αποτελέσματα σχετικά με το motion sickness | 84 |
| 5.5 | Αποτελέσματα σχετικά με τη λειτουργία των hints | 85 |

Κεφάλαιο 1

Εισαγωγή

Η εκμάθηση γλωσσών είναι μια διαδικασία που ξεκινά από την αρχή της ανθρώπινης ζωής και συνεχίζεται σε όλη της τη διάρκεια. Η γλώσσα είναι ίσως το πιο χρήσιμο εργαλείο που έχουμε για να επικοινωνούμε τις σκέψεις μας, τα συναισθήματα και τις εμπειρίες μας, για να συνάψουμε σχέσεις με άλλους ανθρώπους και να μπορούμε να βάζουμε σε μια σειρά το πως αντιλαμβανόμαστε τον κόσμο. Πέρα από τους παραπάνω βασικούς λόγους, στο σύγχρονο κόσμο η γνώση περισσότερων από μία γλώσσες έχει αναδειχθεί ως ένα πολύ σημαντικό εφόδιο. Εκτός από το ότι σου προσφέρει περισσότερες ευκαιρίες να βρεις εργασία, το να μάθεις μια δεύτερη γλώσσα μπορεί να σε εισάγει και σε νέες κουλτούρες. Έτσι, θα είσαι πιο έτοιμος και με περισσότερη αυτοπεποίθηση να ταξιδέψεις στον κόσμο και να εξερευνήσεις τον τρόπο ζωής άλλων ανθρώπων. Με δεδομένα όλα αυτά, τις τελευταίες δεκαετίες η εκμάθηση γλωσσών έχει γίνει μια αρκετά δημοφιλής ενασχόληση ανεξάρτητα από την ηλικία, τις γνώσεις, το πολιτισμικό υπόβαθρο, τα ενδιαφέροντα και άλλα χαρακτηριστικά των ανθρώπων που την επιλέγουν. Συγκεκριμένα για την περίπτωση της ελληνικής γλώσσας οι λόγοι που καθιστούν τη γνώση της χρήσιμη έχουν αυξηθεί τα τελευταία χρόνια για δύο κύριους λόγους. Αν είσαι μετανάστης πρώτης ή δεύτερης γενιάς το να γνωρίζεις ελληνικά είναι πολύ σημαντικό εφόδιο για το ζήτημα της εργασίας, αλλά και για την καλύτερη εξυπηρέτηση των καθημερινών σου αναγκών. Αν είσαι επισκέπτης ή τουρίστας στην Ελλάδα η γνώση της γλώσσας σου δίνει τη δυνατότητα να αλληλεπιδράσεις με τους ανθρώπους με πιο φυσικό τρόπο και να έρθεις σε επαφή με την ντόπια κουλτούρα καλύτερα.

Δύο τεχνολογίες που μπορούν να διευκολύνουν, αλλά και να εμπλουτίσουν σημαντικά τη διαδικασία εκμάθησης γλωσσών είναι αυτές των "διαλογικών βοηθών" (conversational agents ή αλλιώς chatbots) και της εικονικής πραγματικότητας. Ένας σύγχρονος διαλογικός βοηθός μπορεί πλέον να ακούει και να καταλαβαίνει τη φυσική γλώσσα, να απαντά με ανθρώπινη φωνή, να προσαρμόζει την εμφάνιση και τα χαρακτηριστικά του για να ταιριάζει με το θέμα συζήτησης και να χρησιμοποιεί γκριμάτσες και εκφράσεις για να κάνει το διάλογο πιο φυσικό. Ο μακροπρόθεσμος στόχος της τεχνολογίας αυτής είναι οι διαλογικοί βοηθοί να είναι τόσο ρεαλιστικοί που θα μπορούν να έχουν μνήμη, να σκέφτονται και να μιλούν όπως ένας άνθρωπος. Χαρακτηριστικά παραδείγματα σύγχρονων διαλογικών βοηθών τέτοιου τύπου είναι η Siri και η Alexa που επιτρέπουν σιγά σιγά στους ανθρώπους να ελέγχουν και να επικοινωνούν με τα έξυπνα σπίτια, τα αυτοκίνητα και τον ευρύτερο κόσμο που μας περιβάλλει. Από την άλλη, οι τεχνολογίες εικονικής πραγματικότητας κατασκευάζουν περιβάλλοντα που περιέχουν ρεαλιστικές σκηνές και αντικείμενα. Τα περιβάλλοντα αυτά

παρουσιάζονται στον χρήστη μέσω ειδικών συσκευών (βλ. headsets) και έχουν σκοπό να του δημιουργούν την εντύπωση εμπύθισης στον εικονικό χώρο που τον περιβάλλει. Εκτός από τα βήματα που γίνονται στο κομμάτι που αφορά την προβολή και την αναπαράσταση των εικονικών κόσμων, σημαντική πρόοδος έχει σημειωθεί και στο επίπεδο της αλληλεπίδρασης με αυτούς και έχουν να κάνουν με την κίνηση αλλά και διάφορα ερεθίσματα που αναγνωρίζουν οι ανθρώπινες αισθήσεις (δόνηση, θερμοκρασία, ρεύμα αέρα, υγρασία κλπ). Όπως και στην περίπτωση των διαλογικών βοηθών, έτσι και σε αυτήν της εικονικής πραγματικότητας η κατεύθυνση που ακολουθείται είναι να δημιουργηθούν εικονικά περιβάλλοντα με τόσο υψηλά επίπεδα εμπύθισης που θα προσομοιάζουν την πραγματικότητα. Ένα παράδειγμα εφαρμογής, που επιχειρεί να πετύχει κάτι τέτοιο είναι το Metaverse.

Όμως, πρέπει να γίνει πιο σαφής ο τρόπος με τον οποίο οι δύο αυτές τεχνολογίες θα μπορούσαν να συμβάλουν στον σκοπό της εκμάθησης γλωσσών. Γενικά, οι συμβουλές που δίνουν οι ειδικοί για πιο γρήγορες κι αποτελεσματικές μεθόδους όταν μαθαίνεις μια γλώσσα αφορούν τη λήψη περισσότερων ρίσκων και τη χρήση της σε όσες το δυνατόν περισσότερες περιστάσεις μέσα στη μέρα. Ακόμη, σημαντική θεωρείται η προσπάθεια να βρίσκεσαι σε περιβάλλοντα που ομιλείται η γλώσσα αυτή και να τη χρησιμοποιείς σε συνθήκες αλληλεπίδρασης με άλλα άτομα κι όχι απομόνωσης, ώστε να λαμβάνεις άμεσο feedback από έμπειρους χρήστες της και να εξοικειώνεσαι με το να κάνεις λάθη. Αυτά είναι τα σημεία στα οποία ουσιαστικά έρχονται οι δύο τεχνολογίες που συζητάμε και εισάγουν νέα στοιχεία. Για παράδειγμα, διαλογικοί βοηθοί που συνοδεύονται από κατάλληλα διαμορφωμένα περιβάλλοντα εικονικής πραγματικότητας μπορούν να βοηθήσουν τους νέους χρήστες της γλώσσας να αποκτήσουν περισσότερη αυτοπεποίθηση συνομιλώντας με χαρακτήρες που σε συνήθεις συνθήκες δε θα μπορούσαν να αλληλεπιδράσουν. Η συγκεκριμένη ιδέα έχει ήδη αρχίσει να εφαρμόζεται με διάφορες σχετικές εφαρμογές που είναι διαθέσιμες προς χρήση, όπως η Mondly VR και η ImmerseMe να σημειώνουν επιτυχία. Ενώ, όμως, η δημοτικότητα και η θετική ανταπόκριση που λαμβάνουν αυτές οι εφαρμογές από μαθητές κι εκπαιδευτικούς αποτελούν πολύ ενθαρρυντικά σημάδια, ταυτόχρονα αναδεικνύονται ελλείψεις που φαίνεται ότι υπάρχουν και τρόποι που οι δύο αυτές τεχνολογίες και η σύνδεση τους με την εκμάθηση γλωσσών θα μπορούσαν να βελτιωθούν. Υπάρχει, δηλαδή, ακόμη πολύ γόνιμο έδαφος για προσθήκες και καινοτομίες προς την βελτίωση αυτής της εκπαιδευτικής τεχνολογίας.

Στα πλαίσια της παρούσας διπλωματικής εργασίας γίνεται μια τόσο θεωρητική, όσο και πρακτική προσέγγιση των εφαρμογών εικονικής πραγματικότητας και διαλογικών βοηθών. Αναδεικνύονται τα θετικά αποτελέσματα από τις ήδη υπάρχουσες εφαρμογές, γίνεται μια ανάλυση των μεθόδων και των στρατηγικών χρήσης αυτών των τεχνολογιών, εντοπίζονται οι αδυναμίες και οι περιορισμοί που υπάρχουν και γίνονται κάποιες προτάσεις για το μέλλον. Στη συνέχεια, σχεδιάζεται μια εφαρμογή εικονικού περιβάλλοντος και διαλογικής αλληλεπίδρασης για εκμάθηση των ελληνικών και αναπτύσσεται ένα prototype της εφαρμογής αυτής. Τέλος, γίνεται μια αποτίμηση της προσπάθειας αυτής και σκιαγραφούνται οι προσθήκες, οι βελτιώσεις και τα επόμενα βήματα που μπορούν να γίνουν.

Κεφάλαιο 2

Θεωρητικό Υπόβαθρο

Αυτό το κεφάλαιο επιχειρεί καταρχήν να εισάγει τον αναγνώστη στα επιστημονικά αυτά πεδία που αφορούν είτε άμεσα, είτε έμμεσα την εφαρμογή που θα αναπτυχθεί. Για κάθε ένα από αυτά τα πεδία πραγματοποιείται μια καταγραφή του ερευνητικού και του πρακτικού επιπέδου στο οποίο βρίσκονται αυτή τη στιγμή. Τέλος, γίνεται μια ιστορική αναδρομή γύρω από εφαρμογές αντίστοιχες με τη δική μας, από την οποία εξάγονται χρήσιμα συμπεράσματα. Οι ενότητες που αποτελούν αυτό το κεφάλαιο είναι οι εξής:

1. Μια σύντομη εισαγωγή στην Εικονική Πραγματικότητα
2. Εικονική πραγματικότητα και εκπαίδευση
3. Διαλογικά συστήματα
4. Διαλογικά συστήματα στην εκπαίδευση
5. Σχετικές εφαρμογές

2.1 Μια σύντομη εισαγωγή στην Εικονική Πραγματικότητα

Τι ορίζεται ως εικονική πραγματικότητα

Μπορούμε να βρούμε διάφορους ορισμούς για την εικονική πραγματικότητα, αλλά ίσως ο πιο καθολικός από αυτούς να είναι ο εξής: "Ως εικονική πραγματικότητα ορίζεται ένα περιβάλλον, είτε αυτό είναι πραγματικό, είτε αποτέλεσμα προσομοίωσης στο οποίο ο παρατηρητής έχει την αίσθηση ότι βρίσκεται κάπου αλλού" [2]. Η αίσθηση αυτή μεταφράζεται στα αγγλικά ως "telepresence". Ο ορισμός αυτός που δόθηκε από τον Jonathan Steuer φαίνεται ότι προτιμάται σε πολλές περιπτώσεις, γιατί καταφέρνει να δώσει μια εξήγηση της έννοιας χωρίς να αναφέρεται στο ρόλο της τεχνολογίας.

Από την άλλη, ένας ορισμός που χρησιμοποιεί έννοιες της σύγχρονης τεχνολογίας είναι ο εξής. Εικονική πραγματικότητα είναι η χρήση της υπολογιστικής μοντελοποίησης και προσομοιώσεων που δίνει τη δυνατότητα σε ένα άτομο να αλληλεπιδρά με ένα τρισδιάστατο εικονικό περιβάλλον. Ένα ψηφιακά δημιουργημένο περιβάλλον προσομοιώνει την πραγματικότητα μέσω διαδραστικών συσκευών που στέλνουν και λαμβάνουν πληροφορίες. Η αίσθηση του ατόμου πως βρίσκεται σε αυτό το δημιουργημένο περιβάλλον (telepresence)

επιτυγχάνεται με τη χρήση αισθητήρων κίνησης που καταγράφουν τις κινήσεις του και προσαρμόζουν σε πραγματικό χρόνο και με τρόπο αντίστοιχο την οπτική του.

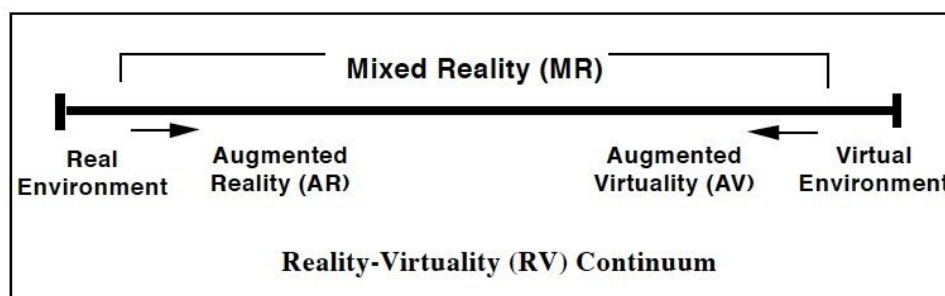
Η βάση της τεχνολογίας αυτής αναδύθηκε τη δεκαετία του 1960 με προσομοιωτές που χρησιμοποιούνταν στην στρατιωτική εκπαίδευση για το χειρισμό αεροπλάνων, οχημάτων, οπλισμού σε συνθήκες μάχης. Η εμπορική της χρήση ήρθε με τη δεκαετία του 1980 και πλέον χρησιμοποιείται σε παιχνίδια, σε εκθέσεις και σε διαστημικούς προσομοιωτές. Ακόμη, η εικονική πραγματικότητα χρησιμοποιείται σε διάφορους άλλους τομείς, όπως είναι η διασκέδαση, η τεχνολογία φαρμάκων, η βιοτεχνολογία, η μηχανική, η σχεδίαση και το μάρκετινγκ[3].

Τέλος, να σημειωθεί ότι η δημοτικότητα της έχει αυξηθεί τα τελευταία χρόνια για δύο βασικούς λόγους. Τη χρήση της στη βιομηχανία παιχνιδιών και στη μείωση του κόστους των συσκευών εικονικής πραγματικότητας. Έτσι, παρόλο που για ένα διάστημα είχε θεωρηθεί εμπορική αποτυχία, μετά από τεχνολογικές βελτιώσεις στις συσκευές, αλλά και γενικότερα το υλικό (processing hardware), κατάφερε να ξαναβγει στην επιφάνεια [4].

Το φάσμα του εικονικού συνεχούς κι άλλες κατηγοριοποιήσεις

Στην εικονική πραγματικότητα, το τρισδιάστατο περιβάλλον που έχει αναπτυχθεί βιώνεται από τους χρήστες με τη φυσική κίνηση σε αυτά, με τη γραπτή ή προφορική επικοινωνία σε κάποιες περιπτώσεις ακόμα και μέσω διαφόρων αισθήσεων, όπως είναι η αφή, η όσφρηση κλπ. Η βασικότερη ταξινόμηση των εφαρμογών της εικονικής πραγματικότητας γίνεται πάνω στο φάσμα του "εικονικού συνεχούς" (virtuality continuum), όπως αυτό ορίστηκε από τους Paul Milgram και Fumio Kishino [1]. Το εικονικό συνεχές εκτείνεται από την ίδια την πραγματικότητα και φτάνει μέχρι την πλήρως εικονική πραγματικότητα που δημιουργείται εξ ολοκλήρου από έναν υπολογιστή [5]. Μέσα σε αυτό βρίσκεται το υποσύνολο που ονομάζεται "μικτή πραγματικότητα" (mixed reality), στο οποίο συμπεριλαμβάνεται οτιδήποτε βρίσκεται ανάμεσα στην πραγματικότητα και την πλήρως τεχνητή εικονική πραγματικότητα. Ένα τέτοιο παράδειγμα είναι η "επαυξημένη πραγματικότητα" (augmented reality)[6]. Ως τέτοια ορίζεται η τεχνική να ενσωματώνονται κάποια στοιχεία του εικονικού κόσμου στον πραγματικό, ώστε να δίνονται περισσότερες πληροφορίες. Υπάρχουν όμως και διαφορές των εφαρμογών εικονικής πραγματικότητας και σε άλλα επίπεδα :

- Εφαρμογές που έχουν πρόσβαση σε ένα μεμονωμένο λογισμικό προβολής που είναι εγκατεστημένο σε κάποιον υπολογιστή ή εφαρμογές που μέσω διαδικτυακής πρόσβασης παρέχουν σε clients το λογισμικό προβολής από έναν απομακρυσμένο server.
- Εφαρμογές που παρέχουν μέσω προσομοίωσης μια αναπαράσταση του πραγματικού κόσμου ή εφαρμογές που απεικονίζουν έναν φανταστικό κόσμο.
- Εφαρμογές που αφορούν έναν μεμονωμένο χρήστη ή εφαρμογές που προορίζονται για χρήση από πολλούς χρήστες ταυτόχρονα [7].



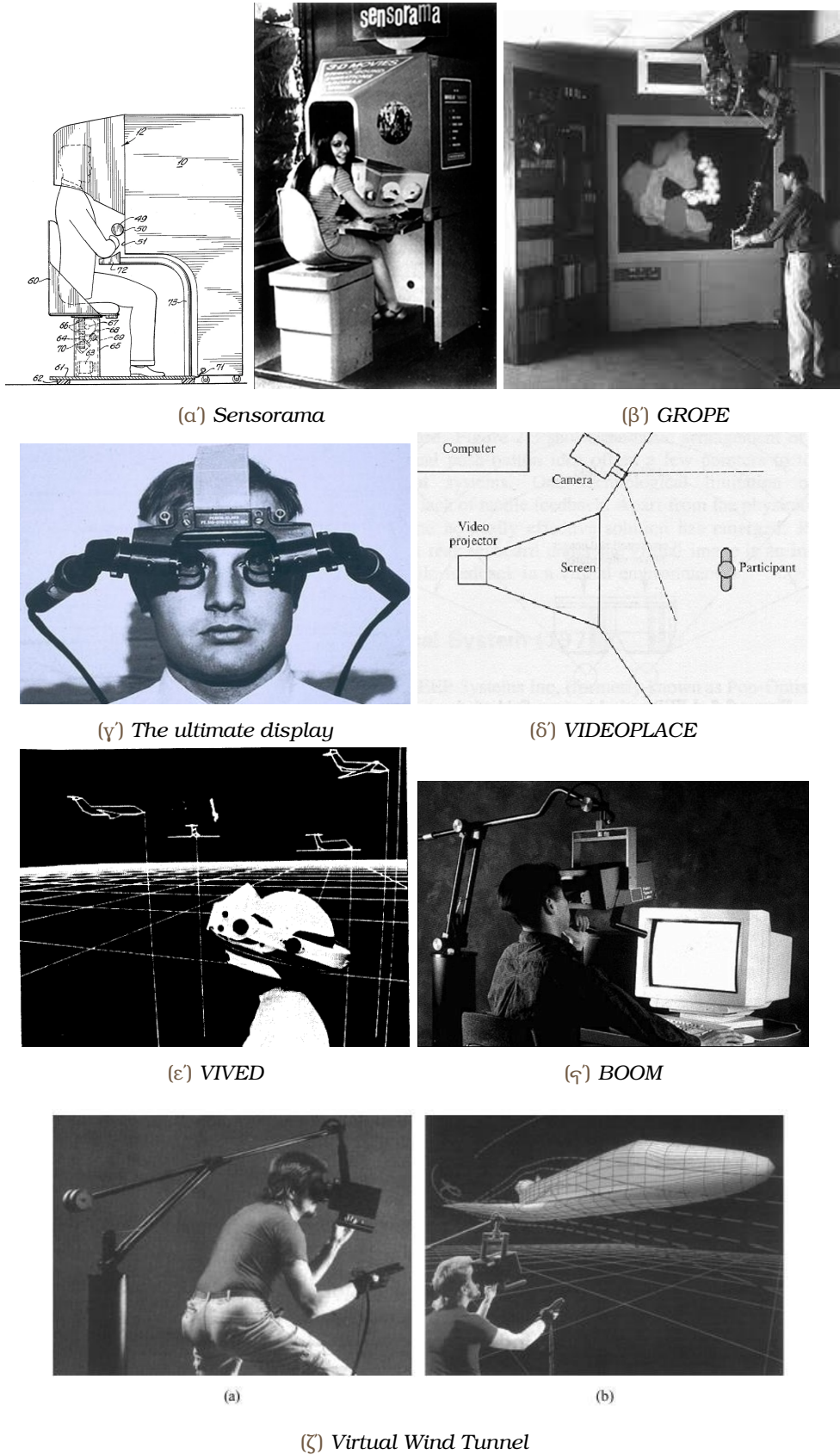
Σχήμα 2.1: Το φάσμα του εικονικού συνεχούς [1]

Κάποιες χαρακτηριστικές προσπάθειες για δημιουργία περιβαλλόντων εικονικής πραγματικότητας

- **Sensorama** - Το 1960-1962 ο Morton Heilig δημιούργησε έναν πολυαισθητηριακό προσομοιωτή. Μια προηχογραφημένη ταινία με χρώμα και στερεοφωνικό ήχο, ενισχύθηκε από διαισθητικό ήχο, μυρωδιά, άνεμο και εμπειρίες δόνησης. Αυτή ήταν η πρώτη προσέγγιση για τη δημιουργία ενός συστήματος εικονικής πραγματικότητας και είχε όλα τα χαρακτηριστικά ενός τέτοιου περιβάλλοντος, αλλά δεν ήταν διαδραστικό [2.2α'](#).
- **The Ultimate Display** - Το 1965 ο Ivan Sutherland πρότεινε την απόλυτη λύση της εικονικής πραγματικότητας: μια ιδέα κατασκευής τεχνητού κόσμου που περιλάμβανε διαδραστικά γραφικά, ανατροφοδότηση δύναμης, ήχο, όσφρηση και γεύση [2.2γ'](#).
- **The Sword of Damocles** - Το πρώτο σύστημα εικονικής πραγματικότητας που υλοποιήθηκε σε υλικό και όχι σε λογισμικό. Ο Ivan Sutherland κατασκεύασε μια συσκευή που θεωρείται η πρώτη Head Mounted Display (HMD) με κατάλληλη παρακολούθηση των κινήσεων του κεφαλιού. Υποστήριζε μια στερεοσκοπική προβολή που ενημερωνόταν ανάλογα με τη θέση και τον προσανατολισμό του κεφαλιού του χρήστη.
- **GROPE** - Το πρώτο prototype ενός συστήματος "force-feedback" λειτουργίας που υλοποιήθηκε στο Πανεπιστήμιο της Βόρειας Καρολίνας (YN) το 1971. Με τον παραπάνω όρο εννοείται η προσομοίωση φυσικών κανόνων, όπως είναι ο νόμος της βαρύτητας, σε περιβάλλοντα εικονικής πραγματικότητας. Έτσι, δίνεται η δυνατότητα στον χρήστη να αλληλεπιδρά άμεσα με αντικείμενα του εικονικού χώρου [2.2δ'](#).
- **VIDEOPLACE** - Τεχνολογία εικονικής πραγματικότητας που δημιουργήθηκε το 1975 από τον Myron Krueger - "ένα εννοιολογικό περιβάλλον, χωρίς ύπαρξη". Σε αυτό το σύστημα οι σιλουέτες των χρηστών που συλλαμβάνονταν από τις κάμερες προβάλλονταν σε μια μεγάλη οθόνη. Οι συμμετέχοντες ήταν σε θέση να αλληλεπιδρούν ο ένας με τον άλλον χάρη στις τεχνικές επεξεργασίας εικόνας που προσδιόριζαν τις θέσεις τους στο 2D χώρο της οθόνης [2.2δ'](#).
- **VCASS** - Ο Thomas Furness στο Armstrong Medical Research της Πολεμικής Αεροπορίας των ΗΠΑ Laboratories ανέπτυξε το 1982 τον Visually Coupled Airborne Systems Simulator - ένα προηγμένο προσομοιωτή πτήσης. Ο πιλότος του μαχητικού

αεροσκάφους φορούσε ένα HMD που ενίσχυε το θέα από το παράθυρο με τα γραφικά που περιγράφουν πληροφορίες στόχευσης ή βέλτιστης διαδρομής πτήσης.

- **VIVED** - Πρόκειται για ένα στερεοσκοπικό μονόχρωμο HMD που κατασκευάστηκε στο ερευνητικό κέντρο NASA Ames το 1984 με ήδη υπάρχουσες τεχνολογίες [2.2ε'](#).
- **VPL** - Η εταιρεία VPL κατασκεύασε το δημοφιλές DataGlove (1985) και το Eyephone HMD (1988) - τις πρώτες εμπορικά διαθέσιμες συσκευές VR.
- **BOOM** - Βγήκε στην αγορά το 1989 από την Fake Space Labs. Το BOOM είναι ένα μικρό κουτί που περιέχει δύο οθόνες CRT οι οποίες μπορούν να προβληθούν μέσω των οπών των ματιών. Ο χρήστης μπορούσε να πιάσει το κουτί, να το κρατήσει από τα μάτια και να κινηθεί μέσα στον εικονικό κόσμο, καθώς ο μηχανικός βραχίονας κατέγραφε τη θέση και τον προσανατολισμό του κουτιού [2.2ς'](#).
- **UNC Walkthrough project** - Στο δεύτερο μισό της δεκαετίας του 1980 στο Πανεπιστήμιο της Βόρειας Καρολίνας αναπτύχθηκε μια εφαρμογή αρχιτεκτονικής περιήγησης. Αρκετές συσκευές VR κατασκευάστηκαν για να βελτιωθεί η ποιότητα αυτού του συστήματος, όπως: HMDs, οπτικοί ανιχνευτές και η μηχανή γραφικών Pixel-Plane.
- **Virtual Wind Tunnel** - Αναπτύχθηκε στις αρχές της δεκαετίας του 1990 στην εφαρμογή του NASA Ames που επέτρεπε την παρατήρηση και τη διερεύνηση πεδίων ροής με τη βοήθεια των BOOM και DataGlove [2.2ζ'](#).
- **CAVE** - Παρουσιάστηκε το 1992 και ήταν ένα σύστημα εικονικής πραγματικότητας και επιστημονικής απεικόνισης. Αντί να χρησιμοποιεί HMD προβάλλει στερεοσκοπικές εικόνες στους τοίχους του δωματίου (ο χρήστης πρέπει να φοράει γυαλιά με LCD κλείστρα). Η προσέγγιση αυτή εξασφάλιζε ανώτερη ποιότητα και ανάλυση των προβλλόμενων εικόνων και ευρύτερο οπτικό πεδίο σε σύγκριση με τα συστήματα που βασίζονταν σε HMD.
- **Augmented Reality (AR)** - Μια τεχνολογία που "παρουσιάζει έναν εικονικό κόσμο που εμπλουτίζει, αντί να αντικαθιστά τον πραγματικό κόσμο". Αυτό επιτυγχάνεται με τη βοήθεια διαφανών HMD που τοποθετούν εικονικά τρισδιάστατα αντικείμενα πάνω σε πραγματικά. Η τεχνολογία αυτή χρησιμοποιήθηκε στο παρελθόν για τον εμπλουτισμό της όρασης του πιλότου μαχητικού αεροσκάφους με πρόσθετες πληροφορίες πτήσης (VCASS). Χάρη στις μεγάλες δυνατότητές της -την ενίσχυση της ανθρώπινης όρασης- η επαυξημένη πραγματικότητα έχει βρεθεί στο επίκεντρο πολλών ερευνητικών έργων από τις αρχές της δεκαετίας του 1990 [8].



Σχήμα 2.2: Ιστορικές φωτογραφίες εφαρμογών εικονικής πραγματικότητας

2.2 Εικονική πραγματικότητα και εκμάθηση γλωσσών

Εκπαίδευση μια ενεργητική γνωστική διαδικασία

Η εκπαίδευση ορίζεται γενικά ως η υιοθέτηση μιας μόνιμης αλλαγής στη συμπεριφορά ενός ανθρώπου, μέσω εμπειριών που έχει ο ίδιος[9]. Η μάθηση είναι μια γνωστική διαδικασία που απαιτεί την ενεργή συμμετοχή του ατόμου. Κάποιος άλλος θα έλεγε πως η μάθηση είναι μια διαδικασία κατά την οποία το άτομο διαμορφώνει τον εαυτό του. Από αυτές τις απόψεις για την εκπαίδευση και τη διδασκαλία, φαίνεται ότι οι αλλαγές στις οποίες το άτομο έχει ενεργό ρόλο μέσω των εκπαιδευτικών διαδικασιών, έχουν να κάνουν με τη συμπεριφοράς του, τις γνώσεις του και γενικότερα στο πως αντιλαμβάνεται τον κόσμο. Η τεχνολογική πρόοδος σε διάφορους τομείς μπορεί να προσφέρει εργαλεία, που βασισμένα στους παραπάνω ορισμούς, θα συμβάλλουν στην εκπαιδευτική διαδικασία και θα βοηθήσουν τα άτομα να μαθαίνουν πιο αποτελεσματικά. Αδιαμφισβήτητα, ένα από αυτά τα εργαλεία φαίνεται ότι είναι η εικονική πραγματικότητα [10].

Στην πράξη έχουν κατατεθεί πολλές θετικές οπτικές γύρω από εφαρμογές της εικονικής πραγματικότητας στην εκπαίδευση. Μάλιστα, το πλήθος αυτών αυξάνεται ακόμη περισσότερο αν η χρήση της εικονικής πραγματικότητας στην εκπαίδευση αναγνωριστεί ως φυσική εξέλιξη της υποβοηθούμενης από τον υπολογιστή εκπαίδευσης (computer-aided education). Ένα βασικό στοιχείο προς υπεράσπιση αυτής της άποψης είναι η δυνατότητα που δίνουν οι εφαρμογές αυτές στα άτομα να αναγνωρίζουν τους εαυτούς τους ως μέρος του περιβάλλοντος που παρατηρούν αυξάνοντας με αυτόν τον τρόπο το ενδιαφέρον τους για το σχετικό μάθημα. Αυτή η αίσθηση εξερεύνησης και εμπλοκής του ατόμου στον εικονικό κόσμο το κάνει και πιο συγκεντρωμένο σε αυτό, ενώ ταυτόχρονα το απομονώνει και από τα ερεθίσματα του πραγματικού κόσμου. Ακόμη, η χρήση των τεχνολογιών εικονικής πραγματικότητας διευκολύνει την αναπαράσταση γεγονότων ή φαινομένων, τα οποία είναι από δύσκολο έως αδύνατο να παρατηρηθούν στα πλαίσια ενός μαθήματος. Ουσιαστικά δηλαδή δημιουργεί μια εικόνα του πραγματικού κόσμου που δεν μπορεί να βιωθεί με άλλον τρόπο, όπως είναι ένα αστρονομικό φαινόμενο ή η ανατομία ενός ανθρώπινου σώματος, ένα παρελθοντικό ιστορικό γεγονός κλπ. Επιπλέον, εκτός από το ότι οι τεχνολογίες αυτές μπορούν να αναπαραστήσουν καταστάσεις που δε θα μπορούσαν να βιωθούν αλλιώς, είναι σημαντικό το ότι μπορούν να δημιουργούν γύρω από αυτές κι ένα αίσθημα ασφάλειας που πολλές φορές είναι απαραίτητο για την εκπαιδευτική διαδικασία [11].

Εκπαιδευτικές μέθοδοι που υιοθετούνται

Στα πλαίσια μελετών γύρω από τις εφαρμογές αυτές, έχουν γίνει προσπάθειες να τυποποιηθούν και να κατηγοριοποιηθούν οι τρόποι με τους οποίους διαφορετικές δραστηριότητες στον εικονικό κόσμο επιδρούν διαφορετικά στα άτομα και κατ' επέκταση ευνοούν διαφορετικές εκπαιδευτικές διαδικασίες. Αυτή η κατηγοριοποίηση κρίνεται ως πολύ χρήσιμη, γιατί δίνει τη δυνατότητα στο σχεδιαστή να εξατομικεύει τις λειτουργίες και τα χαρακτηριστικά των εφαρμογών εικονικής πραγματικότητας, ανάλογα με τους στόχους που έχουν τεθεί. Αντίστοιχα, αυτή η ανάλυση είναι χρήσιμη και για τους εκπαιδευτικούς και τους μαθητές, μιας και θα μπορούν να διακρίνουν πιο εύκολα ποια εφαρμογή ή ποια στοιχεία της κάθε εφαρμογής

είναι χρησιμότερα για το σκοπό του μαθήματος. Ακολουθούν έξι τρόποι με τους οποίους οι εφαρμογές εικονικής πραγματικότητας χρησιμοποιούνται στην εκπαίδευση:

- **Exploring** - Ένας βασικός τρόπος που οι εφαρμογές αυτές πετυχαίνουν τον εκπαιδευτικό τους σκοπό είναι μέσω της εξερεύνησης των χώρων και των ιδιοτήτων του εικονικού κόσμου.
- **Collaborating** - Μια άλλη μέθοδος που χρησιμοποιείται είναι αυτή τη συνεργασίας. Σε αυτήν την περίπτωση, προκειμένου να επιτευχθούν οι στόχοι της εφαρμογής ο χρήστης θα πρέπει εργαστεί ομαδικά στον εικονικό κόσμο. Η συνεργασία του χρήστη μπορεί να γίνεται είτε με άλλους πραγματικούς χρήστες, όπως είναι ο ίδιος, είτε με φανταστικούς χαρακτήρες της εφαρμογής.
- **Being** - Η μέθοδος αυτή είναι ίδια με την exploring με τη διαφορά πως εδώ το άτομο συμπεριλαμβάνεται το ίδιο στο πεδίο εξερεύνησης. Δηλαδή δεν είναι απλά παρατηρητής του εικονικού κόσμου, αλλά συμμετέχει ενεργά σε αυτό με τη δική του ταυτότητα και χαρακτηριστικά (role-play). Μια τέτοια περίπτωση είναι όταν το άτομο έχει το δικό του ανθρώπινο avatar.
- **Building** - Αυτή η μέθοδος εφαρμόζεται όταν ο χρήστης καλείται να συμμετάσχει ενεργά στη διαμόρφωση του περιβάλλοντος χώρου γύρω του όπως για παράδειγμα όταν δημιουργεί αντικείμενα στον εικονικό κόσμο.
- **Championing** - Πρόκειται για την εμπλοκή του ατόμου σε δραστηριότητες του εικονικού κόσμου. Αυτό όμως που κάνει τη συγκεκριμένη περίπτωση ενδιαφέρουσα είναι το γεγονός ότι το περιεχόμενο και οι σκοποί των δραστηριοτήτων αυτών είναι με τέτοιο τρόπο διαμορφωμένοι, ώστε να παρακινούν τον χρήστη να τους ολοκληρώνει με ενδιαφέρον. Ουσιαστικά δηλαδή, να ανταποκρίνονται στα πραγματικά ενδιαφέροντα του ατόμου λες και έχει να κάνει με την πραγματική του ζωή.
- **Expressing** - Τέλος, αυτή η μέθοδος, όπως και η προηγούμενη, αποτελεί έναν τρόπο σύνδεσης των δραστηριοτήτων της εφαρμογής με την πραγματική ζωή του χρήστη. Αυτή, όμως, δεν έχει να κάνει απλά με το ενδιαφέρον του χρήστη αλλά με το ότι τα επιτεύγματα και οι επιδόσεις αυτού, αν και συμβαίνουν εξ ολοκλήρου στον εικονικό κόσμο, μπορούν να προβληθούν και στον πραγματικό. Αυτή η προβολή μπορεί να συμβεί με διάφορα μέσα, όπως είναι παρουσιάσεις, podcasts κ.ά. [12].

Το βήμα για την εκμάθηση γλωσσών

Όσο η τεχνολογία προοδεύει και λόγω της παγκοσμιοποίησης η πολυπολιτισμικότητα γίνεται αναπόσπαστο κομμάτι της ζωής και της εργασίας των ανθρώπων. Έτσι, η ανάγκη για γνώση μιας δεύτερης ξένης γλώσσας και η υιοθέτηση διαπροσωπικών σχέσεων με άτομα με διαφορετικές κουλτούρες γίνεται μεγαλύτερη. Ως αποτέλεσμα της ανάπτυξης των νέων τεχνολογιών, η εκπαίδευση στις ξένες γλώσσες σιγά σιγά ενσωματώνει νέα εργαλεία και γίνεται όλο και περισσότερο μια διαδικασία υποβοηθούμενη από υπολογιστές. Ειδικότερα στην περίπτωση της εικονικής πραγματικότητας τα νέα αυτά περιβάλλοντα είναι ένα πολύ

χρήσιμο εργαλείο και για την εκμάθηση μιας γλώσσας, αλλά και για την εμπύθιση σε νέες κουλτούρες. Επιπλέον, η πρόσφατη πρόοδος στα διαλογικά συστήματα για τα οποία θα μιλήσουμε στην επόμενη ενότητα, και που μιμούνται ανθρώπινες συζητήσεις φέρνουν αυτή την προοπτική ακόμα πιο κοντά. Μια ρεαλιστική αναπαράσταση διαλόγου σε ένα εικονικό περιβάλλον για παράδειγμα κάνει την εκμάθηση γλωσσών και την ανταλλαγή πολιτισμικών στοιχείων ευκολότερες [13].

Στην περίπτωση της εκμάθησης γλωσσών, το κύριο πλεονέκτημα είναι η εμπυθιστική φύση της εικονικής πραγματικότητας. Η εμπύθιση αυτή επιτυγχάνεται εισάγοντας τον χρήστη σε έναν εικονικό κόσμο όπου όλες οι συζητήσεις είναι στη γλώσσα που μαθαίνει. Αυτή η διαδικασία, δε δημιουργεί απλά ένα χώρο που ο χρήστης μπορεί να εξασκήσει τη γλώσσα, αλλά τον ενθαρρύνει να εμπλακεί περισσότερο με όλη του την προσοχή. Η Ellen Bialystok στη μελέτη της για το ζήτημα εκμάθησης ξένων γλωσσών επισημαίνει τέσσερις στρατηγικές που βοηθούν τη διαδικασία αυτή και που σχετίζονται ως εξής:

- τυπική έναντι της λειτουργικής εξάσκησης και
- παρακολούθηση έναντι της εξαγωγής συμπερασμάτων [14].

Η τυπική εξάσκηση στοχεύει στην ακριβή χρήση της γλώσσας, ενώ η λειτουργική αναδεικνύει τρόπους για την επίτευξη επικοινωνίας χρησιμοποιώντας την. Η παρακολούθηση περιλαμβάνει την εξέταση και αξιολόγηση της χρήσης της γλώσσας, κάτι που αποτελεί την παραδοσιακή στρατηγική, ενώ η μέθοδος εξαγωγής συμπερασμάτων ενδιαφέρεται περισσότερο για το αν ο χρήστης αντιλήφθηκε το νόημα μιας φράσης ακόμα, ειδικά όταν αυτό δεν εξηγείται ρητά. Πέρα όμως από τις εκπαιδευτικές στρατηγικές, μια γλώσσα δεν είναι κάτι το στατικό, αλλά διαμορφώνεται συνεχώς από την κοινωνία που τη χρησιμοποιεί. Συνεπώς, για κάποιον που μαθαίνει μια γλώσσα είναι σημαντικό να έρθει σε επαφή και με την κουλτούρα που τη συνοδεύει. Η κουλτούρα αυτή περιλαμβάνει ιδέες, συνήθειες, δεξιότητες, τέχνες, πιστεύω και αξίες. Άρα, με τη χρήση εφαρμογών εικονικής πραγματικότητας, μπορεί κανείς να εξασκήσει μια γλώσσα με λειτουργικό και συμπερασματικό τρόπο και ταυτόχρονα να εισαχθεί στην αντίστοιχη κουλτούρα [15].

Για να γίνουμε πιο συγκεκριμένοι, το βασικό πλεονέκτημα που αναγνωρίζεται στη χρήση της εικονικής πραγματικότητας ως εκπαιδευτικού εργαλείου είναι το εξής. Το ότι λύνει ένα από τα βασικά προβλήματα που πάντα υπήρχαν στην εκμάθηση μιας ξένης γλώσσας: την απουσία εμπύθισης-έκθεσης του μαθητή σε ένα πραγματικό περιβάλλον χρήσης της νέας γλώσσας. Έτσι, μέσα από ρεαλιστικά σενάρια χρήσης της υπό μελέτη γλώσσας ο χρήστης ετοιμάζεται ψυχολογικά κι ενθαρρύνεται να χρησιμοποιήσει δεξιότητες και γνώσεις που ήδη κατέχει, ώστε να βελτιωθεί στην κατανόηση και χρήση της γλώσσας και γενικότερα στην επικοινωνία. Αυτό, δηλαδή, που κάνει κάποιος όταν "αναγκάζεται" να χρησιμοποιήσει αποκλειστικά μια γλώσσα που δε γνωρίζει σε ένα ξένο γί' αυτόν μέρος [16].

Ένα ακόμη βασικό πλεονέκτημα των εφαρμογών αυτών αφορά την έννοια του gamification (παιχνιδοποίηση). Σύμφωνα με αυτήν την έννοια, όταν μια εφαρμογή καταφέρνει να πετύχει το λεγόμενο gamification της μαθησιακής διαδικασίας, τότε εντοπίζονται τα εξής θετικά αποτελέσματα:

- Συνδυάζει την εξάσκηση πάνω σε δεδομένα και πληροφορίες που έχουν σχεδιαστεί εκ των προτέρων και με σκοπό την παρακίνηση του χρήστη προς τον αυτοσχεδιασμό.
- Αποτελεί μια ενεργή μορφή εξάσκησης, κατά την οποία ο χρήστης εμπλέκεται στη διαδικασία καθόλη τη διάρκεια, κρατώντας το ενδιαφέρον του αμείωτο.
- Έχει αποδειχθεί πως παρακινεί τους χρήστες να εξελίσσουν ένα δικό τους ανεξάρτητο και δημιουργικό τρόπο για την επίλυση προβλημάτων ή εν γένει για την αντιμετώπιση καταστάσεων.
- Αυξάνει την κινητοποίηση και όρεξη του χρήστη, ώστε να εμπλακεί στη μαθησιακή διαδικασία. Εκτός, δηλαδή, από το ότι ο χρήστης συντηρεί το ενδιαφέρον του για τη χρήση της εφαρμογής, βρίσκει ακόμη ένα κίνητρο να πετύχει το σκοπό χρήσης της που είναι η εκπαίδευση [17].

Ακολουθούν τρεις κατηγορίες μεθόδων εκμάθησης γλωσσών σε περιβάλλοντα εικονικής πραγματικότητας. Η διάκριση αυτών των κατηγοριών γίνεται με γνώμονα τη βασική σχέση που έχει ο χρήστης με τον εικονικό κόσμο της εφαρμογής:

1. **immersive (εμβυθιστική μέθοδος)**: όταν ο χρήστης περιβάλλεται από ένα εικονικό περιβάλλον, είτε αυτό είναι αληθινό, είτε φανταστικό και αλληλεπιδρώντας με αυτό πραγματοποιείται η μάθηση.
2. **social (κοινωνική μέθοδος)**: όταν οι τρισδιάστατοι χώροι αποτελούν περιβάλλοντα όπου οι χρήστες αλληλεπιδρούν με άλλους χρήστες και συμμετέχουν σε δραστηριότητες εκμάθησης γλωσσών, είτε αυτές έχουν ξεκάθαρο εκπαιδευτικό χαρακτήρα, είτε απλά εξασκούν τη γλώσσα επικοινωνώντας.
3. **creative (δημιουργική μέθοδος)**: όταν η εκμάθηση γλωσσών στο εικονικό περιβάλλον περιλαμβάνει και τη δημιουργία αντικειμένων/οντοτήτων ως δραστηριότητες. Αυτή η κατηγορία είναι η λιγότερο αναπτυγμένη από τις τρεις και δεν υπάρχει αρκετή σχετική έρευνα [12].

Περιορισμοί

Τι περιορισμοί και τι αδυναμίες προκύπτουν από τη χρήση της εικονικής πραγματικότητας σε μια εκπαιδευτική διαδικασία, αλλά και συγκεκριμένα στην εκμάθηση ξένων γλωσσών; Ίσως το βασικότερο μειονέκτημα χρήσης της εικονικής πραγματικότητας γενικά έχει να κάνει με το κόστος του εξοπλισμού που είναι απαραίτητος για να τρέξουν οι εφαρμογές. Ακόμη, σημαντικός παράγοντας είναι ο χρόνος ο οποίος απαιτείται για την εξάσκηση χρήσης του εξοπλισμού και στο επίπεδο του υλικού, αλλά και σε αυτό του λογισμικού, μιας και πρόκειται για εργαλεία με τα οποία ο μέσος χρήστης δεν έχει προηγούμενη εμπειρία. Επιπλέον, περιοριστικό ρόλο στη χρήση εφαρμογών εικονικής πραγματικότητας παίζουν διάφορα ζητήματα ασφαλείας ή κίνδυνου στην υγεία που μπορεί να προκύψουν. Για παράδειγμα, κατά τη διάρκεια χρήσης μιας εφαρμογής όσο μεγαλύτερη είναι η εμβύθιση του χρήστη στον εικονικό κόσμο, τόσο μικρότερη ή πιο αδύναμη είναι η αντιληπτική του ικανότητα στον πραγματικό κόσμο κι αν αυτό συνδυαστεί με έλλειψη των απαραίτητων μέτρων ασφαλείας, αυτό μπορεί

να οδηγήσει σε ατυχήματα. Ατυχήματα μπορούν βέβαια να προκύψουν και λόγω της ναυτίας (motion sickness), η οποία παρατηρείται συχνά ειδικά σε αρχάριους χρήστες. Τέλος, δεν μπορεί να αγνοηθεί η απροθυμία που εμφανίζεται σε πολλούς ανθρώπους που καλούνται να υιοθετήσουν και να χρησιμοποιήσουν νέες τεχνολογίες, ειδικά όταν η χρήση τους αφορά διαδικασίες που δεν είναι καινούργιες, όπως αυτή της εκπαίδευσης. Όλα αυτά τα ζητήματα αποτελούν πραγματικούς λόγους που μπορεί να αποθαρρύνουν κόσμο από το να χρησιμοποιήσει εφαρμογές εικονικής πραγματικότητας για εκπαιδευτικούς σκοπούς. Υπάρχει, όμως, η εκτίμηση ότι με τη συνεχή ανάπτυξη και βελτίωση των τεχνολογιών αυτών, τη μείωση του κόστους του εξοπλισμού, την επακόλουθη αύξηση της χρήσης τους σε τομείς πέρα από την εκπαίδευση και την εξοικείωση που ο κόσμος θα αποκτήσει με αυτές, κάθε ένα από αυτά τα ζητήματα θα ξεπεραστούν [18].

2.3 Διαλογικά Συστήματα

Ορισμοί

Έχουν χρησιμοποιηθεί διάφοροι όροι αντί του chatbot, τον οποίο επινόησε ο Michael L. Mauldin το 1994 [19]. Κάποιοι από αυτούς είναι σύστημα μηχανικής συνομιλίας (machine conversation system), εικονικός πράκτορας (virtual agent), διαλογικό σύστημα (dialogue system), chatterbot κ.ά.. Σύμφωνα με τον απλουστευμένο ορισμό των Shawar και Atwell τα chatbots είναι εφαρμογές συνομιλίας που βασίζονται στην τεχνητή νοημοσύνη και των οποίων οι λειτουργίες κυμαίνονται από την απάντηση απλών ερωτήσεων μέχρι τη συμμετοχή σε πολύπλοκες συνομιλίες [20]. Ανάλογα με την περίπτωση, τα chatbots μπορούν να συμμετέχουν είτε σε γραπτές, είτε σε προφορικές συνομιλίες. Στην αρχή τα chatbots δεν ήταν και τόσο έξυπνα από την άποψη ότι μπορούσαν απλά να ανταποκριθούν σε προκαθορισμένες ερωτήσεις, στις οποίες μάλιστα έδιναν και συγκεκριμένες απαντήσεις. Στις περισσότερες των περιπτώσεων, δεν είχαν καμιά χρησιμότητα πέρα από την ακαδημαϊκή έρευνα σε αυτά ή τη διασκέδαση και συχνά αποθάρρυναν τους χρήστες από το να αλληλεπιδράσουν ξανά μαζί τους. Με την πρόοδο της τεχνητής νοημοσύνης, ωστόσο, άρχισαν να γίνονται καλύτερα σε βαθμό που μπορούσαν αξιοπρεπώς να αντικαταστήσουν τον άνθρωπο σε κάποιες θέσεις που αφορούσαν βασικές επικοινωνιακές λειτουργίες, όπως σε κάποιους εκπαιδευτικούς τομείς [21]. Σύμφωνα με τον Britz, τα chatbots πλέον είναι προγράμματα που είναι ικανά να διεξάγουν συζητήσεις παρόμοιες με αυτές που αναπτύσσονται ανάμεσα σε ανθρώπους [22]. Συνήθως, για παράδειγμα χρησιμοποιούνται για να αυτοματοποιήσουν ή να βελτιώσουν μια εμπορική διαδικασία. Οι τύποι των chatbots διακρίνονται στους απλούς και στους πιο πολύπλοκους. Στην πρώτη περίπτωση, τα απλά bots χειρίζονται βασικά μηνύματα και αιτήματα χρηστών, με τις απαντήσεις σου να είναι προκαθορισμένες και το στυλ τους στη χρήση της γλώσσας να μην είναι ιδιαίτερα εξεζητημένο. Για παράδειγμα, στην ερώτηση ενός πελάτη για κάποιο προϊόν, η απάντηση του bot θα είναι ένα link με σχετικές πληροφορίες. Στη δεύτερη περίπτωση, τα πιο πολύπλοκα chatbots είναι αποτελεσματικότερα και μπορούν να εμπλακούν σε πιο σύνθετες συζητήσεις. Μέσω πρακτικών μηχανικής μάθησης, προγραμματίζονται ώστε να μαθαίνουν από παραδείγματα συζητήσεων που έχουν ήδη γίνει. Μάλιστα, αυτά τα συστήματα πλέον υποστηρίζονται και από μηχανισμούς αναγνώρισης και επεξεργασίας φυσικής γλώσσας (NLP - Natural Language Processing or Understanding).

Μια ιστορική επισκόπηση των chatbots

Το Turing Test, που ουσιαστικά πραγματεύεται την ερώτηση "Μπορούν οι μηχανές να σκεφτούν;", συστάθηκε από τον Alan Turing τη δεκαετία του 1950 σε μια εποχή που ο όρος του chatbot είχε αρχίσει να αποκτά δημοσιότητα [23]. Το αναγνωρισμένο ως πρώτο chatbot στην ιστορία είναι το ELIZA δημιουργήθηκε το 1966 και ο σκοπός του ήταν να συμπεριφέρεται ως ψυχοθεραπευτής, επιστρέφοντας λέξεις ή φράσεις του χρήστη ως ερωτήσεις. Ουσιαστικά, χρησιμοποιούσε ένα σύστημα απαντήσεων βασισμένο σε τεχνικές αναγνώρισης προτύπων για να απαντά στα ερωτήματα του χρήστη [24]. Μετά το ELIZA, το 1972, δημιουργήθηκε ένα ακόμη chatbot με την ονομασία PARRY, ένα πρόγραμμα προσομοίωσης ασθενή με συμπτώματα παράνοιας. Το μοτίβο που χρησιμοποιούσε για να εξελίσσεται η συζήτηση

ήταν έκφραση κυρίως φόβων και αγωνιών που έχει. Σκοπός του ήταν να δει αν άλλοι ψυχίατροι μπορούσαν να ξεχωρίσουν αν πρόκειται για πραγματικούς παρανοϊκούς ασθενείς ή για ένα πρόγραμμα [25]. Ο συνδυασμός αυτών των δύο αποτέλεσε και την πρώτη συνομιλία ανάμεσα σε δύο διαφορετικά chatbots που είχε πολύ ενδιαφέρον [23]. Το 1995 δημιουργήθηκε το chatbot ALICE, το οποίο μάλιστα βραβεύτηκε με το βραβείο Loebner τρεις χρονιές, το 2000, το 2001 και το 2004. Το ALICE έγινε το πρώτο chatbot που χαρακτηρίστηκε ως “ανθρώπινος υπολογιστής”, γράφτηκε στην AIML (Artificial Intelligence Markup Language) και χρησιμοποίησε αναγνώριση προτύπων για να ορίσει τις βασικές του λειτουργίες [26].

Ένα πολύ σημαντικό βήμα σε αυτές τις τεχνολογίες έγινε το 2003, όταν οι χρήστες του AIM, μιας καινοτόμας για την εποχή διαδικτυακής εφαρμογής συνομιλίας σε πραγματικό χρόνο, μπορούσαν να αλληλεπιδράσουν με την ZoeOnAOL, ένα chatbot που “ζούσε” στην πλατφόρμα AIM. Αφού η Zoe προστέθηκε στη λίστα των δυνατών επαφών, οι χρήστες της εφαρμογής μπορούσαν να συνομιλήσουν μαζί της. Οι περισσότερες ερωτήσεις αφορούσαν πληροφορίες που μπορούσαν να αποκτηθούν και από διάφορες μηχανές αναζήτησης. Αναγνώριζε λέξεις-κλειδιά και έκανε αναζητήσεις όπως “Τι καιρό έχει σήμερα στο Ντάλας” και έδινε άμεσα τη σωστή απάντηση. Με δεδομένο ότι η Zoe μπορούσε να συμμετέχει μόνο σε πολύ βασικές συζητήσεις, οι χρήστες αναγνώριζαν ότι επικοινωνούσαν με ένα πρόγραμμα. Στη συνέχεια, αντικαταστάθηκε από το SmarterChild, το οποίο ήταν διαθέσιμο στις υπηρεσίες AOL, Yahoo και MSN. Από εκείνο το σημείο και ύστερα, τα chatbots αναγνωρίστηκαν ως ένα χρήσιμο εργαλείο για εμπορικούς και διαφημιστικούς σκοπούς [27]. Έτσι, με την εξέλιξη της τεχνολογίας ακολουθεί η δημιουργία πιο σύγχρονων chatbots, όπως για παράδειγμα η Siri της Apple, η Alexa της Amazon, ο Watson της IBM η Cortana της Microsoft και ο βοηθός της Google.

Τέλος, μια πτυχή των chatbots που έχει προσεγγίσει τελευταία αρκετό ερευνητικό ενδιαφέρον είναι η ενσυναίσθηση που μπορούν να αναπτύξουν. Όπως είπαμε και προηγουμένως, το PARRY ήταν το πρώτο chatbot που έδειχνε συναισθήματα στον συνομιλητή του. Αντίστοιχα, τώρα ένας από τους στόχους της τεχνητής νοημοσύνης είναι ο σχεδιασμός chatbots με την ικανότητα να ανιχνεύουν τα συναισθήματα του ανθρώπινου συνομιλητή και να επιδεικνύουν ή να προσομοιώνουν συναισθηματική συμπεριφορά. Αυτή η ερευνητική τάση συνδυάζεται και με την παρατήρηση ότι, αν και τα περισσότερα chatbots σχεδιάζονται προκειμένου να διαχειρίζονται αλληλεπιδράσεις εμπορικής φύσης, για παράδειγμα μέσω ενός ηλεκτρονικού καταστήματος, η χρήση chatbots στην ιατρική έχει γνωρίσει μεγάλη έκρηξη. Πράγματι, όπως έχει παρατηρηθεί ήδη από τη δεκαετία του 1970 η χρήση διαλογικών βοηθών και εν γένει υπολογιστών σε διάφορα στάδια επικοινωνίας με τους ασθενείς φαίνεται να έχει θετικά αποτελέσματα. Ένας από τους λόγους που συμβαίνει αυτό είναι η άνεση που μπορεί να νιώθει ο ασθενής με τον υπολογιστή που χρησιμοποιεί, μιας και το τερματικό του υπολογιστή είναι ένα απρόσωπο στοιχείο χωρίς κοινωνικά χαρακτηριστικά τύπου τάξης, φύλο, φυλή και κατ’ επέκταση η διαδικασία μπορεί και ξεπερνάει εμπόδια που οφείλονται σε αυτούς τους διαχωρισμούς [28]. Ακόμη, μια συνομιλία μεταξύ ενός chatbot και ενός ασθενούς επιτρέπει, σε ορισμένες κλινικές καταστάσεις, την καλύτερη διάγνωση και θεραπεία, τη βελτίωση της παρακολούθησης του ασθενούς και τη βελτίωση της συλλογής δεδομένων από αυτόν, προσφέροντας καλύτερη και πιο εξατομικευμένη φροντίδα. Για παράδειγμα, στην ψυχολογία και την ψυχιατρική, υπάρχουν μελέτες που προτείνουν τη χρήση chatbots

κατά τη διενέργεια ψυχολογικών αξιολογήσεων και για τη θεραπεία της ψυχικής υγείας. Ένα παράδειγμα τέτοιου chatbot είναι το LENNA. Η ψυχική κατάσταση ενός υποκειμένου, δηλαδή η διάθεση, μπορεί να επηρεαστεί από το περιβάλλον ή από την αλληλεπίδρασή του με άλλα άτομα, για παράδειγμα, μέσω μιας συνομιλίας μεταξύ του υποκειμένου και άλλων ατόμων [29]. Αυτό το ανοιχτό πεδίο έρευνας στα chatbots και οι ενδεχόμενες εξελίξεις που θα προκύψουν θα μπορούσαν να εφαρμοστούν και σε άλλους τομείς, όπως είναι η εκπαίδευση.

Σύγχρονα διαλογικά συστήματα

Μετά τα διαλογικά συστήματα που βασίζονταν αμιγώς στην επεξεργασία γραπτού λόγου, το πρώτο διαλογικό σύστημα που μιλούσε δημιουργήθηκε στα πλαίσια του προγράμματος DARPA το 1977 [30]. Μετά από αυτό το σύστημα, δημιουργήθηκε και το πρώτο διαλογικό σύστημα που μπορούσε να μιλήσει περισσότερες από μία γλώσσες, αφού εκτός των αγγλικών μιλούσε γαλλικά, γερμανικά και ιταλικά [31]. Αυτά τα πρώτα φωνητικά διαλογικά συστήματα χρησιμοποιήθηκαν στη βιομηχανία των τηλεπικοινωνιών για την παροχή τηλεφωνικής εξυπηρέτησης.

Η αρχιτεκτονική 2.3, τα στοιχεία που αποτελούν το διαλογικό σύστημα και οι αρμοδιότητες αυτών διαφέρουν από περίπτωση σε περίπτωση. Βασικό στοιχείο κάθε διαλογικού συστήματος, όμως, είναι το μέρος αυτό που διαχειρίζεται την κατάσταση, τη ροή και τη στρατηγική του διαλόγου. Το μέρος αυτό μπορεί να θεωρηθεί ως μια εκδοχή των παραδοσιακών chatbots, όπως αυτά περιγράφηκαν από πάνω. Ένας τυπικός κύκλος λειτουργίας σε ένα διαλογικό σύστημα περιλαμβάνει τις εξής φάσεις: [32]

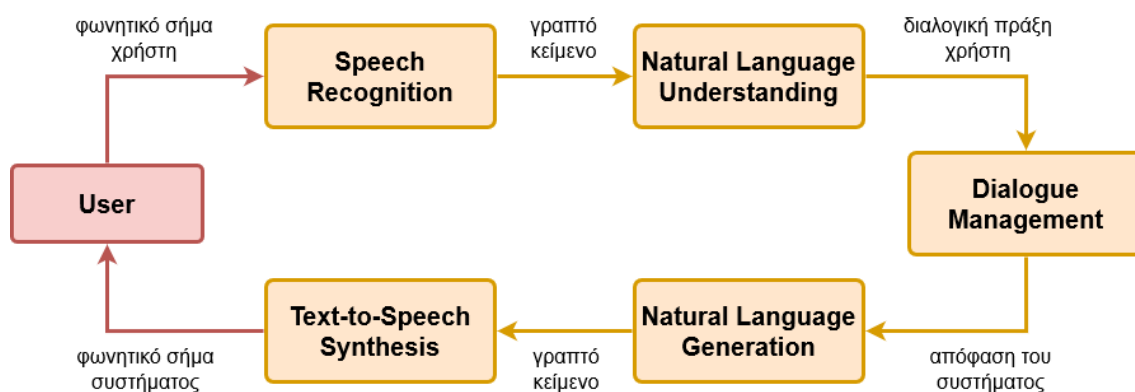
1. Ο χρήστης μιλάει και η φωνητική είσοδος μετατρέπεται σε απλό κείμενο από το αντίστοιχο μέρος του συστήματος, το οποίο μπορεί να περιλαμβάνει:
 - Αυτόματη αναγνώριση ομιλίας (ASR).
 - Αναγνώριση χειρονομιών (gestures).
 - Αναγνώριση γραφής (handwriting).
2. Το κείμενο αναλύεται από ένα σύστημα κατανόησης της φυσικής γλώσσας (NLU), η οποία μπορεί να περιλαμβάνει:
 - Ορθή αναγνώριση ονομάτων και οντοτήτων.
 - Αναγνώριση και χαρακτηρισμό των μερών του λόγου.
 - Συντακτικό / σημασιολογικό αναλυτή.
3. Οι σημασιολογικές πληροφορίες αναλύονται από το μέρος που διαχειρίζεται τον διάλογο, το οποίο διατηρεί το ιστορικό και την κατάσταση του διαλόγου και ελέγχει τη γενική ροή της συνομιλίας. Συνήθως, ο διαχειριστής αυτός επικοινωνεί με έναν ή περισσότερους άλλους διαχειριστές εργασιών, οι οποίοι έχουν γνώση των λειτουργιών που πρέπει να επιτελεστούν.
4. 5. Ο διαχειριστής διαλόγου παράγει την έξοδο χρησιμοποιώντας εργαλεία που μπορεί να περιλαμβάνουν:

- Παραγωγή φυσικής γλώσσας (NLG).
- Εξαγωγή χειρονομιών.
- Διαχείριση της μορφής και της διεπαφής με τον χρήστη.

5. Τέλος, δίνεται η απάντηση προς το χρήστη ως έξοδος, η οποία μπορεί να αποδοθεί μέσω:

- Σύνθεσης φωνής (μετατροπή κειμένου σε ομιλία - TTS).
- Αναπαράστασης με ρομπότ ή avatar.

Τα διαλογικά συστήματα που λειτουργούν αποκλειστικά με διεπαφή κειμένου (π.χ. γραπτή συνομιλία) περιέχουν μόνο τα στάδια 2-4.



Σχήμα 2.3: Η τυπική αρχιτεκτονική ενός διαλογικού συστήματος

2.4 Διαλογικά συστήματα και εκπαίδευση

Τι γίνεται με την είσοδο τους στην εκπαίδευση;

Η πρόοδος στο πεδίο της τεχνητής νοημοσύνης και ιδιαίτερα στα chatbots έχει οδηγήσει σε διάφορες νέες δυνατότητες σε πολλούς κλάδους. Μάλιστα στην εκπαίδευση, τα chatbots χρησιμοποιούνται ήδη για εξάσκηση των μαθητών στο επίπεδο της αλληλεπίδρασης, αλλά και για να βοηθήσουν το διδακτικό προσωπικό με την εισαγωγή της αυτοματοποίησης [33]. Η χρήση των chatbots στις εκπαιδευτικές διαδικασίες προβλέπεται ότι θα αυξήσει την συνδεσιμότητα, την αποτελεσματικότητα και θα μειώσει σε μεγάλο βαθμό την αβεβαιότητα των εκπαιδευτικών αλληλεπιδράσεων. Θα δημιουργηθεί, δηλαδή, ένα μαθησιακό περιβάλλον που θα είναι πιο προσωποποιημένο στον ίδιο το μαθητή και πιο στοχευμένο στον σκοπό της εκπαίδευσης [21]. Εν ολίγοις, τα chatbots υπόσχονται πως θα φέρουν επανάσταση στην εκπαίδευση αφενός εμπλέκοντας πιο ενεργά τους μαθητές, προσωποποιώντας τις εκπαιδευτικές δραστηριότητες, υποστηρίζοντας το έργο των εκπαιδευτών και εξελίσσοντας μια πιο βαθιά ματιά στις δυνατότητες και τις ανάγκες του κάθε μαθητή. Επιπλέον, μπορούν να εξυπηρετούν πολλούς μαθητές την ίδια στιγμή, δίνοντάς τους εκπαιδευτικό περιεχόμενο, ασκήσεις, ερωτήσεις κατανόησης και πηγές μελέτης. Ακόμη, σε συνδυασμό με άλλες τεχνολογίες τα chatbots μπορούν υποστηρίξουν διαδραστικές δραστηριότητες που περιλαμβάνουν χρήση κειμένου, ομιλίας, γραφικών, της αίσθησης της αφής, εκφράσεων του προσώπου και άλλες μορφές επικοινωνίας. Πέρα, όμως, από τις παραδοσιακές μορφές εκπαίδευσης με τη δια ζώσης παρουσία σε τμήματα κ.λπ., η χρήση των chatbots κρίνεται ακόμα πιο χρήσιμη στα διαδικτυακά και εξ αποστάσεως μαθήματα. Μάλιστα, μετά την πρόσφατη εμπειρία της πανδημίας η υιοθέτηση τέτοιων καινοτόμων πρακτικών στην εκπαίδευση φαίνεται πιο ελκυστική. Μια υιοθέτηση που τα τελευταία χρόνια, με τη μείωση του κόστους σχετικού εξοπλισμού, γίνεται όλο και πιο βιώσιμη οικονομικά.

Όσον αφορά το είδος των εκπαιδευτικών chatbots σχετικές συγκεντρωτικές έρευνες για τους εκπαιδευτικούς κλάδους στους οποίους χρησιμοποιούνται τα chatbots δείχνουν τα εξής αποτελέσματα:

- το 33.33% των εκπαιδευτικών chatbots εξειδικεύονται στην επιστήμη των υπολογιστών,
- το 22.77% είναι γενικού σκοπού και χρησιμοποιούνται για διάφορα εκπαιδευτικά περιεχόμενα
- το 16,66% των chatbots χρησιμοποιούνται στην εκμάθηση γλωσσών. Μάλιστα σε αυτήν την περίπτωση οι δημιουργοί των chatbots ενσωματώνουν στο σύστημά τους κι άλλες λειτουργίες, όπως είναι η αναπαραγωγή ήχου και βίντεο και η αναγνώριση φωνής [34].

Ηθικοί προβληματισμοί

Παρά τις αισιόδοξες επιδράσεις που έχουν τα chatbots στον εκπαιδευτικό μετασχηματισμό, φαίνεται να υπάρχει μια διαμάχη όσον αφορά τα πλεονεκτήματα και συνεπακόλουθα οφέλη τους συγκριτικά με τις ήδη υπάρχουσες μεθόδους και τα εκπαιδευτικά εργαλεία που χρησιμοποιούνταν μέχρι στιγμής [21]. Ενώ, η χρήση των chatbots στην εκπαίδευση εν γένει

απολογίζεται ως μια θετική εξέλιξη, γίνεται αποδέκτης κάποιων ηθικών προβληματισμών [35]. Οι προβληματισμοί αυτοί πηγάζουν από το ενδεχόμενο κατάχρησης ή εξαπάτησης (ακόμη κι αν αυτές γίνονται μη ηθελημένα), γιατί οι χρήστες μπορεί να νομίζουν λανθασμένα πως συνομιλούν με κάποιον πραγματικό άνθρωπο, ενώ στην πραγματικότητα μιλούν με ένα chatbot. Ακόμη, τα chatbots είναι μέσα που δεν έχουν από μόνα τους κάποια ηθική και είναι μη ανεξάρτητα και παίρνουν μέρος σε μη πραγματικές συζητήσεις. Αυτός ο προβληματισμός δεν έρχεται από μία οπτική που είναι αρνητικά προκατειλημμένη με την ύπαρξη των chatbots a priori, αλλά αντίθετα προσπαθεί να αναδείξει κάποια κενά ή αδυναμίες που υπάρχουν σε αυτήν τη νέα τεχνολογία και πρέπει να προσεχθούν.

Τι κάνει τα chatbots ελκυστικά στην εκμάθηση γλωσσών;

Όσον αφορά το ζήτημα της γλώσσας στην σύγχρονη τεχνολογική εποχή υπάρχουν οι εξής τρεις τοποθετήσεις που είναι ενδιαφέρουσες. Η ανάγκη για διαλογικούς βοηθούς συνέπεσε με τη διαδεδομένη χρήση προσωπικών ηλεκτρονικών μέσων, των οποίων οι χρήστες έχουν την επιθυμία να χρησιμοποιούν στη φυσική και μητρική τους γλώσσα [36]. Ακόμη, όσο οι άνθρωποι χρησιμοποιούν κάποια γλώσσα για την διαπροσωπική τους επικοινωνία, τόσο θα θέλουν να χρησιμοποιούν την ίδια γλώσσα και για την επικοινωνία τους με τους υπολογιστές [37]. Τέλος, ο καλύτερος τρόπος να διευκολύνεις την αλληλεπίδραση ανθρώπου - μηχανής (HCI - Human Computer Interaction) είναι να επιτρέψεις στους χρήστες να εκφράσουν το ενδιαφέρον, τις επιθυμίες ή τα ερωτήματα τους με τρόπο άμεσο και φυσικό μέσω την ομιλίας, του πληκτρολογίου ή της χειρονομίας [38].

Η εκμάθηση μια γλώσσας δεν είναι εύκολη διαδικασία. Ακόμα και όταν αυτή γίνεται κάτω από τις καλύτερες συνθήκες, οι μαθητές αντιμετωπίζουν διάφορα εμπόδια όμως είναι οι πολιτισμικές διαφορές, προβλήματα στην προφορά τους, η έλλειψη κινήτρων, η απουσία επικοινωνιακής κριτικής ή επιβεβαίωσης, η ανάγκη να μάθεις ειδικό λεξιλόγιο και πολλά άλλα. Οι μαθητές που διδάσκονται μια νέα ξένη γλώσσα συνήθως αντιμετωπίζουν όλα αυτά τα προβλήματα, ενώ έχουν από λίγη έως και καθόλου τη δυνατότητα να χρησιμοποιήσουν τη γλώσσα που μαθαίνουν πέρα από τη διαδικασία εκπαίδευσης (π.χ. έξω από την τάξη). Ειδικά οι μαθητές μιας γλώσσας έχουν συνήθως πολύ λίγα μέσα για να εξασκήσουν τις νέες τους γνώσεις, όπως είναι οι δραστηριότητες listening, όπου συνήθως αφού ακούσουν κάποιες ηχογραφήσεις τις επαναλαμβάνουν ή τις καταγράφουν γραπτές. Στις περιπτώσεις που η διδασκαλία γίνεται σε τμήματα, οι μαθητές μπορούν στην καλύτερη περίπτωση να εξασκήσουν τη χρήση της γλώσσας με τους συμμαθητές τους, ενώ μόνο σε μικρά τμήματα έχουν την ευκαιρία να την εξασκήσουν με κάποιον έμπειρο χρήστη, όπως είναι ο καθηγητής. Ακόμα κι έτσι όμως, η εξάσκηση αυτή σπάνια είναι πολύ διαδραστική και συνήθως αντιμετωπίζει προβλήματα χάρη στην έλλειψη αυτοπεποίθησης, στο αίσθημα ντροπής ή ακόμα κι εξαιτίας λαθών στη γραμματική και την προφορά που δε διορθώνονται ποτέ [39].

Νέες δυνατότητες

Έξι τρόποι με τους οποίους τα chatbots αποδεικνύουν ότι η χρησιμότητά τους στην εκμάθηση γλωσσών υπερβαίνει όποια δυσκολία μπορεί να προκύψει είναι οι εξής:

1. Οι μαθητές νιώθουν πιο άνετα όταν συνομιλούν με έναν υπολογιστή από ότι με έναν άνθρωπο.
2. Τα chatbots είναι πρόθυμα να επαναλάβουν ξανά και ξανά το ίδιο πράγμα μέχρι ο μαθητής να νιώσει ότι το κατάλαβε. Αυτό προφανώς συμβαίνει, γιατί τα bots δε βαριούνται, δεν κουράζονται, αλλά ούτε και χάνουν την υπομονή τους.
3. Πολλά bots μπορούν να παράγουν και γραπτό και προφορικό λόγο. Με αυτόν τον τρόπο, επιτρέπουν στους μαθητές να εξασκούν και τις ικανότητες ανάγνωσης, αλλά και κατανόησης του προφορικού λόγου.
4. Τα chatbots είναι κάτι το καινούργιο και το ενδιαφέρον για τους μαθητές. Μάλιστα, στις περισσότερες των περιπτώσεων, το να συνομιλούν μαζί τους, τους φαίνεται ως κάτι το αστείο και το διασκεδαστικό. Αυτές οι θετικές εμπειρίες με τα chatbots θα μπορούσαν να δημιουργήσουν νέο ή να ανανεώσουν το ήδη υπάρχον ενδιαφέρον στην εκμάθηση γλωσσών και να αναζωπυρώσουν το κίνητρο των μαθητών. Ακόμη, όταν κάποιο chatbot παλιώσει ή οι μαθητές αποκτήσουν τόσο μεγάλη οικειότητα μαζί του που δεν τους προκαλεί το ενδιαφέρον, τότε θα μπορούσε απλά να αντικατασταθεί από ένα νέο με καινούργια χαρακτηριστικά και προσωπικότητα. Θα μπορούσε βέβαια πάντα να χρησιμοποιείται ένα chatbot το οποίο βελτιώνεται όσο χρησιμοποιείται και αξιοποιεί τα καινούργια δεδομένα από τις συνομιλίες που κάνει για να εξελίσσεται και να ανανεώνεται με αυτόν τον τρόπο.
5. Οι μαθητές έχουν τη δυνατότητα να χρησιμοποιούν τη γλώσσα με ποικιλία και στη δομή, αλλά και στο λεξιλόγιο με έναν τρόπο που υπό κανονικές συνθήκες δε θα μπορούσαν. Τέτοια παραδείγματα χρήσης της γλώσσας είναι διάφορες εκφράσεις της καθημιλουμένης (slang) ή λέξεις / φράσεις που θεωρούνται taboo. Μάλιστα, αυτού του είδους η γλώσσα αν και είναι σημαντική για τους μαθητές, ώστε να έρθουν σε επαφή και να την κατανοήσουν, σπάνια διδάσκεται και είναι ακόμα πιο δύσκολο να την εξασκήσουν σε συνθήκες καθημερινής χρήσης της γλώσσας.
6. Τέλος, τα chatbots μπορούν να δημιουργηθούν με τέτοιο τρόπο, ώστε όχι μόνο να συνομιλούν με τους χρήστες, αλλά να δίνουν και feedback στους μαθητές για ενδεχόμενα γραμματικά ή ορθογραφικά λάθη που κάνουν.

Τρόποι που διευκολύνουν τους εκπαιδευτικούς

Τα chatbots είναι εν δυνάμει και πολύ χρήσιμα στην εκπαιδευτική διαδικασία όσον αφορά τον τρόπο που διευκολύνουν τους δασκάλους. Παρόλο που δεν γνωρίζουμε κανένα που να έχει δημιουργηθεί με σκοπό να αντικαθιστά το ρόλο του δασκάλου στην τάξη ή να είναι ένα αυτοτελές εργαλείο εκμάθησης γλώσσας, τα chatbots έχουν ή μπορούν να έχουν ήδη πολλές χρήσεις. Τέσσερις από αυτές τις χρήσεις είναι οι εξής:

1. **Εξάσκηση της ελεύθερης ομιλίας.** Σε μια συνθήκη που σε μια τάξη κάθε θρανίο έχει πρόσβαση σε υπολογιστή οι μαθητές θα μπορούσαν να πειραματιστούν συνομιλώντας με το chatbot πάνω σε όποιο θέμα θέλουν αυτοί ή ο δάσκαλος τους.

2. **Επανάληψη.** Το πιο σύνηθες για τους μαθητές που έρχονται σε επαφή με νέο υλικό είναι να μην έχουν την ευκαιρία να το χρησιμοποιήσουν πραγματικά. Με τη χρήση των chatbots, μετά από κάθε μάθημα θα έχουν την ευκαιρία να συζητήσουν με αυτό και να δοκιμάσουν τις καινούργιες τους γνώσεις.
3. **Αυτο-έλεγχος.** Αυτό θα μπορούσε να είναι το ιδανικό μέσο για τους μαθητές να αξιολογήσουν τους εαυτούς τους, τους συμμαθητές τους ή ακόμη και το ίδιο το chatbot.
4. **Για το δάσκαλο.** Ένας εκπαιδευτικός θα μπορούσε να παρακολουθεί τις συζητήσεις που αναπτύσσουν οι μαθητές του με το chatbot και να έχει μια καλή εικόνα για την πρόοδο της τάξης, τα σημεία στα οποία οι μαθητές του είναι πιο αδύναμοι και χρειάζονται βοήθεια ή ακόμη και ποια στοιχεία της γλώσσας είναι αυτά που τους κινούν περισσότερο το ενδιαφέρον [39].

Περιορισμοί

Το όραμα με τα chatbots να παίζουν τον ίδιο ρόλο με τα εκπαιδευτικά βιβλία για τα παιδιά, να διδάσκουν ξένες γλώσσες και γενικά να είναι αυτά που θα εκπαιδεύουν είναι ακόμα ζωντανό [40]. Παρόλα αυτά η αλήθεια είναι πως απέχουμε πολύ ακόμα από το να αντικατασταθεί ο ρόλος του δασκάλου από ένα chatbot, μιας και αυτοί αποτελούν τη ραχοκοκαλιά της εκπαιδευτικής διαδικασίας. Η τεχνολογία μέσω των διαλογικών βοηθών και με τις λειτουργίες που έχουν αναπτυχθεί μέχρι τώρα, δεν μπορεί παρά να ενισχύσει και όχι να αντικαταστήσει το ρόλο ενός καλού καθηγητή [41]. Συνεπώς, έχει περισσότερο νόημα οι σχεδιαστές chatbots να τα υλοποιούν με τρόπο που να διευκολύνουν τις εργασίες των ανθρώπων και την αλληλεπίδρασή τους με τους υπολογιστές και όχι να προσπαθούν να αντικαταστήσουν πλήρως το ανθρώπινο στοιχείο. Αυτή η παρατήρηση προφανώς δεν έχει να κάνει με μελλοντικές επεκτάσεις και συνδυασμούς μέσω του τομέα της τεχνητής νοημοσύνης και των διαλογικών συστημάτων, αλλά στο τι εργαλεία έχουν αναπτυχθεί μέχρι τώρα και ποια φαίνεται ότι είναι η καλύτερη τους χρήση. Αντίστοιχα, η κύρια άποψη των εκπαιδευτικών είναι πως ο καλύτερος τρόπος να μάθει κανείς μια γλώσσα είναι να την εξασκεί μέσω συζητήσεων. Ακολουθώντας, η πρόκληση που εμφανίζεται όσον αφορά τις διάφορες μορφές υποβοηθούμενης εκμάθησης γλωσσών από τους υπολογιστές (CALL - Computer-Assisted Language Learning) και τον τρόπο που θα χρησιμοποιηθούν από αυτές οι διαλογικοί βοηθοί είναι να ξεπεράσουν τα βασικά προβλήματα που συνήθως εμφανίζονται [42]. Προβλήματα που έχουν να κάνουν με την εξάσκηση της γλώσσας σε συνθήκες καθημερινής χρήσης αυτής. Σε αυτήν την περίπτωση οι διαλογικοί βοηθοί πρέπει να αναβαθμίσουν το ρόλο τους σε πιο πλήρεις συνομιλητές, είτε αφορούν κάποιο συγκεκριμένο θέμα είτε είναι γενικού σκοπού.

Ένα χρήσιμο framework

Ακολουθεί ένα μοντέλο που παρουσιάστηκε στα πλαίσια του συνεδρίου "WI 2022" και μάλιστα στον τομέα "Human Computer Interaction & Digital Assistance Systems" [43]

2.4. Το μοντέλο αυτό ουσιαστικά είναι ένα framework που προτείνει ένα συγκεκριμένο τρόπο προσέγγισης κι εργασίας στη διαδικασία σχεδίασης και ανάπτυξης ενός διαλογικού βοηθού που θα έχει εκπαιδευτικό ρόλο. Η ανάγκη για την κατασκευή αυτού του framework

προκύπτει από την παρατήρηση ότι παρόλο που τα chatbots μπορούν να παίξουν ένα πολύ θετικό ρόλο συμβάλλοντας στην εκπαιδευτική διαδικασία, φαίνεται ότι στην πράξη υπάρχουν κάποια προβλήματα. Οι διαλογικοί βοηθοί δεν αποτελούν μια προκαθορισμένη και σπάντα λύση που εφαρμόζεται παντού με τον ίδιο τρόπο. Μάλιστα, η πεποίθηση αυτή σε συνδυασμό με την πολλές φορές πρόχειρη σχεδίαση τους, έχουν ως αποτέλεσμα αρνητικές εμπειρίες των χρηστών ή και μη επίτευξη των αρχικών στόχων τους.

Το μοντέλο αυτό, λοιπόν, διαρθρώνεται σε ένα σχεδιασμό τριών επιπέδων που το κάθε επίπεδο αφορά τις εξής θεματικές:

1. Στόχοι και μέσα.
2. Ρόλος και συσχέτιση.
3. Αξιολόγηση και αναγνώριση.

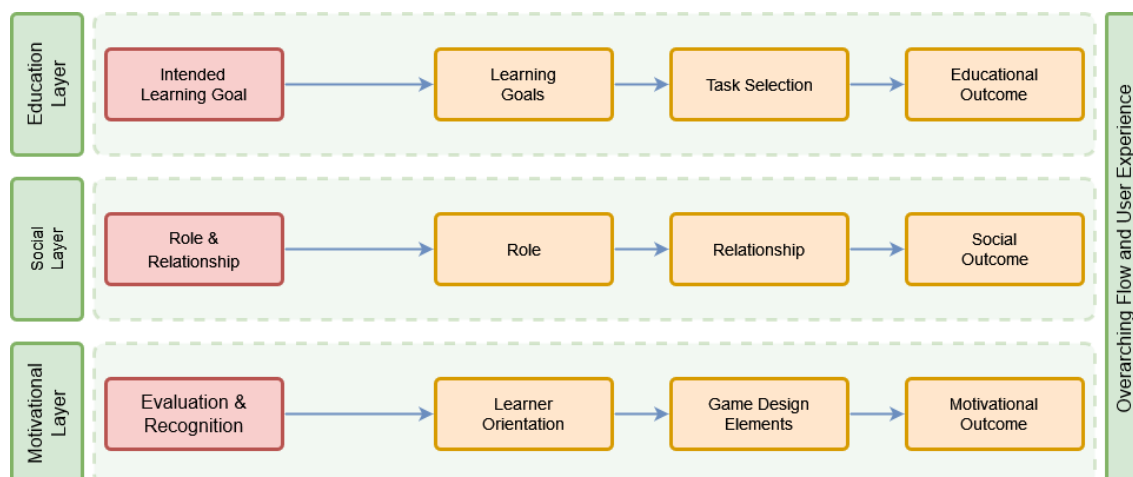
Η σχεδίαση της εκπαιδευτικής εμπειρίας χρήσης του διαλογικού βοηθού ξεκινάει με την επιλογή των στόχων που αναπαριστά το πρώτο από τα τρία επίπεδα. Ανάλογα με την επιλογή των γνωστικών στόχων υπονοούνται ή επιλέγονται αντίστοιχα και οι δραστηριότητες (π.χ. κουίζ) μέσω των οποίων θα επιτευχθούν. Εν ολίγοις, σε αυτό το στάδιο ξεκαθαρίζονται ποιες θα είναι οι γνωστικές ασκήσεις και ποια τα εκπαιδευτικά τους αποτελέσματα.

Το επόμενο βήμα αφορά την κοινωνική διάσταση του στόχου. Δηλαδή, από την επιλογή των στόχων μπορούν να προκύψουν ο ρόλος και η σχέση που θα έχει ο διαλογικός βοηθός με τον μαθητή και κατ'επέκταση το κοινωνικό αποτέλεσμα της δραστηριότητας. Η ανάλυση αυτή βασίζεται στην θεωρία του κοινωνικού παράγοντα του Lowenthal [44]. Ο στόχος του συγκεκριμένου σταδίου είναι να καθοριστούν ο ρόλος του βοηθού (π.χ. να έχει εκπαιδευτικό ρόλο) και η σχέση του με τη διαδικασία (π.χ. να έχει ως βασική δραστηριότητα να διδάσκει). Από τα παραπάνω μπορεί εύκολα να βγει ως συμπέρασμα πως όσο λιγότερο καθορισμένος είναι το πρώτο επίπεδο (στόχοι και μέσα επίτευξης), τόσο περισσότερο ο ρόλος του βοηθού θα τείνει προς την παραδοσιακή του έκφανση (δάσκαλος) και τόσο πιο τυπική θα είναι η σχέση με τους μαθητές. Από την άλλη, όσο πιο σαφώς ορισμένοι είναι οι στόχοι, τόσο πιο ελεύθερος θα μπορεί να είναι ο ρόλος του βοηθού. Η συγκεκριμένη αναλογία αποτελεί άμεση απόρροια της μελέτης των Anderson και Krathwohl πάνω στην ταξινόμηση των εκπαιδευτικών στόχων [45]. Τουτέστιν, όταν οι μαθητές καλούνται να αναλύσουν και να αξιολογήσουν μια περίπτωση, τότε ενδεχομένως να χρειάζονται για βοηθό έναν πιο σύνθετο ρόλο που να μπορεί να τους δείξει πως να εφαρμόσουν τις γνώσεις τους για να λύσουν ένα πρόβλημα. Από την άλλη, όταν οι γνώσεις εμφανίζονται ως πιο απλοϊκές και η θέση τους στο γνωστικό οικοδόμημα είναι να τεθούν τα θεμέλια ενός θέματος, μπορεί να μην απαιτείται μια τόσο σύνθετη σχέση ανάμεσα στο δάσκαλο και το μαθητή, αλλά απλά μια τυπική παρουσία με την παραδοσιακή έννοια.

Τέλος, το τελευταίο επίπεδο του framework και το τελευταίο βήμα στη διαδικασία σχεδίασης του βοηθού είναι αυτό της αξιολόγησης και της αναγνώρισης. Στην πράξη, αυτό το επίπεδο έχει να κάνει με την επιλογή των μηχανισμών και των στοιχείων αυτών που θα εξασφαλίσουν την κινητοποίηση του μαθητή για την επίτευξη των στόχων που έχουν τεθεί. Τα θεωρητικά εργαλεία για αυτό το βήμα αντλούνται από την ανάλυση της έννοιας της παιχνιδιοποίησης (gamification), με την οποία ασχολούμαστε και στις προηγούμενες ενότητες για τις εκπαιδευτικές εφαρμογές της εικονικής πραγματικότητας. Με απλά λόγια, οι εκπαιδευτικοί

στόχοι δεν πρέπει να είναι ούτε πολύ βαρετοί, ούτε υπερβολικοί. Η αλληλεπίδραση με τον διαλογικό βοηθό δεν πρέπει να είναι ούτε τετριμμένη, αλλά ούτε και να δυσκολεύει το μαθητή (π.χ. προκαλώντας του άγχος ή πίεση). Πρέπει, δηλαδή, να υπάρχει μια ισορροπία ανάμεσα στη βαρεμάρα και το άγχος διατηρώντας ροή καθόλη τη διάρκεια των δραστηριοτήτων και της αλληλεπίδρασης με τον βοηθό [46].

Εδώ να σημειωθεί πως η συγκεκριμένη εργασία πάνω στη σχεδίαση και την υλοποίηση εκπαιδευτικών διαλογικών βοηθών είναι απλά μια θεωρητική μεθοδολογία που προσπαθεί να καλύψει το κενό ανάμεσα στους εκπαιδευτικούς στόχους και τα εργαλεία ή τα στοιχεία που κάθε φορά ο σχεδιαστής θέλει να προσδώσει στο εργαλείο που δημιουργεί. Μένει, λοιπόν, να αποδειχθεί στην πράξη πως αυτό το framework θα αξιοποιηθεί και στη δική μας περίπτωση που ο διαλογικός βοηθός δεν θα είναι απλά ένα chatbot, αλλά ένα ολόκληρο περιβάλλον εικονικής πραγματικότητας με ενσωματωμένο ένα διαλογικό σύστημα.



Σχήμα 2.4: Framework σχεδίασης ενός εκπαιδευτικού διαλογικού βοηθού.

2.5 Σχετικές εφαρμογές

Πρώτα ιστορικά βήματα

Πρόκειται για εφαρμογές που ο σκοπός τους ήταν κυρίως ψυχαγωγικός, είτε από τη σκοπιά των video games, είτε από αυτή των κοινωνικών δικτύων, που όμως κατέληξαν εκτός των άλλων να χρησιμοποιηθούν και ως εργαλεία εκμάθησης ξένων γλωσσών [47].

Ενδιαφέρον έχουν οι περιπτώσεις του "Colossal Cave Adventure" και των "Traveler", "Active Worlds", αλλά και του αρκετά ειδικευμένου "Babel-M: A Virtual Environment For the Promotion Of Language Learning". Το πρώτο ήταν μια text-only προσομοίωση (1970s), όπου ο χρήστης επικοινωνούσε με τον υπολογιστή αποκλειστικά γράφοντας εντολές στο πληκτρολόγιο. Στα πλαίσια αυτού του παιχνιδιού, αναπτύχθηκαν τα λεγόμενα MOOs (Multi-user domains object-oriented), τα οποία καθηγητές γλωσσών χρησιμοποιούσαν για να διδάσκουν ξένες γλώσσες. Τα άλλα τρία (1990s) ήταν το επόμενο μεγάλο βήμα, όπου πλέον δινόταν η δυνατότητα και για ηχητική επικοινωνία ανάμεσα στα avatars, τα οποία αναπαρίστανται ως ασώματα κεφάλια σε ένα τρισδιάστατο τοπίο.

Η ενδιαφέρουσα περίπτωση του Second Life

Το videogame "Second Life" (κυκλοφόρησε το 2003 και παρόλο που δεν είναι πια τόσο διαδεδομένο, όσο ήταν την περίοδο που υπήρχε hype γύρω από αυτό, αποτελεί ακόμα χαρακτηριστικό παράδειγμα εφαρμογών που υλοποιούν έναν virtual world) έχει χρησιμοποιηθεί ευρύτατα για το σκοπό της εκμάθησης γλωσσών. Μάλιστα, σύμφωνα με τους ίδιους τους δημιουργούς του παιχνιδιού "η εκμάθηση γλωσσών είναι η πιο συχνή εκπαιδευτική δραστηριότητα στο Second Life". Πιο συγκεκριμένα, πολλά πανεπιστήμια, γλωσσολογικά ινστιτούτα και ερευνητικά κέντρα έχουν "νησιά" στο σύμπαν του Second Life, ειδικευμένα στην εκμάθηση γλωσσών.

Σύγχρονες εφαρμογές

Με την πάροδο των χρόνων και την ολοένα και επιταχυνόμενη βελτίωση του εξοπλισμού και των λογισμικών που χρησιμοποιούνται για εμπειρίες εικονικής πραγματικότητας, έχουν συντελεστεί ταυτόχρονα, αν και ακόμη αρκετά δευτερεύοντα, βήματα στον τομέα εφαρμογών εκμάθησης γλωσσών. Ακολουθούν έξι σύγχρονες τέτοιες εφαρμογές που αξίζουν να αναφερθούν:

- Η **VR Learn English** είναι μια εφαρμογή για μελέτη του αγγλικού λεξιλογίου. Σε αυτήν ο χρήστης κινείται στο εσωτερικό ενός δωματίου και μπορεί να ακούσει την προφορά των ονομάτων διαφόρων αντικειμένων που βρίσκονται γύρω του.
- Η **Gold Lotus** είναι παρόμοια με την "VR Learn English", με τη διαφορά ότι τα μέρη και οι χώροι, στα οποία ο χρήστης μπορεί να περιηγηθεί, ποικίλουν. Το διδακτικό υλικό ακόμη εκτός από το λεξιλόγιο που έχει, συμπεριλαμβάνει και μαθήματα γραμματικής.
- Η **Mondly VR** είναι η πρώτη εφαρμογή που χρησιμοποιεί ταυτόχρονα τεχνολογίες επαυξημένης πραγματικότητας (augmented reality), διαλογικών συστημάτων (chatbot),

καθώς και αναγνώρισης φωνής (ASR) για την εκμάθηση γλωσσών. Μέχρι τώρα, διαθέτει πακέτα δραστηριοτήτων για τριάντα τρεις τουλάχιστον γλώσσες και πέρα από τα χαρακτηριστικά που έχουν οι προηγούμενες δύο εφαρμογές, συμπεριλαμβάνει και ασκήσεις κατανόησης κι ανάπτυξης προφορικού λόγου (listening and speaking). Αυτό όμως που κάνει τη συγκεκριμένη εφαρμογή πραγματικά ξεχωριστή είναι η δυνατότητα του χρήστη να παίρνει μέρος σε συζητήσεις με εικονικούς χαρακτήρες σε ρεαλιστικές συνθήκες. Εξαιτίας των καινοτομιών αυτής της εφαρμογής, αλλά και της μεγάλης εμπορικής της επιτυχίας, η περίπτωση της έχει καταστεί σημείο αναφοράς. Για το λόγο αυτό, στη συνέχεια της ενότητας θα μελετήσουμε εκτενώς τα χαρακτηριστικά της, ενώ τα συμπεράσματα που θα εξαγάγουμε θα χρησιμοποιηθούν στο κομμάτι του σχεδιασμού της δικής μας εφαρμογής.

- Η **VRSpeech** είναι μια εφαρμογή που η χρήση της στοχεύει κυρίως στην προφορική εξάσκηση και στον εμπλουτισμό του λεξιλογίου του χρήστη. Περιλαμβάνει δύο διαφορετικά ενδεχόμενα χρήσης. Το πρώτο είναι το στάδιο εκπαίδευσης (learning mode) του χρήστη και το δεύτερο είναι το στάδιο εξάσκησης (situation mode). Όλα τα σενάρια στα οποία θα εμπλακεί ο χρήστης είναι ρεαλιστικά: παραγγελία σε ένα εστιατόριο, συνέντευξη για εργασία, παρουσίαση κλπ. Επίσης, διαθέτει υψηλής ποιότητας αναγνώριση φωνής και εξειδικευμένο λεξιλόγιο ανά την περίπτωση.
- Αντίστοιχων χαρακτηριστικών περίπτωση είναι και η **VirtualSpeech**. Αυτό που κάνει τη συγκεκριμένη εφαρμογή, όμως, να ξεχωρίζει είναι η πολύ αξιόλογη δουλειά που έχει γίνει στην κατεύθυνση του να γίνει η VR εμπειρία ακόμη πιο εμβυθιστική. Έχουν σχεδιαστεί ρεαλιστικοί χώροι εκδηλώσεων με κοινό, παρουσία θορύβου και αποσπάσεων, ώστε ο χρήστης να αισθάνεται πως είναι σε ένα ρεαλιστικό σενάριο.
- Τέλος, αξίζει να αναφέρουμε και την **AltspaceVR** η οποία είναι μια εφαρμογή που δεν εστιάζει στην εκμάθηση γλωσσών, αλλά είναι γενικού σκοπού. Κόσμος τη χρησιμοποιεί για να διοργανώσει ή να παρευρεθεί σε εικονικές εκδηλώσεις κάθε είδους [48].

Mondly VR

Η εφαρμογή αυτή βγήκε στην αγορά για πρώτη φορά το 2017, ενώ το 2021 έγινε διαθέσιμη και στο Oculus 2 (συσκευή για VR εμπειρίες για την οποία θα μιλήσουμε στη συνέχεια). Με το που ξεκινήσει να τρέχει, καλεί τον χρήστη να επιλέξει μητρική γλώσσα, καθώς και τη γλώσσα που μαθαίνει. Η Mondly VR έχει δύο επιλογές λειτουργίας. Η πρώτη αφορά εξάσκηση πάνω σε συγκεκριμένο κάθε φορά λεξιλόγιο, ενώ η δεύτερη περιλαμβάνει ασκήσεις διαλόγων [49].

Στην πρώτη περίπτωση η εξάσκηση βασίζεται κυρίως στη σωστή προφορά των χρησιμων λέξεων, με την αξιολόγηση αυτή να γίνεται μέσω ενός συστήματος αναγνώρισης φωνής (ASR). Η εξάσκηση γίνεται με πρακτικό και βολικό τρόπο, καθώς ο χρήστης έχει το αντικείμενο μπροστά του και αλληλεπιδρώντας με αυτό, μαθαίνει πώς γράφεται και πώς προφέρεται. Βέβαια, οι λέξεις δε χρησιμοποιούνται στα πλαίσια κάποιας πρότασης ή έκφρασης. Εξαιτίας αυτού, αν και οι ασκήσεις λεξιλογίου προετοιμάζουν κάπως τον χρήστη για σχετικές με το κάθε σενάριο λέξεις, δεν είναι πολύ χρήσιμες στο πώς να τις χρησιμοποιήσει στο λόγο

του. Παρόλο που οι ασκήσεις λεξιλογίου γίνονται σε ένα χώρο που είναι σχετικός με τις λέξεις που παρουσιάζονται (π.χ. σε έναν πάγκο μαθαίνεις τα φρούτα, στο διαστημικό σταθμό μαθαίνεις για το διάστημα), δε συνδέονται άμεσα με κάποια συζήτηση. Δεν γίνεται, δηλαδή, ούτε υπονοείται πάντα μια συσχέτιση των λέξεων με τις συνθήκες συζήτησης που αυτές χρησιμοποιούνται (όπως όταν παραγγέλνεις φαγητό ή παίρνεις ένα ταξί κ.λπ.).

Μετά τις ασκήσεις λεξιλογίου ακολουθούν οι ασκήσεις διαλόγου. Αυτές διακρίνονται σε οκτώ διαφορετικά σενάρια :

1. Χαιρετισμός (γνωριμία με έναν άνθρωπο σε ένα βαγόνι τρένου).
2. Εξυπηρέτηση σε ένα ταξί.
3. Κράτηση ενός δωματίου σε ξενοδοχείο.
4. Αίτημα για αλλαγή του δωματίου.
5. Επίσκεψη σε ένα εστιατόριο.
6. Αγορά ρούχων σε κατάστημα.
7. Κλείσιμο ραντεβού για το βράδυ.
8. Αγορά ενός εισιτηρίου για το τρένο.

Στην αρχή του κάθε σεναρίου, ένας άνθρωπος καλωσορίζει τον χρήστη καλώντας έτσι να εμπλακεί σε έναν σύντομο διάλογο. Ο κάθε ένας από αυτούς τους διαλόγους έχει τα εξής χαρακτηριστικά. Για να μιλήσει ο χρήστης πρέπει να πατήσει το κουμπί του μικροφώνου, που ενεργοποιεί το ASR. Το πλήθος των αποδεκτών απαντήσεων είναι κάθε φορά τρεις. Ενώ, το μονοπάτι του διαλόγου είναι πολύ συγκεκριμένο. Με το που ολοκληρωθεί το μονοπάτι διαλόγου, το σενάριο τερματίζεται και εμφανίζεται ένα νέο παράθυρο που δείχνει πόσο καλά τα πήγε ο χρήστης σύμφωνα με το ASR. Για τα λόγια του χαρακτήρα του σεναρίου καθώς και για τις αποδεκτές απαντήσεις σε κάθε γύρο του διαλόγου, υπάρχει κι ένα σχετικό παράθυρο που δίνει τη δυνατότητα στον χρήστη να ακούσει την κάθε πρόταση, καθώς και να διαβάσει τη μετάφραση της, εφόσον το επιθυμεί. Τέλος, υπάρχει η δυνατότητα να ακουστεί η απάντηση του χαρακτήρα με πιο αργή ομιλία, καθώς και η δυνατότητα να παραλειφθεί κάποιο βήμα.

Στο κομμάτι του **πώς αξιολογείται η εφαρμογή** αυτή μπορούν να ειπωθούν τα εξής. Τα σενάρια κρίνονται ως εμπυθιστικά (immersive) και φαίνεται ότι πετυχαίνουν να κρατηθεί το ενδιαφέρον του χρήστη. Παρόλο, όμως, που η εφαρμογή δίνει τη δυνατότητα στον χρήστη να επιλέξει το επίπεδο του, από επίπεδο σε επίπεδο δεν εντοπίζεται καμία διαφορά στο περιεχόμενο των σεναρίων. Τα σενάρια σε επίπεδο διαλόγου δεν ξεπερνούν τις δεκαπέντε ερωτήσεις κι απαντήσεις, με πολλές από αυτές μάλιστα να πρόκειται για φράσεις του τύπου "Γεια", "Σας ευχαριστώ", "Αντίο" κ.λπ.. Αυτό έχει ως αποτέλεσμα παρόλο που η προσέγγιση του εκπαιδευτικού ρόλου της εφαρμογής να βασίζεται στην ύπαρξη ρεαλιστικών σεναρίων, η επιτυχία του χρήστη και η αξιολόγησή του στο επίπεδο της ολοκλήρωσης των σεναρίων να εξαρτάται κυρίως από την ικανότητά του να απομνημονεύει τους διαλόγους. Μια κίνηση που ο κάθε χρήστης θα κάνει αναπόφευκτα μιας και τα σενάρια, έτσι όπως έχουν υλοποιηθεί, δεν επιτρέπουν οποιαδήποτε απόκλιση από τα μονοπάτια διαλόγου που έχουν.

Τέλος, **οι μελλοντικές επεκτάσεις και βελτιώσεις** που απορρέουν από τις παραπάνω παρατηρήσεις είναι οι εξής. Θα βοηθούσε η αλληλεπίδραση του χρήστη με το περιβάλλον του εικονικού κόσμου να είναι πιο ελεύθερη. Η εμπειρία, δηλαδή, κρίνεται ότι θα ήταν πολύ πιο πλούσια **αν οι χρήστες μπορούσαν να κινούνται στο χώρο**, να μετακινούν αντικείμενα, να συλλέγουν πληροφορίες από το περιβάλλον τους, να διαβάζουν γράμματα και καταλόγους κ.λπ.. Επίσης, αν και η προηγούμενη επέκταση φαίνεται πολύ σημαντική, το μεγαλύτερο μειονέκτημα της εφαρμογής αυτής είναι οι **περιορισμένοι διάλογοι** και τα σύντομα σενάρια. Το ASR δεν είναι γενικού σκοπού και είναι περιορισμένες οι προτάσεις που μπορούν να αναγνωριστούν από τα λόγια του χρήστη σε αυτές που μπορούν να συνεχίσουν στο προκαθορισμένο μονοπάτι διαλόγου. Το NLP (Natural Language Processing) από την άλλη είναι πολύ αδύναμο κι αυτό και εξυπηρετεί μόνο όσες φράσεις εντάσσονται στο πλαίσιο του κάθε σεναρίου. Τέλος, μια πολύ ενδιαφέρουσα επέκταση στο Mondly VR θα ήταν η περαιτέρω σύνδεση των σεναρίων με την κουλτούρα που έχει η κάθε γλώσσα. Μια επέκταση που ουσιαστικά θα αξιοποιούσε και το βασικό πλεονέκτημα χρήσης της εικονικής πραγματικότητας στον τομέα της εκμάθησης γλωσσών, την άμεση επαφή με ρεαλιστικά σενάρια και περιβάλλοντα που αλλιώς δε θα γινόταν.

Συμπεράσματα

Από τις παραπάνω περιπτώσεις ενδεχομένως δόθηκε μια εικόνα του πόσο καλές προοπτικές έχουν οι τεχνολογίες εικονικής πραγματικότητας για την εκμάθηση γλωσσών. Φάνηκε, όμως, ότι έχουν πολύ δρόμο ακόμη να διανύσουν, μέχρι να καθιερωθούν ως χρήσιμο εργαλείο στον τομέα αυτό. Περιπτώσεις σαν αυτή της εφαρμογής Mondly VR είναι πολύ καλές για να φανούν τα πλεονεκτήματα και προοπτικές τους, τα σημεία στα οποία πρέπει να βελτιωθούν, αλλά και κάποιους κινδύνους που ελλοχεύουν και πρέπει να αποφευχθούν. Οι VR εμπειρίες των πιο επιτυχημένων εφαρμογών στοχεύουν στο να αναπαριστούν σχεδόν ρεαλιστικούς διαλόγους σε πραγματικό χρόνο με ένα avatar. Ωστόσο, πέρα από το να βελτιώσουν τους διαλόγους αυτούς ακόμη περισσότερο, ώστε να γίνουν πιο ρεαλιστικοί και να πλησιάσουν την πραγματική ζωή, θα μπορούσαν να εντάξουν στις λειτουργίες τους και την αλληλεπίδραση ανάμεσα σε πολλαπλούς χρήστες σε ένα εμπυθιστικό περιβάλλον. Με άλλα λόγια, αντί ένας χρήστης απλά να αλληλεπιδρά με έναν τεχνητό ανθρώπινο χαρακτήρα, θα μπορούσε να μιλά σε πραγματικό χρόνο με άλλους χρήστες, είτε αυτοί είναι μαθητές, είτε δάσκαλοι, είτε απλοί άνθρωποι που έχουν ως μητρική γλώσσα τη γλώσσα που μαθαίνει ο ίδιος. Υπάρχουν εφαρμογές που προσπαθούν να δημιουργήσουν τέτοια περιβάλλοντα, αλλά δεν υπάρχει καμία που να εμπεριέχει και τις δύο περιπτώσεις αλληλεπίδρασης.

Κεφάλαιο 3

Σχεδίαση

Σε αυτό το κεφάλαιο γίνεται μια σκιαγράφηση των σχεδιαστικών αποφάσεων που αφορούν την εφαρμογή που υλοποιήθηκε. Οι αποφάσεις αυτές πάρθηκαν με οδηγό τα συμπεράσματα από τη μελέτη της θεωρίας του προηγούμενου κεφαλαίου και την επισκόπηση των σχετικών εφαρμογών που ήδη υπάρχουν. Οι ενότητες που αποτελούν αυτό το κεφάλαιο είναι οι εξής:

1. Baseline μιας VRALL εφαρμογής και τι θέλουμε να βελτιώσουμε.
2. Χαρακτηριστικά της VRALL εφαρμογής.

3.1 Baseline μιας VRALL εφαρμογής και τι θέλουμε να βελτιώσουμε

Από την επισκόπηση σχετικών με την εκμάθηση γλωσσών εφαρμογών που έγινε σε προηγούμενη ενότητα, είναι ξεκάθαρο πως η περίπτωση της **Mondly VR** είναι η πιο ενδιαφέρουσα. Όσον αφορά τα τεχνικά στοιχεία της εφαρμογής αυτό που τη διακρίνει από τις περισσότερες είναι ότι προσπαθεί να δημιουργήσει πραγματικά εμπυθιστικές εμπειρίες με την δημιουργία χώρων και σεναρίων σε πλήρως VR περιβάλλοντα. Αυτό όμως που την κάνει να θεωρείται μια state-of-the-art εφαρμογή στο χώρο είναι η εκπαιδευτική μέθοδος που υιοθετεί. Πιο συγκεκριμένα, με βάση και τη διάκριση που γίνεται στο αντίστοιχο κομμάτι της θεωρίας, η Mondly VR δίνει περισσότερη βαρύτητα στην εξάσκηση των γνώσεων του χρήστη μέσω αναπαράστασης ρεαλιστικών συνθηκών συζήτησης, παρά σε μια παραδοσιακή προσέγγιση εκμάθησης γλωσσών. Με δεδομένο, λοιπόν, πως ο σκοπός της παρούσας εργασίας είναι η δημιουργία μιας τέτοιας εφαρμογής εικονικής πραγματικότητας που δε θα προσπαθεί να αντικαταστήσει τη συνήθη εκπαιδευτική διαδικασία, αλλά να προσφέρει ένα εργαλείο που θα τη διευκολύνει και θα της δίνει νέες δυνατότητες φαίνεται ότι η Mondly VR είναι το baseline για τη VRALL εφαρμογή που σχεδιάζουμε. Θα αποτελεί, δηλαδή, το βασικό σημείο αναφοράς για το τι ακριβώς θέλουμε να βελτιώσουμε σε σχέση με το τι υπάρχει εκεί έξω και τι θέλουμε να δοκιμάσουμε ως καινοτομία. Βεβαίως, και οι άλλες εφαρμογές που σχολιάστηκαν έχουν ενσωματώσει στοιχεία τα οποία είναι πολύ χρήσιμα, όπως είναι τα επίπεδα εμπυθιστικότητας στην **VirtualSpeech**, ή το αρκετά εξειδικευμένο λεξιλόγιο σχετικό με το σενάριο και το αντίστοιχο σύστημα αναγνώρισης ομιλίας που υποστηρίζει το **VRSpeech**. Καμία όμως από

τις δύο, αν και σε αυτά τα συγκεκριμένα χαρακτηριστικά τους πετυχαίνουν καλύτερα αποτελέσματα, δεν καταφέρνει να συνδυάσει τόσο υψηλό επίπεδο σε συνδυαστικούς τομείς, όπως η Mondly VR.

Όμως, παρόλο που τα χαρακτηριστικά της εφαρμογής αυτής είναι πολύ καλά κι αυτό αντικατοπτρίζεται και στο μερίδιο που καλύπτει στη σχετική αγορά, υπάρχουν σημεία τα οποία χρήζουν βελτίωσης (βλ. αξιολόγηση σχετικών εφαρμογών). Με την ανάπτυξη του prototype της δικής μας VRALL εφαρμογής θα προσπαθήσουμε όχι μόνο να φτάσουμε το baseline που μόλις ορίσαμε, αλλά σε κάποια από αυτά τα αδύναμα σημεία να προσφέρουμε καλύτερες λύσεις.

Κίνηση

Το πρώτο στοιχείο που θέλουμε να βελτιώσουμε είναι η παρουσία του χρήστη στον εικονικό κόσμο της εφαρμογής. Αναγνωρίζοντας πως ένα από τα βασικότερα μειονεκτήματα της εφαρμογής του baseline είναι το γεγονός πως ο χρήστης βρίσκεται συνεχώς καθηλωμένος στο ίδιο σημείο. Έτσι, βασικό σχεδιαστικό στοιχείο είναι η ικανότητα του χρήστη να κινείται στο χώρο και να έχει όσο το δυνατόν μεγαλύτερη αυτονομία στην κίνηση του αυτή. Αυτό σημαίνει ότι ο χώρος του σεναρίου δε θα είναι ένα φόντο απομακρυσμένο από τον χρήστη, αλλά ένας εικονικός τρισδιάστατος χώρος που θα προσφέρεται για παρατήρηση κι εξερεύνηση.

Διάρκεια και ποικιλία διαλόγων

Το δεύτερο σημείο στο οποίο θέλουμε να εμβαθύνουμε είναι αυτό των διαλόγων. Μάλιστα, στο σημείο αυτό της διαλογικής αλληλεπίδρασης η Mondly VR δέχεται τη μεγαλύτερη κριτική και φαίνεται ότι έχει τις περισσότερες προοπτικές βελτίωσης. Για παράδειγμα, θέλουμε η εφαρμογή που σχεδιάζουμε να προσφέρει στον χρήστη διαλόγους με μεγαλύτερη διάρκεια αλλά και ποικιλία. Επίσης, θέλουμε να υπάρχει ελαστικότητα στο πως εξελίσσονται οι διάλογοι. Για να επιτευχθεί αυτό θα πρέπει να στηθεί ένα διαλογικό σύστημα που τα μονοπάτια συζητήσεων δεν είναι προκαθορισμένα με τον τρόπο που είναι στην Mondly VR, αλλά εμφανίζουν ευελιξία όπως γίνεται και στην πραγματική ζωή.

Hints

Επίσης, η λειτουργία των hints αναγνωρίζεται ως πολύ σημαντικό στοιχείο της εφαρμογής. Αυτό σημαίνει ότι δεν πρέπει να είναι απλά ένα μέσο που δίνει τη δυνατότητα στον χρήστη να συνεχίζει τη συζήτηση, αλλά να αναδειχθεί ως αναπόσπαστο κομμάτι της εκπαιδευτικής μεθόδου. Με δεδομένο ότι θα έχει δημιουργηθεί ένα πιο ικανό διαλογικό σύστημα που θα μπορεί να υποστηρίξει μεγαλύτερες και πολυπλοκότερες συζητήσεις με το χρήστη, θα μπορεί να στηθεί κι ένα συνοδευτικό σύστημα με hints που δε θα λέει απλά στον χρήστη τι περιμένει το διαλογικό να ακούσει από αυτόν για να συνεχίσει, αλλά θα του δίνει μια ή περισσότερες προτάσεις για το πως θα μπορούσε να εξελιχθεί η συζήτηση.

Αναγνώριση φωνής

Από τις παραπάνω προσθήκες φαίνεται ότι η εφαρμογή που σχεδιάζεται χρειάζεται κι ένα πιο ικανό σύστημα αναγνώρισης φωνής. Πιο συγκεκριμένα, με την ποικιλία και το ενδεχόμενο βάθος των συζητήσεων να έχουν αυξηθεί, δεν μπορούμε να βασιζόμαστε σε ένα σύστημα ASR αρκετά πιο διευρυμένο από αυτό που έχει το Mondly ή έστω που δεν είναι γενικού σκοπού. Θέλουμε, δηλαδή, να πετύχουμε αναγνώριση ομιλίας του χρήστη σε όσο το δυνατόν πιο υψηλό επίπεδο και όχι απλά στο βαθμό που χρειάζεται για να εξελίσσεται κάθε φορά ο διάλογος.

Σημειώνοντας πραγματική πρόοδο στα σημεία που αναφέρονται από πάνω, η εφαρμογή θα έχει ολοκληρώσει ένα πολύ σημαντικό βήμα προς την κατεύθυνση των πραγματικά ρεαλιστικών σεναρίων χρήσης της γλώσσας. Όσο σημαντικό ρόλο παίζουν τα οπτικά, τα ακουστικά και γενικά τα αισθητικά χαρακτηριστικά ενός εικονικού κόσμου για την εμπύθιση σε αυτόν, άλλο τόσο κι ενδεχομένως ακόμη περισσότερο παίζουν τα στοιχεία της ανθρώπινης αλληλεπίδρασης σε αυτόν. Σε γενικές γραμμές αυτά είναι τα βασικά στοιχεία στα οποία θέλουμε να πετύχουμε καλύτερα αποτελέσματα.

3.2 Χαρακτηριστικά της VRALL εφαρμογής

Μέχρι τώρα ασχοληθήκαμε με τα χαρακτηριστικά της εφαρμογής που αναγνωρίζονται ως σημεία βελτίωσης και δοκιμής καινοτομιών σε σχέση με το baseline, έτσι όπως το ορίσαμε στην προηγούμενη ενότητα. Πρέπει, όμως, να προσδιορίσουμε και ένα σύνολο από κάποια άλλα χαρακτηριστικά που είναι απαραίτητα για τη συνολική σχεδίαση της εφαρμογής και την προετοιμασία της υλοποίησης μας. Κάποια από τα χαρακτηριστικά αυτά έχουν να κάνουν με την κατηγοριοποίηση της εφαρμογής με βάση τα στοιχεία εικονικής πραγματικότητας, κάποια άλλα αφορούν τις εκπαιδευτικές μεθόδους που θα ακολουθηθούν, έτσι όπως αυτές προκύπτουν από την αντίστοιχη θεωρία και κάποια άλλα έχουν να κάνουν με το πιο τεχνικό κομμάτι της υλοποίησης και αφορούν την σκιαγράφηση της αρχιτεκτονικής της εφαρμογής. Επίσης, πρέπει να γίνουν πιο συγκεκριμένα βασικά στοιχεία της εφαρμογής, όπως το πώς θα διαρθρώνεται η ίδια, πόσα σενάρια θα υλοποιηθούν και τι χαρακτηριστικά θα έχει το καθένα κ.ά..

3.2.1 Κύριες λειτουργίες

Ο κύριος σκοπός της εφαρμογής είναι η **εξάσκηση** της χρήσης των ελληνικών και όχι η εκμάθηση αυτών. Αυτό σημαίνει πως η χρήση της προϋποθέτει την ύπαρξη ενός υποδάθρου γνώσεων. Δηλαδή, αν θα μπορούσε να συσχετιστεί με μια εκπαιδευτική διαδικασία, η σχέση τους θα ήταν συμπληρωματική και όχι σχέση αντικατάστασης της μίας από την άλλη. Συνεπώς, δε θα περιέχει τόσο έντονα στοιχεία που αφορούν την εκμάθηση γλωσσών, όπως είναι η μετάφραση λεξιλογίου ή η παρουσίαση γραμματικών και συντακτικών κανόνων. Αντίθετα, θα προσπαθεί να προσομοιώσει ρεαλιστικές συνθήκες χρήσης της γλώσσας, καλώντας τους χρήστες να αντεπεξέλθουν σε αυτές. Για αυτό το λόγο, αποφασίστηκε η χρήση της να είναι αποκλειστικά στα ελληνικά.

Για λόγους οικονομίας αποφασίστηκε η VRALL εφαρμογή που θα αναπτυχθεί στα πλαίσια της συγκεκριμένης εργασίας να μην είναι σε ολοκληρωμένη μορφή, αλλά να αποτελεί ένα prototype στο οποίο θα υλοποιηθούν τα βασικά λειτουργικά της στοιχεία και θα δοκιμαστούν οι επεκτάσεις και οι βελτιώσεις που περιγράφονται πιο πάνω. Πιο συγκεκριμένα, το prototype αυτό θα περιέχει ένα μόνο σενάριο, το οποίο θα διαδραματίζεται στο λόμπι ενός ξενοδοχείου. Ο μόνος ανθρῶπινος χαρακτήρας που θα υπάρχει στον εικονικό χώρο θα είναι η υπάλληλος της υποδοχής με την οποία ο χρήστης θα μπορεί να συνομιλήσει. Ο χρήστης θα μπορεί είτε να κινηθεί στο χώρο και να εξοικειωθεί με το εικονικό περιβάλλον της εφαρμογής, είτε να μιλήσει στην υπάλληλο.

Κατ' επέκταση, οι βασικές λειτουργίες που θα υλοποιηθούν για αυτό το σενάριο και θα σχετίζονται με την αλληλεπίδραση του χρήστη με τη ρεσεψιονίστ θα είναι οι εξής:

- Αναγνώριση φωνής γενικού σκοπού που θα είναι ικανό να αναγνωρίζει οτιδήποτε πει ο χρήστης. Καθόλη τη διάρκεια της συζήτησης θα υπάρχει σχετικό παράθυρο στο οπτικό πεδίο του χρήστη που θα εμφανίζονται τα λόγια του χρήστη.
- Διαλογικό σύστημα που επεξεργάζεται τα λόγια του χρήστη και θα ανταποκρίνεται κατάλληλα δίνοντας την απάντηση της υπαλλήλου.

- Παραγωγή των hints που θα πρέπει να πληρούν τις προϋποθέσεις που έχουν οριστεί.
- Σύνθεση φωνής για τα λόγια της υπαλλήλου της ρεσεψιόν καθώς και για τα hints, προκειμένου να μπορεί ο χρήστης να τα ακούσει. Και για την απάντηση της υπαλλήλου και για τα hints θα υπάρχουν αντίστοιχα παράθυρα, για να βοηθούν τον χρήστη.
- Ο χαρακτήρας της υπαλλήλου δε θα πρέπει να είναι ένα ακίνητο avatar, αλλά θα πρέπει να μπορεί να κινείται. Μιας και ο σκοπός της εφαρμογής είναι η εξάσκηση στην ομιλία μιας γλώσσας, μια από τις δυνατές κινήσεις του χαρακτήρα της υπαλλήλου πρέπει να αφορά τα χείλη της (lipsync).

3.2.2 Εκπαιδευτικές μέθοδοι

Τα στοιχεία που θέλουμε να έχει η εφαρμογή και οι εκπαιδευτικές μέθοδοι με τις οποίες θα συνδέονται είναι τα εξής.

- **Exploring:** Ο χρήστης μπορεί να κινηθεί και σε ένα στοιχειώδη βαθμό να εξερευνήσει το χώρο γύρω του
- **Collaborating:** Για την επίτευξη των στόχων του σεναρίου ο χρήστης θα πρέπει να συνεργαστεί με έναν εικονικό χαρακτήρα συζητώντας.
- **Championing:** Ο χρήστης θα πρέπει να εμπλέκεται με ενεργό τρόπο στη συζήτηση με τον χαρακτήρα του σεναρίου. Για αυτό οι διάλογοι πρέπει να ανταποκρίνονται σε σενάρια της πραγματικής ζωής κι έτσι θα έχει επιπλέον κίνητρο να πετύχει.
- **Λειτουργική εξάσκηση:** Στόχος της εφαρμογής δεν θα είναι η τυπική εξάσκηση προκαθορισμένων φράσεων μέχρι να αναπαράγονται με σωστό τρόπο, αλλά η λειτουργική εξάσκηση με συμμετοχή σε συζητήσεις.
- **Gamification και εμπύθιση:** Τα σενάρια χρήσης της ελληνικής γλώσσας θα πρέπει να δημιουργούν ένα ρεαλιστικό περιβάλλον γύρω από τον χρήστη. Επίσης, οι εκπαιδευτικοί στόχοι δε θα πρέπει να έχουν ένα σαφή και κεντρικό χαρακτήρα στην εμπειρία της εφαρμογής, όπως στις παραδοσιακές μεθόδους εκμάθησης γλωσσών, αλλά να επιτυγχάνονται με έμμεσο και πιο διασκεδαστικό τρόπο.

Με αυτές τις μεθόδους υπόψιν, αλλά και το σχετικό σχεδιαστικό πλαίσιο που είδαμε στη θεωρία θα ορίσουμε το στόχο, τα μέσα και κάποια βασικά στοιχεία της εφαρμογής. Ο κύριος εκπαιδευτικός **στόχος** της είναι η εξάσκηση του χρήστη στα ελληνικά και μάλιστα στο επίπεδο της διαπροσωπικής επικοινωνίας. Το κύριο **μέσο** με το οποίο θα επιτυγχάνεται ο εκπαιδευτικός σκοπός είναι η συμμετοχή του χρήστη σε ρεαλιστικούς διαλόγους, όπου θα έχει την ευκαιρία να χρησιμοποιήσει κάθε φορά ειδικό για την περίπτωση λεξιλόγιο. Κατά τη χρήση της εφαρμογής θα υπάρχουν διάφορα μέσα αλληλεπίδρασης με τον χρήστη, με τα βασικότερα από αυτά να είναι ένας ανθρώπινος χαρακτήρας και κάποια στοιχεία UI (παράθυρα). Ο **ρόλος** τους, με κεντρικό στοιχείο να είναι ο χαρακτήρας του σεναρίου, θα είναι να συνομιλεί με τον χρήστη. Δηλαδή, ο γύρω χώρος και τα παράθυρα με τις οδηγίες θα είναι η εισαγωγή του μαθητή στο περιεχόμενο του σεναρίου, ενώ ο συνομιλητής θα είναι

το κύριο μέσο για την εξάσκηση των γλωσσικών του ικανοτήτων. Το τελευταίο κομμάτι του framework είναι αυτό των μηχανισμών που θα εξασφαλίσουν την κινητοποίηση του χρήστη για την επίτευξη των στόχων που έχουν τεθεί με βάση τη θεωρία του **gamification**. Οι διάλογοι και οι θεματικές που θα έχει ο καθένας από αυτούς δε θα πρέπει να είναι τόσο εξεζητημένοι, αλλά ούτε τόσο βαρετοί. Ακόμη, ο τρόπος που θα στηθεί όλο το σενάριο (χώρος, ανθρώπινο avatar, UI κ.ά.) θα πρέπει να δημιουργούν την αυτοπεποίθηση στον χρήστη ότι μπορεί να τα διαχειριστεί, χωρίς όμως αυτό να γίνεται τόσο εύκολο που να γίνονται τετριμμένα. Με βάση τα παραπάνω, βγαίνουν οι εξής σχεδιαστικές αποφάσεις. Στην αρχή του σεναρίου θα πρέπει να ορίζονται οι στόχοι του χρήστη και να δίνονται οι απαραίτητες οδηγίες για την επίτευξη αυτών. Αυτοί οι στόχοι θα έχουν να κάνουν με τις θεματικές των διαλόγων που θα αναπτυχθούν και τα πράγματα που θα πρέπει να επιτευχθούν μέσω αυτών. Επίσης, κατά τη διάρκεια χρήσης της εφαρμογής και ανά πάσα στιγμή, ο χρήστης θα μπορεί να δει την καταγραφή με τις συζητήσεις που έχει πραγματοποιήσει μέχρι εκείνη τη στιγμή με τον χαρακτήρα του σεναρίου.

3.2.3 Τεχνικά χαρακτηριστικά

Οι χώροι της VRALL εφαρμογής που σχεδιάζουμε θα είναι πλήρως εικονικοί και φανταστικοί. Θα απεικονίζεται, δηλαδή, ένας φανταστικός κόσμος, ο οποίος όμως επειδή θα περιέχει ρεαλιστικά σενάρια, θα υιοθετεί τις ιδιότητες και τα χαρακτηριστικά του πραγματικού. Αυτό σημαίνει ότι στο εικονικό συνεχές, θα τοποθετείται στο άκρο της πλήρους εικονικής πραγματικότητας. Ακόμη, η εφαρμογή προορίζεται για χρήση από ένα άτομο κάθε φορά. Αυτό σημαίνει ότι όλοι οι χαρακτήρες με τους οποίους ο χρήστης θα αλληλεπιδρά στα σενάρια της εφαρμογής δεν θα είναι κάποιιοι άλλοι πραγματικοί χρήστες σαν αυτόν, αλλά φανταστικοί (bots). Επίσης, το λογισμικό προβολής της εφαρμογής δεν θα βρίσκεται σε κάποιον απομακρυσμένο server κι από εκεί θα μεταδίδεται στη συσκευή του χρήστη, αλλά θα είναι εγκατεστημένο και θα τρέχει τοπικά σε αυτήν. Βέβαια, αυτό αφορά την υποστήριξη του εικονικού περιβάλλοντος. Όλες οι υπόλοιπες λειτουργίες, δηλαδή αυτές που έχουν να κάνουν με το διαλογικό σύστημα, την υπηρεσία αναγνώρισης φωνής, την υπηρεσία σύνθεσης φωνής κ.λπ., δε θα μπορούν να τρέχουν τοπικά στη συσκευή προβολής του χρήστη, αλλά θα βρίσκονται σε απομακρυσμένους servers με τους οποίους το λογισμικό προβολής του εικονικού περιβάλλον θα επικοινωνεί. Με την παραπάνω διάκριση γίνεται μια σκιαγράφηση της αρχιτεκτονικής της εφαρμογής με το frontend μέρος της να είναι το εικονικό περιβάλλον και οι λειτουργίες που ενθυλακώνονται σε αυτό και το backend μέρος να έχουν να κάνουν με το διαλογικό σύστημα και τις λοιπές υπηρεσίες.

Κεφάλαιο 4

Υλοποίηση

Αυτό το κεφάλαιο αφορά την ανάπτυξη και την υλοποίηση της εφαρμογής, όπως αυτή σχεδιάστηκε στο προηγούμενο κεφάλαιο. Αποτελείται από τις ακόλουθες ενότητες:

1. Επιλογή των εργαλείων ανάπτυξης.
2. Εικονικό περιβάλλον της εφαρμογής.
3. Ανθρώπινα avatars.
4. River - VA Backend.
5. Δημιουργία του μοντέλου για το RASA.
6. Παραγωγή των hints.
7. Στήσιμο του client της εφαρμογής και συντονισμός της frontend λειτουργίας.

4.1 Επιλογή των εργαλείων ανάπτυξης

Σε αυτήν την ενότητα θα διακρίνουμε τα εργαλεία που χρησιμοποιήσαμε σε δύο βασικές κατηγορίες. Η πρώτη αφορά το κομμάτι του περιβάλλοντος εικονικής πραγματικότητας και περιλαμβάνει τις εργασίες σχεδίασης τρισδιάστατων χώρων, αντικειμένων και χαρακτήρων, το στήσιμο ενός client που θα επικοινωνεί με τον server του διαλογικού συστήματος και την υλοποίηση του σεναρίου της εφαρμογής. Η δεύτερη αφορά την ανάπτυξη και λειτουργία ενός server διαλογικού συστήματος ο οποίος θα υποστηρίζει τις υπηρεσίες αναγνώρισης ομιλίας του χρήστη (ASR), δημιουργία γραπτής απάντησης προς το χρήστη και προτάσεων (hints) για επόμενες αποκρίσεις αυτού και παραγωγή ομιλίας στα ελληνικά (TTS).

Oculus Quest 2

Ένα από τα βασικότερα κριτήρια για την επιλογή των εργαλείων σχεδίασης και ανάπτυξης εφαρμογών εικονικής πραγματικότητας είναι το σε ποιες συσκευές ή πλατφόρμες θα χρησιμοποιηθεί η εφαρμογή που αναπτύσσεται (target devices). Στην περίπτωση μας η συσκευή αυτή είναι το αυτόνομο headset της Oculus "Meta Quest 2"¹. Το headset και τα δύο χειριστήρια του Quest 2 είναι πλήρως ασύρματα κάτι το οποίο προσφέρει στον χρήστη ελευθερία

¹<https://store.facebook.com/quest>

κινήσεων στο χώρο. Αυτό σε συνδυασμό με τα αναβαθμισμένα τεχνικά χαρακτηριστικά του συγκεκριμένου set ενισχύουν την αίσθηση εμπύθισης στη VR εμπειρία. Ακόμη, η Oculus προσφέρει διάφορα εργαλεία για την υποστήριξη του Χυεστ, όπως είναι το "Oculus Developer Hub"², που βοηθά στην ανάπτυξη και δοκιμή εφαρμογών γι' αυτή τη συσκευή. Τέλος, το γεγονός ότι η συσκευή αυτή είναι πολύ διαδεδομένη και καλύπτει μεγάλο κομμάτι της αγοράς έχει βοηθήσει στο να δημιουργηθεί μεγάλη κι ενεργή κοινότητα προγραμματιστών³ γύρω της.

Unity

Ένα πολύ διαδεδομένο εργαλείο για την ανάπτυξη εφαρμογών εικονικής πραγματικότητας και μάλιστα που θα τρέχουν στο Oculus Quest 2 είναι το Unity [50]. Η συγκεκριμένη εφαρμογή θεωρείται εύκολη στη χρήση και είναι διάσημη για την ανάπτυξη σε αυτήν πολλών δισδιάστατων (2D) και τρισδιάστατων (3D) indie games (independent video games). Αυτό συμβαίνει, γιατί ο unity editor έχει φτιαχτεί με τέτοιο τρόπο, ώστε να δίνει τη δημιουργική ελευθερία στον προγραμματιστή να πειραματίζεται με πειραματικές δομές παιχνιδιού (gameplay), να εξερευνά το μέσο παράγοντας μοναδικές εμπειρίες σε εναλλακτικά πεδία, όπως αυτά της τέχνης και της εκπαίδευσης κι εν γένει να παίρνει ρίσκα εφαρμόζοντας καινοτομίες που συνήθως δεν επιτρέπεται σε AAA (triple-A) παραγωγές παιχνιδιών [51].

Όσον αφορά τη συμβατότητα του Unity με τον εξοπλισμό της Oculus, προσφέρεται ένα SDK⁴ (κιτ ανάπτυξης λογισμικού) ως επέκταση στο περιβάλλον του Unity editor. Χρησιμοποιώντας αυτό το SDK, μπορείς να αξιοποιήσεις βασικά VR στοιχεία του headset, όπως είναι η διαχείριση δεδομένων από το χρήστη (user input), έλεγχο των χειριστηρίων (controllers), όπως και βελτιστοποίηση παραμέτρων οθόνης (rendering), για να δημιουργήσεις εμπυθιστικές εμπειρίες. Επίσης, η Unity πέρα από εκτενή documentation που επικεντρώνονται στην ανάπτυξη VR εφαρμογών και για το Quest 2, διαθέτει και μεγάλη κι ενεργή κοινότητα προγραμματιστών και σχεδιαστών. Τέλος, υπάρχει και το δικό της ψηφιακό κατάστημα (Unity Asset Store)⁵ με μια μεγάλη γκάμα από γραφιστικά αντικείμενα, scripts για διάφορες λειτουργίες, μέχρι και ολόκληρα APIs.

Makehuman - Blender

Για τη δημιουργία των ανθρώπινων avatars που εμφανίζονται στο σενάριο χρησιμοποιήθηκαν τα εργαλεία MakeHuman⁶ και Blender⁷.

Το MakeHuman είναι ένα καινοτόμο λογισμικό ανοιχτού κώδικα που χρησιμοποιείται για τη μοντελοποίηση τρισδιάστατων ανθρωποειδών χαρακτήρων. Για τη διευκόλυνση του δημιουργού διαθέτει ένα αρκετά φιλικό γραφικό περιβάλλον χρήστη (GUI), ενώ η δημιουργία των χαρακτήρων γίνεται μέσω παραμετροποίησης διαφόρων παραμέτρων που επιτρέπουν το συνδυασμό διαφορετικών ανθρώπινων χαρακτηριστικών. Πιο συγκεκριμένα, τα χαρακτη-

²<https://developer.oculus.com/documentation/unity/ts-odh/>

³<https://forums.oculusvr.com/>

⁴<https://developer.oculus.com/documentation/unity/unity-overview/>

⁵<https://assetstore.unity.com/>

⁶<http://www.makehumancommunity.org/>

⁷<https://www.blender.org/>

ριστικά που καθορίζεις και προσαρμόζεις, διακρίνονται σε δύο κατηγορίες, τα macro και τα detail. Τα macro χαρακτηριστικά είναι πιο γενικά κι έχουν να κάνουν με το φύλο, την ηλικία, το ύψος, το βάρος και την εθνικότητα κ.α., ενώ τα detail επιτρέπουν στον χαρακτήρα να διαμορφωθεί με μεγαλύτερη λεπτομέρεια δίνοντας έμφαση σε πράγματα όπως είναι το σχήμα των ματιών ή το μήκων των δακτύλων. Ο στόχος των δημιουργών του συγκεκριμένου project είναι αρκετά φιλόδοξος, μιας και προσπαθούν να προσφέρουν ένα εργαλείο που επιτρέπει τη διαχείριση όλων των πραγμάτων που απαιτούνται για να φτιάξεις έναν ρεαλιστικό ανθρώπινο χαρακτήρα. Για το λόγο αυτό κάποια από τα εργαλεία που περιλαμβάνονται, όπως αυτά που αφορούν τις στάσεις, την κινησιολογία ή τις εκφράσεις προσώπου των χαρακτήρων, είναι σε αρκετά πρώιμα στάδια. Προκειμένου να συμπληρωθούν τα κενά που υπάρχουν σε αυτά τα σημεία, το MakeHuman προσφέρει εργαλεία εξαγωγής (export) των εικονικών χαρακτήρων σε άλλα λογισμικά μοντελοποίησης κι επεξεργασίας τρισδιάστατων σχεδίων, όπως το Blender.

Το Blender είναι κι αυτό ένα ελεύθερο λογισμικό ανοιχτού κώδικα που διατίθεται για κάθε χρήση (εμπορική ή εκπαιδευτική). Μέσω του Blender και με την εγκατάσταση σε αυτό του Makehuman add-on (MPFB), ο δημιουργός μπορεί να επεξεργαστεί σε βάθος και με κάθε λεπτομέρεια τον χαρακτήρα του, να του δώσει κίνηση και να δημιουργήσει μέχρι και animations clips με αυτόν.

4.2 Εικονικό περιβάλλον της εφαρμογής

Unity Editor

Το γραφικό περιβάλλον της εφαρμογής με τον τρισδιάστατο χώρο, τα αντικείμενα, οι περιοχές-επιφάνειες διεπαφής με το χρήστη (User Interface - UI) σχεδιάστηκαν και υλοποιήθηκαν όλα στα πλαίσια χρήσης του Unity. Γι' αυτό το λόγο, κρίνεται χρήσιμο να γίνει μια γρήγορη εισαγωγή σε κάποιες βασικές έννοιες και ορολογίες που χρησιμοποιούνται από το σύστημα αυτό. Κάποια από αυτά είναι τα **Assets**, τα **Scenes**, τα **GameObjects** μαζί με τα **Components** αυτών και τέλος τα **Scripts**.

Asset στο Unity ονομάζεται οτιδήποτε μπορεί να θεωρηθεί πηγή ή πόρος. Αυτό περιλαμβάνει τρισδιάστατα μοντέλα (3D models), υλικά, εικόνες ή επιφάνειες και το πώς αυτά προβάλλονται στην οθόνη (materials και textures), ήχοι (audio), κομμάτια κώδικα (scripts) κ.ά.. Αυτά συνήθως δε δημιουργούνται μέσα στο Unity αλλά σε τρίτες εφαρμογές και κατόπιν εισάγονται σε αυτό για να χρησιμοποιηθούν.

Scene στο Unity είναι το μέρος στο οποίο τοποθετούνται τα Assets του project, τα οποία με τις επιθυμητές τοποθετήσεις και διατάξεις δημιουργούν επίπεδα, χώρους κι ό,τι άλλο χρειάζεται για να υπάρξει το σύμπαν της εφαρμογής. Για να γίνει καλύτερα κατανοητό, οτιδήποτε υπάρχει στην εφαρμογή από ένα τρισδιάστατο αντικείμενο μέχρι κάποια λειτουργία του project, τότε αυτό θα πρέπει να εισαχθεί στα Scenes του project.

Το **GameObject** θεωρείται η σημαντικότερη έννοια στον editor του Unity και είναι η κλάση αντικειμένων που χρησιμοποιείται για να αναπαρασταθεί οποιαδήποτε οντότητα σε μια Scene. Είτε είναι χαρακτήρας, είτε στοιχείο φωτός, είτε στοιχείο κάμερας (οπτική χρήση στην εφαρμογή), είτε οποιοδήποτε ειδικό εφέ ή οντότητα αυτού του αντικειμένου αντιστοιχίζεται με ένα GameObject. Από μόνα τους όμως τα GameObject δεν μπορούν να κάνουν πολλά, αλλά ουσιαστικά αποτελούν το στοιχεία που περιλαμβάνουν (containers) τα Components. Τα **Components** είναι στοιχεία που προσδένονται στα GameObjects και του προσδίδουν ιδιότητες και λειτουργίες. Ανάλογα με τι είδους αντικείμενο θες να είναι ένα GameObject του προσθέτεις και διαφορετικό συνδυασμό από Components.

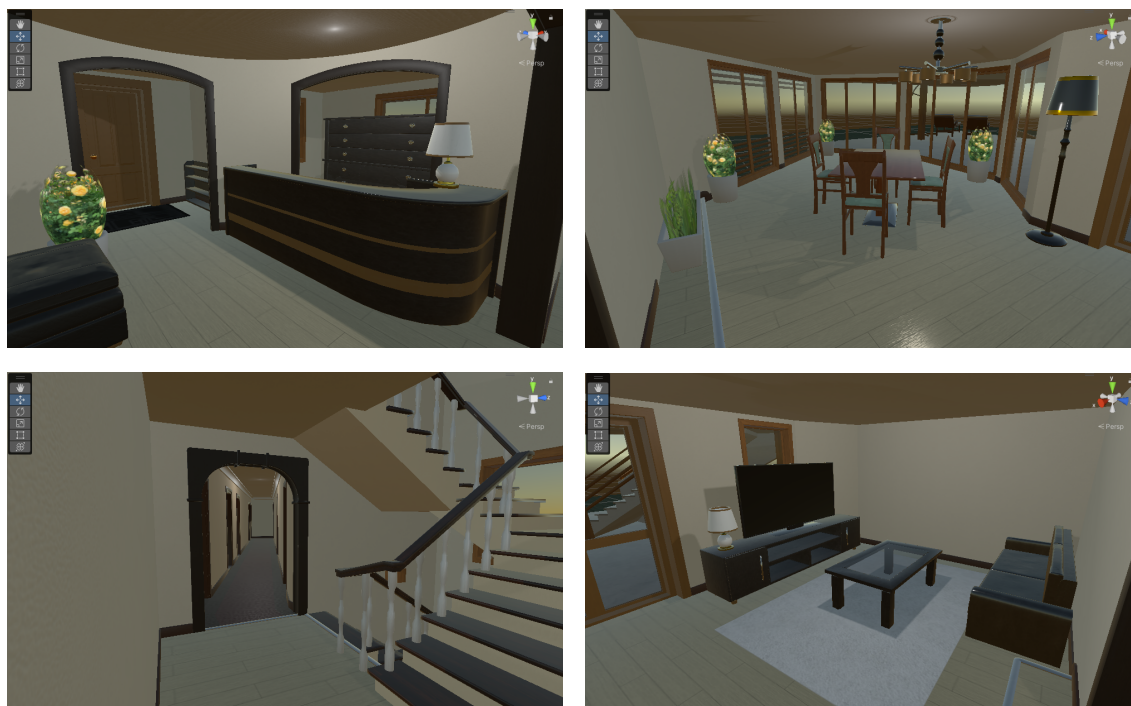
Τέλος, τα **Scripts** μπορούν να αναγνωριστούν ως μια ειδική κατηγορία από Components, τα οποία συνδέονται με GameObjects και χρησιμοποιούνται για να τους προσδώσουν κάποιες συμπεριφορές. Σε κάθε GameObject μπορούν να συνδεθούν πολλά Scripts. Τα Scripts αυτά γράφονται συνήθως στη γλώσσα C# και όλα τα Scripts έχουν πάντα δύο συναρτήσεις (methods), την start() και την update(). Η start() συνάρτηση τρέχει μόνο μια φορά και καλείται όταν το GameObject στο οποίο το Script είναι συνδεδεμένο δημιουργείται (συνήθως όταν φορτώνεται η Scene που το περιέχει). Η update() από την άλλη καλείται μια φορά για κάθε frame, κάθε φορά δηλαδή που ανανεώνεται η εικόνα.

4.2.1 Δημιουργία του χώρου του σεναρίου

Για το χώρο της ρεσεψιόν του ξενοδοχείου χρησιμοποιήθηκε το "Country Hotel" Asset που είναι διαθέσιμο στο Unity Asset Store⁸. Το συγκεκριμένο Asset είναι μια τρισδιάστατη μοντελοποίηση ενός ολόκληρου ξενοδοχείου περιλαμβάνοντας μεταξύ άλλων διάφορους

⁸<https://assetstore.unity.com/packages/3d/environments/urban/country-hotel-136439>

χώρους (αυλή, ρεσεψιόν, δωμάτια, διάδρομος, σκάλες κλπ), διάφορα έπιπλα και όλα αυτά να συνοδεύονται από τα απαραίτητα materials και textures προκειμένου να έχουν την κατάλληλη εμφάνιση. Μετά τη προμήθεια του Asset αυτού, προκειμένου ο χώρος της ρεσεψιόν να είναι έτοιμος, έπρεπε αρχικά να απομονωθεί από όλους τους υπόλοιπους. Αυτό κρίθηκε απαραίτητο γιατί αν έμενε ολόκληρο το μοντέλο του ξενοδοχείου στο Scene, τότε το κόστος της τρεξίματος εφαρμογής από πλευρά μνήμης θα ήταν πολύ μεγάλο ρίχνοντας την ποιότητα χρήσης της. Πράγματι, όταν ολοκληρώθηκε το βήμα αυτό εντοπίστηκαν σημαντικές βελτιώσεις στην εμπειρία χρήσης, αφού εξαφανίστηκαν κάποιες καθυστερήσεις που υπήρχαν. Στη συνέχεια, κάνοντας χρήση των σχεδιαστικών εργαλείων που διαθέτει το Unity, προστέθηκαν κάποιες επιφάνειες (τοιχοί και πατώματα) με σκοπό τη διαμόρφωση ενός πιο βολικού χώρου για το συγκεκριμένο σενάριο. Ακόμη, επειδή ο χώρος ήταν αρχικά πολύ άδειος, έγιναν εκτενείς αλλαγές στη διάταξη των πραγμάτων και των επίπλων. Τέλος, υπήρχαν κάποια προβλήματα με τα προεπιλεγμένα από τον σχεδιαστή του "Country Hotel" στοιχεία φωτός, τα οποία δημιουργούσαν μια περίεργη αίσθηση αντανάκλασης σε όλες τις επιφάνειες του χώρου. Γι' αυτό το λόγο αφαιρέθηκαν όλα κι αντικαταστάθηκαν από διαφορετικές πηγές φωτός που διαμορφώθηκαν κατάλληλα και τοποθετήθηκαν σε όλους τους χώρους. Το αποτέλεσμα από τις παραπάνω εργασίες φαίνεται στις εικόνες που ακολουθούν.



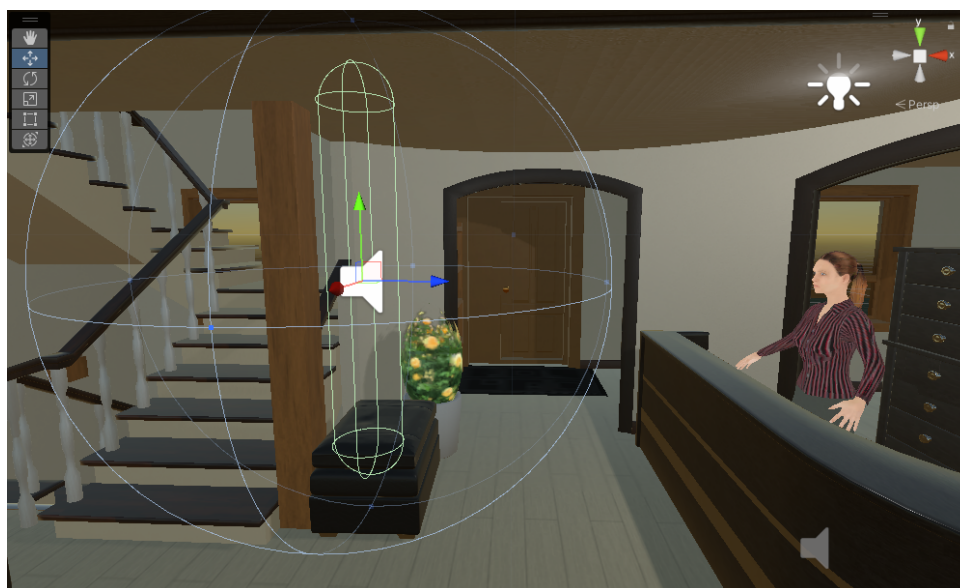
Σχήμα 4.1: Οι χώροι της reception του ξενοδοχείου

4.2.2 Κίνηση του χρήστη στο χώρο

Αφού ετοιμάστηκε ο χώρος στον οποίο διαδραματίζεται το πρώτο σενάριο της εφαρμογής, έπρεπε να υλοποιηθούν οι λειτουργίες που έχουν να κάνουν με τον χρήστη. Αναλυτικά, οι λειτουργίες αυτές αφορούν κυρίως την οπτική του χρήστη και την κίνηση του στο χώρο.

Το Oculus Integration package (SDK)⁹, που αναφέρθηκε σε προηγούμενες ενότητες, εκτός από το εργαλείο για το lipsync περιέχει κι άλλα δύο χρήσιμα για την εφαρμογή στοιχεία σε μορφή έτοιμων GameObject (prefabs). Πρόκειται για τα OVRPlayerController και OVRCameraRig¹⁰. Με τη χρήση αυτών ουσιαστικά αυτοματοποιείται ο έλεγχος του VR εξοπλισμού (headset και χειριστήρια) από την εφαρμογή. Πιο συγκεκριμένα, το GameObject OVRPlayerController είναι αυτό που με τη χρήση του αριστερού μοχλού κινεί το χρήστη στο εικονικό περιβάλλον της εφαρμογής και όπως επίσης αναγνωρίζει και όποιο άλλο κουμπι του χειριστηρίου πατηθεί. Από την άλλη το OVRCameraRig είναι αυτό που με την κατάλληλη παραμετροποίηση μέσω του headset παρακολουθεί τις κινήσεις που κάνει ο χρήστης με το κεφάλι του και κατ' ουσίαν τις μεταφράζει σε κινήσεις στον εικονικό κόσμο.

Αφού ενσωματώθηκαν αυτά τα δύο εργαλεία στο Scene του σεναρίου, τότε έπρεπε να γίνουν κάποιες προσθήκες και σε αυτά αλλά και στα GameObjects που συνθέτουν τον εικονικό χώρο, υπό τη μορφή Components. Οι λειτουργίες που υλοποιήθηκαν ήταν η οριοθέτηση του χώρου στον οποίο ο χρήστης μπορεί να κινηθεί, προκειμένου να μην μπορεί να βγει από αυτόν. Να μην μπορεί δηλαδή να περάσει μέσα από τοίχους κι άλλες επιφάνειες, να μην πέφτει στο κενό όταν φτάνει στο τέλος του επιπέδου κ.λπ.. Αυτό έγινε με τη χρήση των Colliders. Collider είναι ο τρόπος που δίνεται η δυνατότητα από το Unity στον προγραμματιστή να χειρίζεται τις συγκρούσεις αντικειμένων με όρους φυσικής. Ουσιαστικά, πρόκειται για αόρατα τρισδιάστατα γεωμετρικά σχήματα τα οποία προσδένονται στα GameObjects του Scene, χωρίς όμως να χρειάζεται να υιοθετούν το σχήμα και τη μορφή τους (συνήθως έχουν τη μορφή σφαίρας, κύβου ή κάψουλας). Προσθέτοντας, λοιπόν, capsule collider στο GameObject OVRPlayerController, σε όλες τις επιφάνειες κι αντικείμενα του χώρου και προσδίδοντας στους Colliders των αντικειμένων την ιδιότητα να συμπεριφέρονται με βάση τους νόμους της φυσικής οριοθετήθηκε καλά η περιοχή που ο χρήστης κινείται.



Σχήμα 4.2: Ο collider του OVRPlayerController

⁹<https://developer.oculus.com/documentation/unity/unity-utilities-overview/#vr-folder>

¹⁰<https://developer.oculus.com/documentation/unity/unity-add-camera-rig/>

4.2.3 Διεπαφές με το χρήστη

Το επόμενο βήμα στην υλοποίηση του εικονικού περιβάλλοντος και των λειτουργιών του ήταν η δημιουργία των στοιχείων αυτών που θα έδιναν στο χρήστη τη δυνατότητα να αλληλεπιδρά με την εφαρμογή πέρα από το απλά να περιπλανιέται σε αυτήν. Έπρεπε δηλαδή να στηθούν τα διάφορα στοιχεία UI (δείκτες, κουμπιά, παράθυρα, εικόνες, κείμενα κλπ) τα οποία χρησιμοποιεί ο χρήστης καθόλη τη διάρκεια της εμπειρίας του.

Το Unity προσφέρει στον προγραμματιστή τρεις επιλογές για το πώς θα δημιουργήσει τέτοιες διεπαφές χρήστη (UI)¹¹. Αυτές είναι οι εξής, το **UI Toolkit**, το **Unity UI**¹² (**uGUI**) και το **IMGUI**. Το UI Toolkit είναι το εργαλείο που ενημερώνεται συχνά με νέες εκδόσεις και προορίζεται να γίνει το προτεινόμενο σύστημα ανάπτυξης ΥΙ εφαρμογών. Ακόμη, είναι εμπνευσμένο από καθιερωμένες τεχνολογίες στον τομέα της ανάπτυξης web εφαρμογών με τον πυρήνα του να είναι ένα σύστημα διατηρούμενης λειτουργίας (retained-mode API). Από την άλλη, το Unity UI είναι ένα σύστημα που έχει στη βάση του για τη δημιουργία του UI τα GameObjects και χρησιμοποιεί Components για τις ιδιότητές του (θέση, διάταξη, στυλ) και για τις λειτουργίες του. Βέβαια, με δεδομένη την προώθηση του UI Toolkit από την ίδια τη Unity, το uGUI για να χρησιμοποιηθεί απαιτεί τη χρήση κάποιων εργαλείων που έχουν καταργηθεί. Τέλος, είναι και το IMGUI που δεν έχει σχεδιαστεί τόσο για χρήση από πραγματικούς χρήστες σε εφαρμογές, αλλά περισσότερο για debugging χρήσεις και για δημιουργία επεκτάσεων στο ίδιο το Unity. Με αυτά τα δεδομένα αποφασίστηκε να χρησιμοποιηθεί το Unity UI, γιατί η ροή εργασίας σε αυτό είναι πιο κοντά στον τρόπο που έχουν γίνει μέχρι αυτό το σημείο τα πράγματα, αλλά και για κάποιους πιο τεχνικούς λόγους, όπως είναι το ότι το UI Toolkit δεν έχει τόσο καλά ακόμη ανεπτυγμένα τα εργαλεία του για δημιουργία διεπαφών σε περιβάλλοντα VR εφαρμογών.

Όπως ειπώθηκε και προηγουμένως, το βασικό στοιχείο του uGUI είναι τα GameObjects. Αντίστοιχα, το βασικό GameObject από τα στοιχεία του UI που δημιουργούνται είναι το Canvas, το οποίο είναι η περιοχή στην οποία όλα τα υπόλοιπα UI στοιχεία περιέχονται και είναι παιδιά του. Κάποια από αυτά τα στοιχεία είναι Buttons (περιοχές του Canvas που αν αλληλεπιδράσεις με αυτές πυροδοτούν ένα επόμενο γεγονός), Images και Texts (περιοχές του Canvas που περιέχουν κείμενο). Επιπλέον, το Canvas έχει δύο βασικούς τρόπους που μπορεί να εμφανίζεται στο Scene της εφαρμογής. Ο ένας είναι ως μέρος της οθόνης (screen space) και ο δεύτερος είναι ως μέρος του εικονικού κόσμου (world space). Ο τρόπος που εν τέλει χρησιμοποιήθηκε για όλα τα Canvas του σεναρίου ήταν ο δεύτερος.

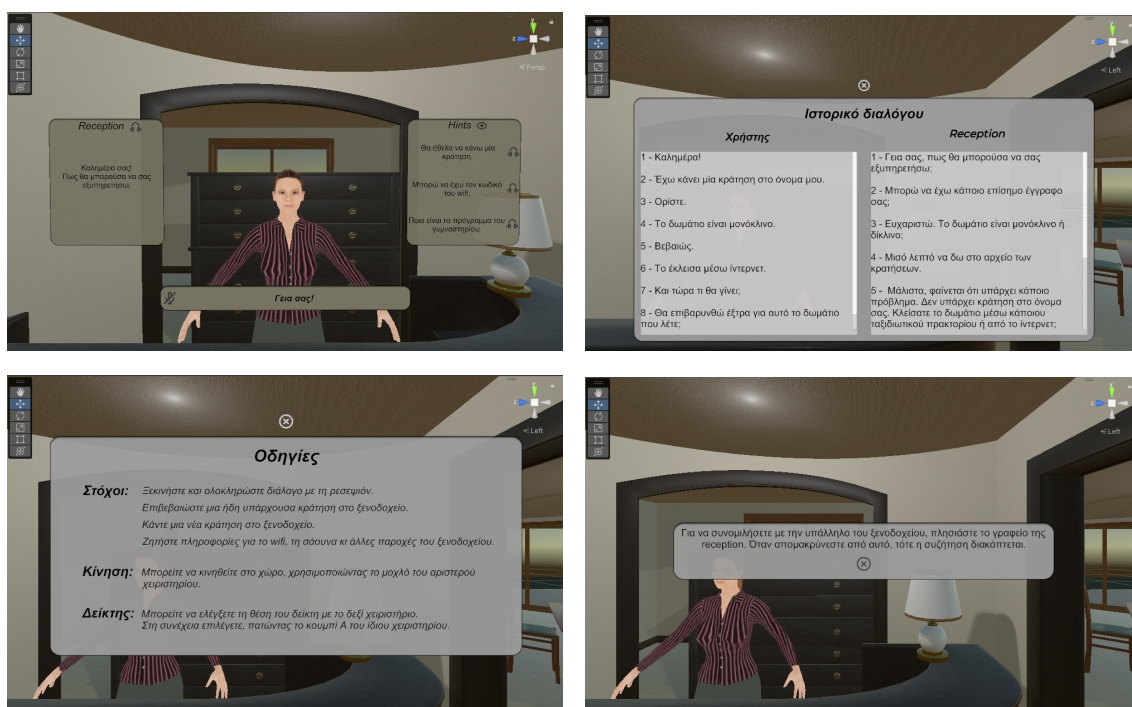
Τα στοιχεία διεπαφής που δημιουργήθηκαν και τα οποία αντιστοιχούν και σε δύο διαφορετικά Canvas είναι τα εξής. Το πρώτο είναι το **menu_Canvas** και το δεύτερο είναι το **reception_Canvas**. Ο τρόπος όμως που υλοποιήθηκαν αυτά τα GameObjects, όπως είπαμε προηγουμένως είναι να τοποθετηθούν ως worldspace στο χώρο. Το πρώτο φαίνεται λες και είναι screenspace, γιατί κατέχει σταθερή θέση στην οθόνη του χρήστη. Αυτό έγινε αρχικοποιώντας σε μια αρχική απόσταση και θέση από την οπτική της κάμερας και στη συνέχεια ορίζοντας το ως παιδί του GameObject OVRCameraRig. Με αυτόν τον τρόπο, όποια μεταβολή στη θέση και στην περιστροφή συμβαίνει στην οπτική της κάμερας, συμβαίνει και σε

¹¹<https://docs.unity3d.com/Manual/UI-system-compare.html>

¹²<https://docs.unity3d.com/Packages/com.unity.ugui@1.0/manual/index.html>

αυτό το Canvas. Το δεύτερο βρίσκεται σταθερά πάνω από τον γκισέ της ρεσεψιόν με τρόπο που περιβάλλει την υπάλληλο κι εμφανίζεται όταν ο χρήστης την πλησιάσει.

Όσον αφορά στο περιεχόμενό τους, το menu_Canvas αποτελείται από τρία διαφορετικά παράθυρα και τρία εικονικά κουμπιά με τα οποία ο χρήστης μπορεί να ανοιγοκλείσει τα παράθυρα που θέλει. Το πρώτο είναι το παράθυρο των οδηγιών το οποίο εμφανίζεται αυτόματα με το που ο χρήστης μεταφέρεται στο περιβάλλον του σεναρίου. Σε αυτό βρίσκονται οι στόχοι του σεναρίου που πρέπει να επιτευχθούν και οδηγίες για το πώς ο χρήστης κινείται στο χώρο και χρησιμοποιεί τον δείκτη του χειριστηρίου. Όταν το παράθυρο αυτό κλείσει για πρώτη φορά, τότε εμφανίζεται το δεύτερο παράθυρο με το οποίο η εφαρμογή καλεί τον χρήστη να κινηθεί προς τη ρεσεψιόν για να ξεκινήσει διάλογο με την υπάλληλο. Αυτό το παράθυρο εμφανίζεται μόνο μια φορά. Τέλος, στο τρίτο παράθυρο φαίνεται το ιστορικό των συζητήσεων που ο χρήστης έχει πραγματοποιήσει με την υπάλληλο στα πλαίσια του συγκεκριμένου γύρου. Το recerption_Canvas περιέχει κι αυτό τρία παράθυρα, τα οποία όμως σε αντίθεση με αυτά του menu_Canvas εμφανίζονται όλα ταυτόχρονα. Το ένα είναι χαμηλά στο κέντρο και περιέχει στα αριστερά ένα σύμβολο που ενημερώνει τον χρήστη τότε το μικρόφωνο του είναι ενεργό, προκειμένου να μιλήσει και στη μέση του παραθύρου τη φράση που αναγνωρίστηκε από την ομιλία του χρήστη. Το δεύτερο βρίσκεται στα αριστερά και περιέχει τη γραπτή απάντηση της υπαλλήλου της ρεσεψιόν την οποία ο χρήστης μπορεί να ακούσει όσες φορές θέλει, πατώντας το σχετικό κουμπί που βρίσκεται στη δεξιά γωνία του παραθύρου. Το τρίτο παράθυρο βρίσκεται στα δεξιά και περιέχει τα hints για το τι θα μπορούσε να πει ο χρήστης, ώστε να συνεχίσει τη συζήτηση. Δίπλα σε κάθε γραπτό hint υπάρχει ένα κουμπί που αν πατηθεί ακούγεται το πώς προφέρεται. Το παράθυρο αυτό των hints είναι ανακλινόμενο, για να δίνεται η δυνατότητα στον χρήστη να το εμφανίζει ή να το κλείνει ανάλογα με το αν θέλει να βλέπει τα hints ή όχι. Από κάτω φαίνονται τα παράθυρα που μόλις περιγράφηκαν.



Σχήμα 4.3: Τα παράθυρα διεπαφής με τον χρήστη.

4.3 Ανθρώπινα avatars

4.3.1 Σχεδιασμός των κινήσεων στο πρόσωπο

Το MakeHuman στην καρτέλα Pose/Animate περιέχει κάποιες προεπιλεγμένες εκφράσεις προσώπου που έχουν να κάνουν με συναισθήματα, όπως είναι η χαρά, η λύπη, ο ενθουσιασμός, ο φόβος, η έκπληξη, η βαρεμάρα, η ευχαρίστηση, η αηδία, κ.ά.^{4.4}.



Σχήμα 4.4: Μερικές εκφράσεις προσώπου από το MakeHuman.

Οι εκφράσεις αυτές, αν και αρκετά ρεαλιστικές δεν περιλαμβάνουν παραμορφώσεις του προσώπου που αντιστοιχούν στις κινήσεις που κάνουμε όταν προφέρουμε διάφορα φωνήματα, σύμφωνα και φωνήεντα. Μέσω της κοινότητας της εφαρμογής βρήκαμε κάποιες προσθήκες σε μορφή targets που αποτελούσαν ένα σύνολο από 39 εκφράσεις προσώπου, συμπεριλαμβανομένων και των εκφράσεων που αντιστοιχούν στα εξής φωνήματα: AA, AE, AH, AO, AW, AY, B, CH, D, DH, EH, ER, EY, F, G, H, IH, IY, JH, K, L, M, N, NG, OW, OY, P, R, S, SH, T, TH, UH, UW, V, W, Y, Z, ZH¹³. Προφανώς, τα φωνήματα αυτά αντιστοιχούν στις γλώσσες που χρησιμοποιούν το λατινικό αλφάβητο, αλλά όπως θα δούμε και στη συνέχεια μπορούμε να απομονώσουμε από αυτά όσα μας είναι απαραίτητα για την αντιστοίχιση με την ελληνική γλώσσα.

Από τις εικόνες που βλέπουμε στο 4.5 διακρίνονται κάποια προβλήματα. Αρχικά, παρατηρείται ότι τα άνω και κάτω χείλη δεν κλείνουν αρκετά με αποτέλεσμα στην κατάσταση ηρεμίας, όπου δεν προφέρεται κανένα φώνημα, να φαίνονται τα δόντια. Ακόμη, αν και οι κινήσεις που αντιστοιχούν στο κάθε φώνημα είναι αρκετά ρεαλιστικές, παρατηρείται μια έλλειψη συμμετρίας στο πώς υλοποιείται η κίνηση. Πιο συγκεκριμένα, η κάτω οδοντοστοιχία φαίνεται πάντα και μάλιστα περισσότερο από ότι θα έπρεπε, ενώ η πάνω σχεδόν καθόλου. Το πρώτο πρόβλημα λύθηκε παραμορφώνοντας ελαφρώς το πλέγμα Body Mesh στις περιοχές που αντιστοιχούν τα χείλη, ώστε σε κατάσταση ηρεμίας να είναι τελείως κλειστά. Το δεύτερο πρόβλημα λύθηκε απομονώνοντας το πλέγμα που αφορά τα δόντια TeethBase Mesh και με-

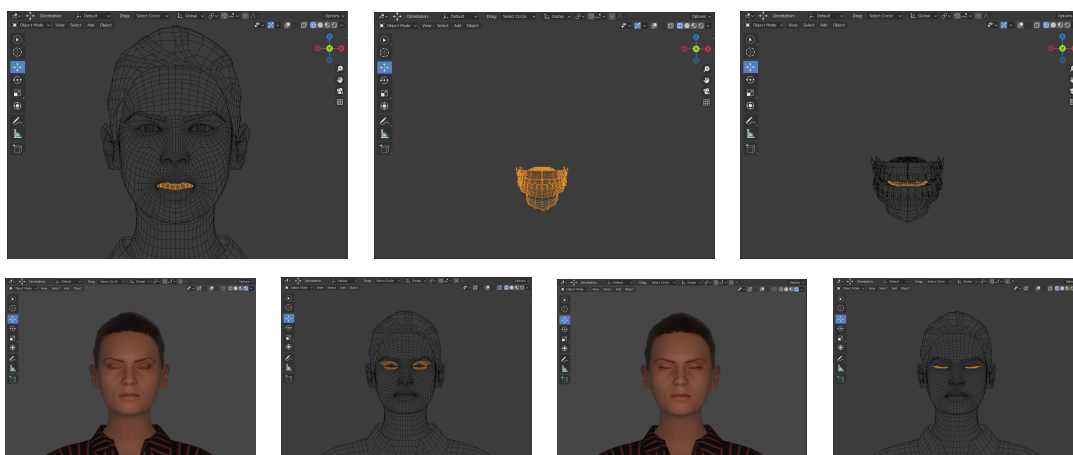
¹³<https://github.com/Hunanbean-Collective/MakeHuman-Additions>



Σχήμα 4.5: Μερικές εκφράσεις προσώπου που αντιστοιχούν σε φωνήματα.

τακινώντας το προς τα κάτω. Κάνοντας αυτό, παρατηρήθηκε ότι η γλώσσα ήταν περισσότερο εκτεθειμένη από πριν, αλλά και περισσότερο από ότι φυσιολογικά θα έπρεπε. Γι' αυτό το λόγο, με παρόμοιο τρόπο απομονώνοντας το πλέγμα που αφορά τη γλώσσα Tongue Mesh μετακίνησα τη γλώσσα κι αυτή προς τα κάτω.

Αντίστοιχα, για την κίνηση των βλεφάρων βρήκαμε κάποια σχετικά targets¹⁴ που αφορούν το κλείσιμο του πάνω και κάτω μέρους των βλεφάρων αντίστοιχα. Το πρόβλημα σε αυτήν την περίπτωση ήταν πως αν και η συγκεκριμένη προσθήκη έλυσε το ζήτημα με την κίνηση των βλεφάρων, οι βλεφαρίδες συνέχιζαν να μένουν ακίνητες ανεξάρτητα από τα βλέφαρα. Αυτό συνέβαινε γιατί τα πλέγματα (meshes) των βλεφάρων και των βλεφαρίδων ήταν διαφορετικά, μιας και με βάση το πως εξήγαγε το MakeHuman τα skins των avatar. Πιο συγκεκριμένα, η παραμόρφωση που προκαλούσε την κίνηση των βλεφάρων γινόταν στο Body Mesh, ενώ η παραμόρφωση που αφορούσε τις βλεφαρίδες γινόταν στο Eyelashes Mesh. Με δεδομένο ότι στην κοινότητα δεν υπήρχε κάποια σχετική προσθήκη, η συγκεκριμένη παραμόρφωση σχεδιάστηκε με το χέρι με τέτοιο τρόπο ώστε να είναι συγχρονισμένη με την κίνηση των βλεφάρων.



Σχήμα 4.6: Προσαρμογή των πλεγμάτων στο πρόσωπο μέσω Blender.

¹⁴<http://www.makehumancommunity.org/forum/viewtopic.php?f=3&t=14160&p=38405&hilit=close+lids#p38382>

Lip sync

Αφού ο χαρακτήρας είναι έτοιμος, πρέπει να βρεθεί ένας τρόπος να γίνει και το lip sync (να συγχρονιστεί δηλαδή ο ήχος που αναπαράγεται με τις κινήσεις που αντιστοιχούν στα κατάλληλα φωνήματα). Για να γίνει αυτό, πρέπει να εξασφαλιστούν τα εξής πράγματα. Αρχικά, θα πρέπει να γίνεται ανάλυση του ήχου που αναπαράγεται (στην περίπτωση μας φωνή της υπαλλήλου στη reception), ώστε να αναγνωρίζονται τα φωνήματα που υπάρχουν. Αυτή η διαδικασία είναι σημαντικό να γίνεται σε πραγματικό χρόνο, γιατί το lip sync του avatar θα συμβαίνει ενώ ταυτόχρονα ο χρήστης της εφαρμογής συνομιλεί με αυτό. Αυτό σημαίνει πως όταν λαμβάνεται η απάντηση - σε γραπτή και προφορική μορφή - από το διαλογικό σύστημα (backend), αυτή θα αναπαράγεται άμεσα. Εφόσον γίνεται αυτό, θα πρέπει να υπάρχει ένας τρόπος να προκαλούνται οι αντίστοιχες κινήσεις στο πρόσωπο του avatar, να έχει γίνει δηλαδή μια μορφή αντιστοίχισης των φωνημάτων που αναγνωρίζονται με τα φωνήματα που έχουν ήδη συνδεθεί, όπως δείξαμε στην προηγούμενη ενότητα με τις κινήσεις του προσώπου. Τέλος, εκτός από τα φωνήματα ένα ακόμη δεδομένο που θα πρέπει να εξάγεται από την ομιλία του avatar είναι η ένταση με την οποία εκφέρεται η κάθε συλλαβή και το κάθε φώνημα, ώστε να επηρεάζεται αντίστοιχα και η έκταση της κίνησης που θα γίνει στο στα χείλη, τα δόντια και τη γλώσσα του avatar.



Σχήμα 4.7: Κινήσεις φωνημάτων μέσω των βλενδοσηπες στο Unity.

Ένα εργαλείο που έχει υλοποιήσει αυτά που συζητάμε από πάνω είναι το "Oculus Lip-sync"¹⁵, το οποίο μάλιστα υπάρχει και σε μορφή plug-in για το Unity ως μέρος του Audio SDK του Oculus Integration Package που αναφέρθηκε σε προηγούμενη ενότητα. Το Oculus Lipsync αναλύει τον ήχο που στριμάρεται (audio input stream) από το μικρόφωνο ή από ένα αρχείο ήχου και προβλέπει ένα σύνολο τιμών που ονομάζει "visemes". Τα "visemes" είναι εκφράσεις των χειλιών και του προσώπου που αντιστοιχίζονται σε συγκεκριμένους ήχους ομιλίας. Ακόμη, το σύστημα αυτό αναγνωρίζει κι αντιστοιχίζει σε "visemes" 15 ξεχωριστά φωνήματα. Αυτά είναι τα εξής: sil, PP, FF, TH, DD, kk, CH, SS, nn, RR, aa, E, ih, oh, και ou. Το φώνημα sil αντιστοιχεί στην ουδέτερη έκφραση (σιωπή). Το σύνολο των φωνημάτων αυτών μπορεί να φαίνεται φτωχότερο από το σύνολο των φωνημάτων που αντιστοιχήθηκαν προηγουμένως στον χαρακτήρα μας, αλλά δεν είναι. Αυτό συμβαίνει γιατί επιλέχθηκαν από την Oculus με δύο βασικά κριτήρια. Το πρώτο ήταν να καλύπτουν το μέγιστο εύρος κίνησης των χειλιών και το δεύτερο ήταν να είναι ανεξάρτητα από την ομιλούμενη γλώσσα. Μάλιστα τα κριτήρια αυτά καθορίστηκαν με βάση το Viseme MPEG-4 Standard [52]. Τέλος, αν και το plug-in αυτό συμπεριλαμβάνει το δικό του avatar, το οποίο έχει σχεδιαστεί με τέτοιο τρόπο, ώστε τα "visemes" που προβλέπονται να αντιστοιχίζονται τέλεια με τη γεωμετρία του προσώπου-στόχου (geometry morph target) ο τρόπος που σχεδιάστηκε ο δικός μας χαρα-

¹⁵<https://developer.oculus.com/documentation/unity/audio-ovrlipsync-unity/>

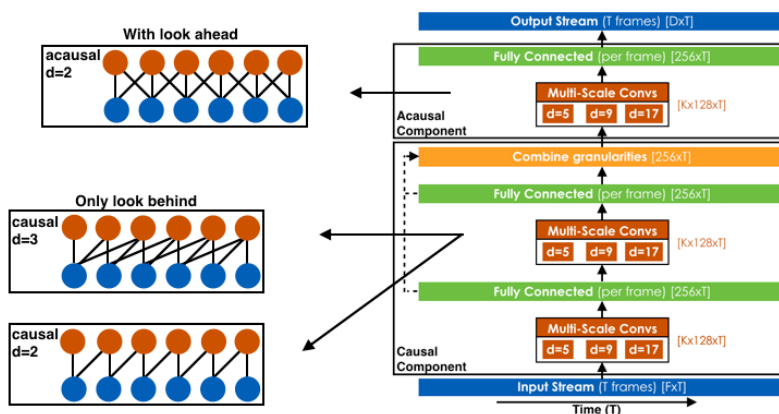
κτήρας και το μεγάλο πλήθος των blendshapes που διαθέτει, βοήθησε ώστε να φτιαχτεί μια ικανοποιητική αντιστοιχία (mapping) "visemes" και avatar.

Μοντέλο της Oculus για την πρόβλεψη των visemes.

Το αρχικό μοντέλο για το Oculus Lipsync χρησιμοποιούσε ένα μικρό και ρηχό νευρωνικό δίκτυο, για να δημιουργήσει μια αντιστοιχισή μεταξύ μικρών τμημάτων της εισόδου ήχου με ανθρώπινη ομιλία (speech segments) και φωνημάτων (visemes). Ενώ αυτό το μοντέλο λειτουργούσε αρκετά καλά για την αγγλική γλώσσα, διαπιστώθηκε ότι αφενός δεν λειτουργούσε καλά για άλλες γλώσσες και αφετέρου δεν ήταν αποτελεσματικό στο να διαχειρίζεται το θόρυβο του περιβάλλοντος. Με αυτό το δεδομένο αναπτύχθηκαν νεότερα μοντέλα μηχανικής μάθησης, και συγκεκριμένα Temporal Convolutional Networks (TCN), τα οποία έχουν δείξει ότι επιτυγχάνουν σημαντικά υψηλότερες επιδόσεις και ανθεκτικότητα και σε άλλους τομείς (π.χ. όραση και γλώσσα). Τα μοντέλα αυτά μπόρεσαν να επιτύχουν πάνω από 30% καλύτερα αποτελέσματα στην οπτικοποίηση σε visemes της αγγλικής γλώσσας και να ξεπεράσουν σημαντικά το προηγούμενο μοντέλο στις περιπτώσεις ομιλιών με διαφορετική από τη συνηθισμένη προφορά, αλλά και σε περιπτώσεις που υπήρχε πολύς θόρυβος στο παρασκήνιο. Στην ερευνητική κοινότητα της "Αναγνώρισης κι Επεξεργασίας Φωνής", αυτά τα μοντέλα αναφέρονται ως ακουστικά μοντέλα και χρησιμοποιούνται συχνά στις αρχιτεκτονικές συστημάτων αναγνώρισης ομιλίας.

Ένα διάγραμμα που απεικονίζει τη γενική αρχιτεκτονική του TCN παρουσιάζεται παρακάτω 4.8. Αυτό το μοντέλο δέχεται ως είσοδο μια ροή από δεδομένα με low-level χαρακτηριστικά του ήχου από το παρελθόν, καθώς και σε ορισμένες περιπτώσεις δεδομένα για από το "μέλλον" για να προβλέψει ένα σύνολο φωνημάτων. Οι ακριβείς παράμετροι (π.χ. αριθμός στρωμάτων και κρυφών καταστάσεων) αυτής της αρχιτεκτονικής παραμετροποιούνται για τη βελτιστοποίηση της υπολογιστικής αποδοτικότητας και επίδοσης, αλλά η γενική διάταξη είναι όπως παρουσιάζεται στην εικόνα.

Παρά το γεγονός ότι το μοντέλο αυτό είναι πολύ πιο πολύπλοκο από την αρχική προσέγγιση lipsync η ομάδα της Oculus κατάφερε να το χρησιμοποιήσει πολύ αποτελεσματικά επιστρατεύοντας μια τεχνική προσωρινής αποθήκευσης (caching technique που μοιάζει με τον Fast WaveNet Generation Algorithm.



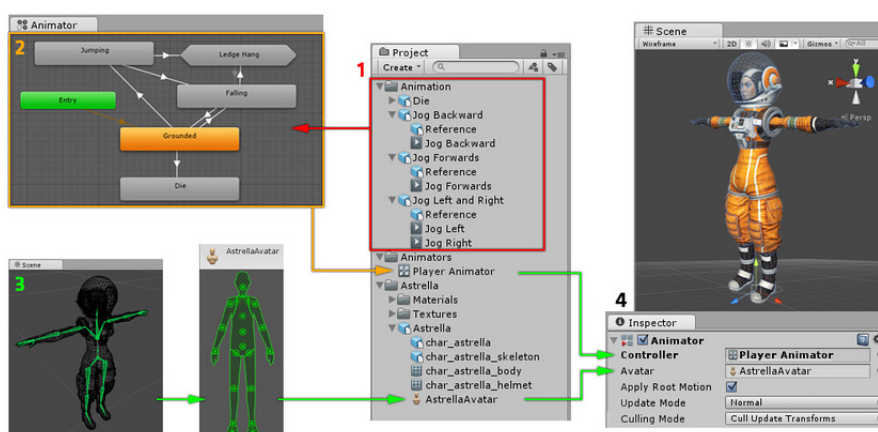
Σχήμα 4.8: Η αρχιτεκτονική του TCN μοντέλου για την αναγνώριση φωνημάτων.

4.3.2 Τα animations στο Unity

Το Unity διαθέτει το δικό του σύστημα για τη δημιουργία και το στήσιμο κίνησης, το οποίο συνήθως αποκαλείται **Mecanim** 4.9. Το σύστημα αυτό έχει τις δικές του έννοιες και ορολογίες, συνεπώς για να γίνει δυνατό να περιγραφεί η υλοποίηση της κίνησης στην περίπτωση της εφαρμογής μας, θα πρέπει να κάνουμε μια σύντομη εισαγωγή για τη ροή εργασίας σε αυτό.

Το Mecanim βασίζεται στην έννοια των **animation clips**, τα οποία περιέχουν πληροφορίες σχετικά με τον τρόπο με τον οποίο ορισμένα αντικείμενα πρέπει να αλλάζουν τη θέση τους, την περιστροφή τους ή άλλες ιδιότητες με την πάροδο του χρόνου. Κάθε animation clip μπορεί να θεωρηθεί ως μια ενιαία γραμμική καταγραφή κίνησης. Τα clips αυτά είτε σχεδιάζονται και δημιουργούνται μέσα από το Unity, είτε γίνεται η εισαγωγή τους από εξωτερικές πηγές, όπως είναι το Asset Store.

Στη συνέχεια, τα animation clips οργανώνονται σε ένα εργαλείο που μοιάζει με διάγραμμα ροής και ονομάζεται **animator controller**. Ο animator controller ελέγχει ποιο clip πρέπει να αναπαράγεται κάθε στιγμή και πότε πρέπει να αλλάζουν ή να συνδυάζονται τα animations. Ο animator controller παρέχει τη δυνατότητα να εφαρμόζεις τις μεθόδους layering (διαφορετικά επίπεδα κίνησης που ισχύουν ταυτόχρονα, είτε γίνεται με παραλληλοποίηση, είτε με επαύξηση της μίας στην άλλη) και masking (εφαρμογή κινήσεων μόνο σε συγκεκριμένα μέρη του χαρακτήρα). Το σύστημα αυτό της Unity έχει επίσης πολλές ειδικές δυνατότητες για το χειρισμό ανθρωποειδών χαρακτήρων. Μία από αυτές είναι η εφαρμογή μιας κίνησης που φτιάχτηκε για έναν συγκεκριμένο χαρακτήρα σε έναν άλλο. Κάποια άλλη είναι η προσαρμογή του τρόπου με τον οποίο έχουν οριστεί οι μύες και ο σκελετός του κάθε χαρακτήρα. Αυτά τα ειδικά χαρακτηριστικά υποστηρίζονται από το εργαλείο του Unity που λέγεται **avatar**, όπου οι ανθρωποειδείς χαρακτήρες και όλα τα επιμέρους στοιχεία τους αντιστοιχίζονται σε μια κοινή εσωτερική μορφή (mapping of body, head, left hand, right hand).



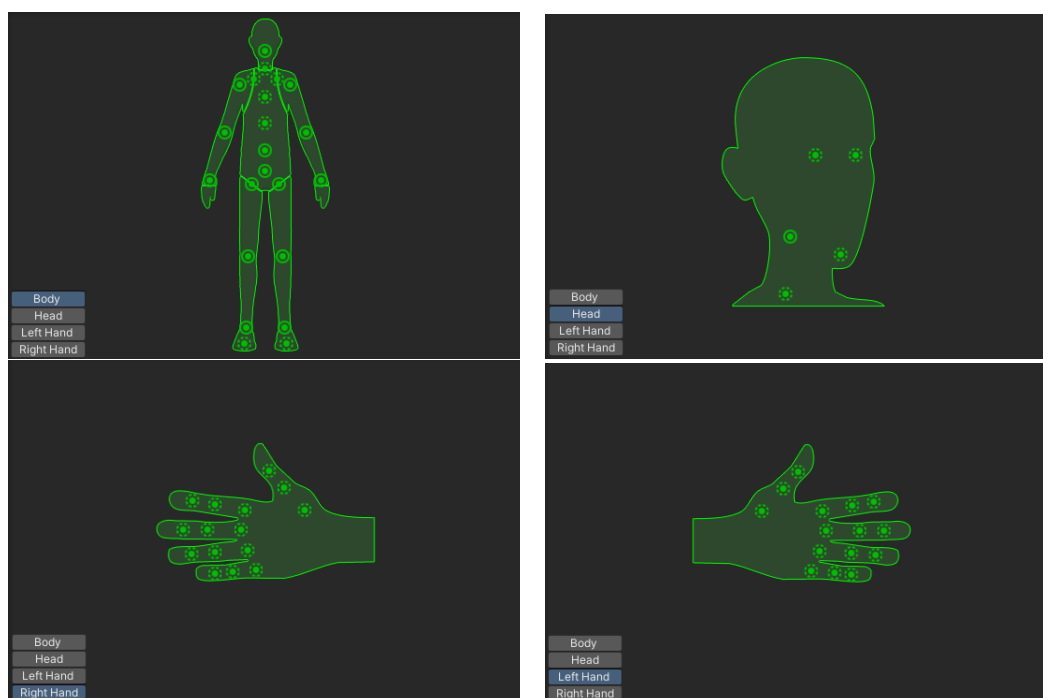
Σχήμα 4.9: Η ροή λειτουργίας του συστήματος Mecanim.

Κάθε ένα από αυτά τα τρία κομμάτια - τα **animation clips**, ο **animator controller** και το **avatar**, συγκεντρώνονται σε ένα GameObject μέσω του component Animator. Αυτό το component έχει έναν δείκτη σε έναν animator controller και (εάν απαιτείται) στο avatar για

αυτό τον χαρακτήρα μοντέλο. Ο animator controller, με τη σειρά του, περιέχει τους δείκτες στα animation clips που χρησιμοποιεί.

4.3.3 Υλοποίηση της κίνησης

Για την εφαρμογή κίνησης στον χαρακτήρα της εφαρμογής έπρεπε να ξεκινήσουμε με τη δημιουργία του avatar 4.10, ώστε να κάνουμε και το αντίστοιχο mapping με όλα τα μέρη του σώματος του. Στη δική μας περίπτωση το avatar είχε κι έπρεπε να αντιστοιχίσουμε τα εξής μέρη (bones). Στον κορμό υπάρχουν οι γοφοί, η σπονδυλική στήλη, ο θώρακας και ο άνω θώρακας. Σε κάθε χέρι (δεξί κι αριστερό) υπάρχουν ο ώμος, το άνω μπράτσο, το κάτω μπράτσο (στην περιοχή του αγκώνα) και ο καρπός και σε κάθε πόδι (δεξί κι αριστερό) υπάρχουν το άνω πόδι, το κάτω πόδι (στην περιοχή του γονάτου), ο αστράγαλος και τα δάκτυλα του ποδιού. Στο κεφάλι υπάρχουν ο λαιμός, η βάση του κεφαλιού, το δεξί κι αριστερό μάτι και το σαγόι. Τέλος, υπάρχουν και τα λεπτομερή σημεία του κάθε χεριού στην περιοχή της παλάμης τα οποία είναι 15 για κάθε χέρι (3 σε κάθε δάχτυλο). Οι αντιστοιχίσεις που κάναμε φαίνονται στις εικόνες από κάτω.



Σχήμα 4.10: Τα μέρη του humanoid avatar στο Unity.

Βασική κινησιολογία

Αφότου ολοκληρώθηκε αυτό, εισάγαμε από το Unity Asset Store κάποια δείγματα από μια συλλογή βασικών κινήσεων για ανθρωποειδή¹⁶. Αυτή η συλλογή αν και είναι μικρή και τα animation clips που περιέχει είναι αρκετά απλοποιημένα ήταν αρκετή στο να δοθεί στο χαρακτήρα της ρεσεψιόν μια στοιχειώδης κίνηση. Συγκεκριμένα, από αυτήν χρησιμοποιήσαμε δύο animation clips. Το πρώτο αφορά την κατάσταση στην οποία ο χαρακτήρας

¹⁶<https://assetstore.unity.com/packages/3d/animations/basic-motions-free-154271>

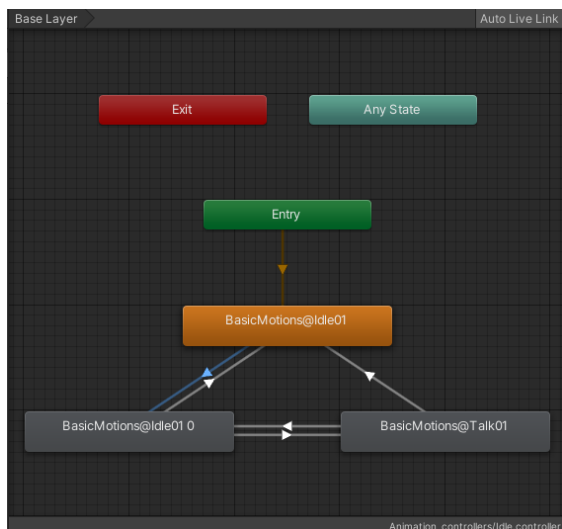
είναι αδρανής (idle) και το δεύτερο αφορά την κατάσταση στην οποία ο χαρακτήρας μιλάει (μια ελαφριά κίνηση στον κορμό, τα χέρια και το κεφάλι). Αυτά τα δύο clips μπήκαν στο κύριο layer του animator controller που συνδέθηκε με τον χαρακτήρα της ρεσεψιόν. Με τον συγκεκριμένο animator controller 4.11a' συνδέθηκαν και δύο παράμετροι (user_in_reception και have_to_talk) οι οποίοι ανάλογα με την τιμή τους (αληθής ή ψευδής) ρυθμίζουν τις μεταβάσεις από το ένα clip στο άλλο. Η τιμή αυτών των παραμέτρων αλλάζει από το script που είναι συνδεδεμένο με το GameObject του player και στο οποίο είναι υλοποιημένος ο client που επικοινωνεί με το διαλογικό σύστημα (backend).

Ανοιγοκλείσιμο των ματιών

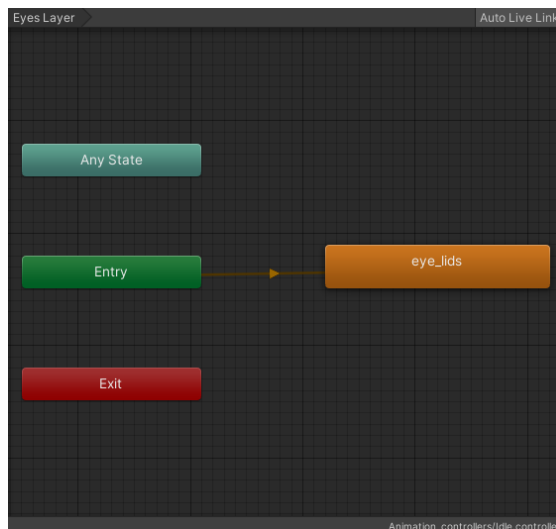
Για το ανοιγοκλείσιμο των ματιών (blinking) δεν υπήρχε τίποτα σχετικό στο Unity Asset Store ή σε κάποια άλλη πηγή. Γι' αυτό το λόγο το animation clip που αφορά αυτήν την κίνηση έπρεπε να δημιουργηθεί από το μηδέν. Ευτυχώς, η συγκεκριμένη κίνηση είναι αρκετά απλούστερη και αποτελείται από τα εξής τρία στάδια. Αρχικά, υπάρχει η κατάσταση ακινησίας που έχει διάρκεια δύομιση δευτερόλεπτα. Μετά το κλείσιμο των βλεφάρων που έχει διάρκεια ένα τέταρτο του δευτερολέπτου. Και τέλος το άνοιγμα των βλεφάρων που έχει κι αυτό διάρκεια ένα τέταρτο του δευτερολέπτου. Η κίνηση αυτή υλοποιήθηκε με χρήση του σχετικού blendshape που περιγράφεται σε παραπάνω ενότητα. Και τα τρία στάδια μαζί συνθέτουν ένα animation clip των τριών δευτερολέπτων. Αυτό το clip μπήκε σε ένα διαφορετικό layer του animator controller 4.11b' προκειμένου να παραλληλοποιηθεί με τα animation clips της κίνησης του σώματος. Επιπλέον, γι' αυτό το layer δημιουργήθηκε και ένα masking 4.11γ' που απομονώνει το κεφάλι του χαρακτήρα από το υπόλοιπο σώμα του, προκειμένου η εφαρμογή του animation clip να μην επηρεάζει μέρη του σώματος άσχετα με την κίνηση.

Περιστροφή του κεφαλιού

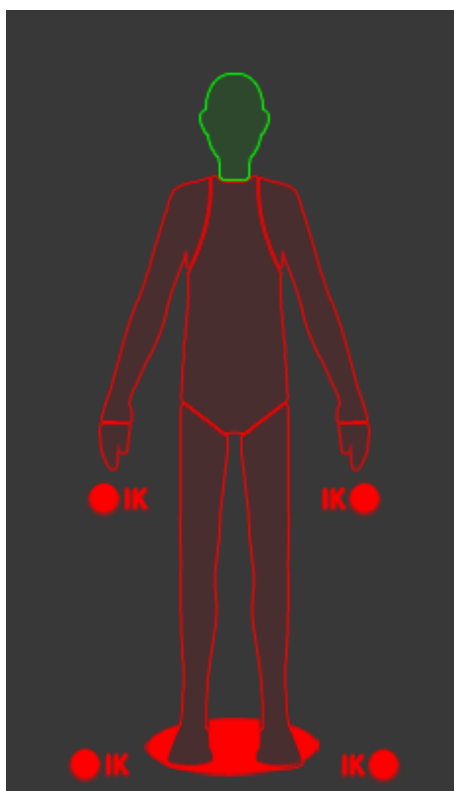
Η τελευταία κίνηση που προσδόθηκε στον χαρακτήρα της ρεσεψιόν ήταν αυτή της περιστροφής του κεφαλιού και του λαιμού. Η κίνηση αυτή ενεργοποιείται, όταν ο controller του χρήστη βρίσκεται κοντά στον γκισέ της ρεσεψιόν. Σε αυτήν την περίπτωση διαφορετικά από όλες τις προηγούμενες, δεν έγινε χρήση του συστήματος του Unity Mecanim. Αντιθέτως, η λειτουργία αυτή υλοποιήθηκε με απευθείας αλλαγή στις τιμές των κατάλληλων μερών του σώματος στο avatar. Πιο συγκεκριμένα, γράφτηκε ένα script που προστέθηκε στο GameObject του χαρακτήρα της ρεσεψιόν, το οποίο είχε ως δείκτες τα bones του avatar που αντιστοιχούσαν στα μάτια (δεξί κι αριστερό), στο λαιμό και το κεφάλι του, αλλά κι έναν δείκτη στη θέση του GameObject που αντιστοιχεί στην οπτική του χρήστη. Με αυτόν τον τρόπο, όταν ο collider του χαρακτήρα της ρεσεψιόν έρχεται σε επαφή με τον collider του GameObject του controller του χρήστη, τότε ενεργοποιείται η διαδικασία περιστροφής όλων αυτών των μερών που αναφέρονται από πάνω, ώστε να κοιτούν προς την οπτικής του χρήστη. Έτσι, δημιουργείται η αίσθηση στο χρήστη ότι ο χαρακτήρας της ρεσεψιόν κοιτάει προς το μέρος του. Τέλος, σε αυτό το component script υλοποιήθηκε και η λειτουργία να δίνεται η επιλογή στον προγραμματιστή να ελέγχει την ταχύτητα με την οποία θα περιστρέφεται το κεφάλι του χαρακτήρα, προκειμένου να βρίσκεται κάθε φορά η προτεινόμενη τιμή για πιο ρεαλιστική αναπαράσταση της κίνησης.



(α) Ο controller για την κίνηση του σώματος.



(β) Ο controller για τις κινήσεις στο κεφάλι.



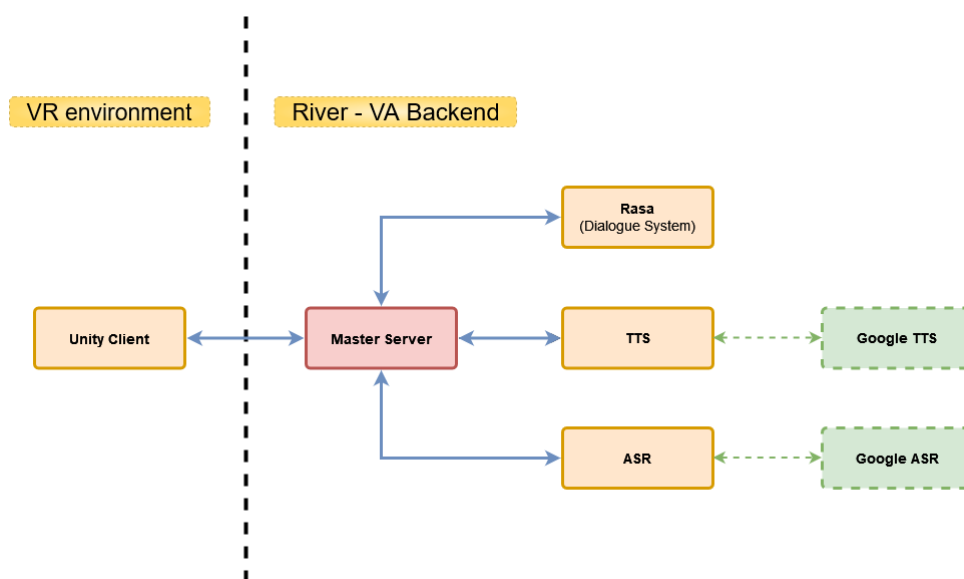
(γ) Το mask για τον δεύτερο controller.

Σχήμα 4.11: Animator Controllers και mask.

4.4 River - VA backend

Στο σημείο της περιγραφής της υλοποίησης μέχρι τώρα και πριν αναφερθούμε στο πώς σχεδιάστηκε και υλοποιήθηκε ο client από την πλευρά του εικονικού περιβάλλοντος, είναι χρήσιμο να περιγραφούν οι εργασίες που έγιναν στο backend σύστημα. Από την ομάδα του ΙΕΛ απέκτησα πρόσβαση στο project "River - VA backend", το οποίο αποτελεί τον πυρήνα του διαλογικού συστήματος και με κάποιες προσθήκες διαμορφώθηκε, ώστε να χρησιμοποιηθεί από την εφαρμογή.

Για αρχή ας γίνει μια περιγραφή της αρχιτεκτονικής και κάποιων από τις επιμέρους λειτουργίες του συστήματος αυτού, πριν τη δικιά μου παρέμβαση. Ουσιαστικά το River είναι ένα σύνολο από υπηρεσίες (set of microservices) που μετατρέπουν ένα ήδη υπάρχον chatbot σε voice chatbot. Τα κύρια μέρη που το αποτελούν είναι τα εξής, ο **Master Server**, ο **Rasa Server**, ο **TTS Server** και ο **ASR Server**. Η δομή του συστήματος που ουσιαστικά αντικατοπτρίζει και την επικοινωνία μεταξύ αυτών των υπηρεσιών είναι δομή αστέρα με τον Master Server να βρίσκεται στο κέντρο του και τα άλλα τρία μέρη να συνδέονται και να επικοινωνούν με αυτόν. Ενδεικτικά, η αρχιτεκτονική του συστήματος της εφαρμογής φαίνεται στην εικόνα που ακολουθεί.



Σχήμα 4.12: Η αρχιτεκτονική της εφαρμογής.

Ο Master Server, λοιπόν, είναι αυτός που μέσω συνδέσεων websocket χειρίζεται την επικοινωνία με τον client, χειρίζεται την επικοινωνία ανάμεσα στα υποσυστήματα του River και συντονίζει τις εργασίες που κάθε φορά εκτελούνται διοχετεύοντας τα μηνύματα του ενός προς το άλλο στην κατάλληλη σειρά και με την κατάλληλη μορφή.

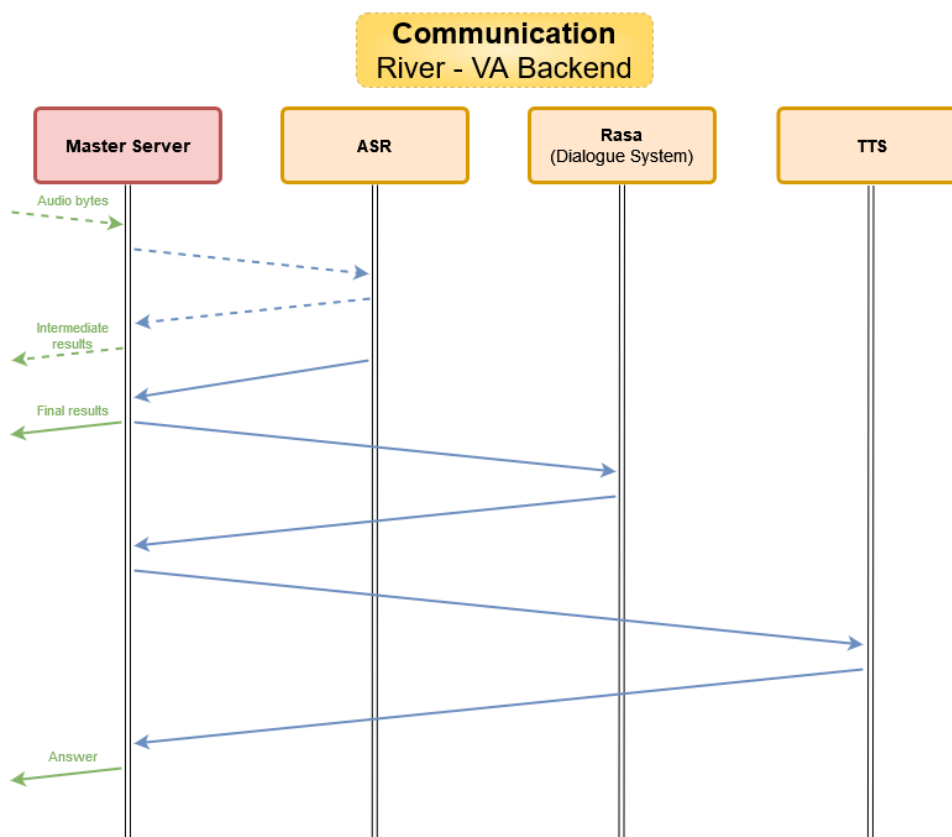
Η σειρά με την οποία γίνονται αυτά είναι η εξής:

1. Ένας client συνδέεται στην κατάλληλη διεύθυνση του Master Server και τον ενημερώνει για το όνομά του και ένα μοναδικό αναγνωριστικό.
2. Τότε ο client ξεκινάει να στέλνει byte ήχου στον Master, ο οποίος με τη σειρά του τα προωθεί στον ASR Server. Ο ASR από την άλλη απαντά άμεσα στον Master με τα πρώτα

ενδιάμεσα αποτελέσματα αναγνωρισμένης φωνής. Αυτά τα αποτελέσματα στέλνονται αντίστοιχα στον client.

3. Όταν ο ASR αναγνωρίσει μια φράση, τότε η φράση αυτή στέλνεται ολόκληρη στον client, ο οποίος άμεσα ενημερώνεται πως πρέπει να σταματήσει να στέλνει byte ήχου και να περιμένει.
4. Παράλληλα, ο Master στέλνει τη φράση που αναγνωρίστηκε στον Rasa Server, ο οποίος μετά από λίγο επιστρέφει την απάντηση του διαλογικού.
5. Τότε ο Master προωθεί την απάντηση από τον Rasa στον TTS Server, ώστε αυτός να συνθέσει το αντίστοιχο φωνητικό μήνυμα, το οποίο και στέλνει πίσω στον Master σε μορφή MP3 με κωδικοποίηση base64.
6. Τέλος, ο Master στέλνει στον client τη γραπτή απάντηση με το φωνητικό μήνυμα.
7. Η ροή εκτέλεσης επιστρέφει στο δεύτερο βήμα με τον client να στέλνει τα byte ήχου, εκτός και αν η απάντηση από τον Rasa συνοδεύεται από το κατάλληλο αναγνωριστικό που σηματοδοτεί το τέλος της συζήτησης.

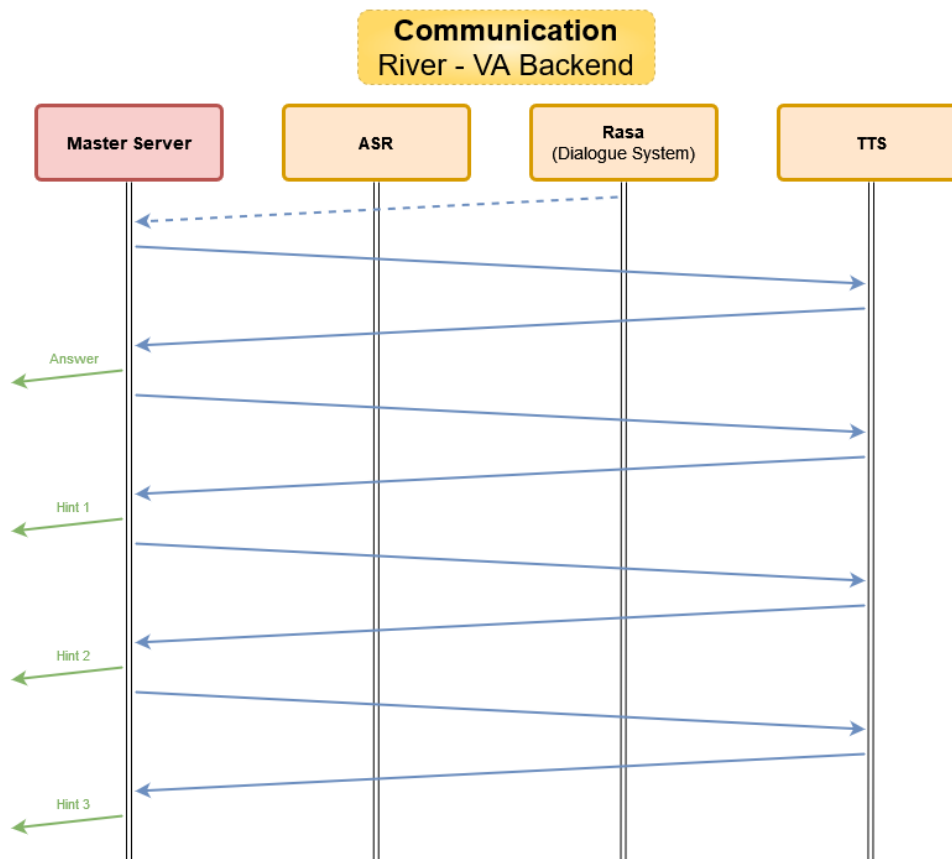
Τα παραπάνω βήματα οπτικοποιούνται στο διάγραμμα που ακολουθεί.



Σχήμα 4.13: Η επικοινωνία στο εσωτερικό του River.

Διαχείριση και αποστολή στον client των hints που παράγει ο Rasa

Μια σημαντική προσθήκη που έγινε στο River και στη ροή εκτέλεσής του, που περιγράφηκε από πάνω, είναι αφενός ο τρόπος με τον οποίο ο Master Server επικοινωνεί με τον Rasa Server κι αφετέρου η διαχείριση και η αποστολή στον client των hints που παράγονται από τον Rasa. Αρχικά λοιπόν ο Master παραμετροποιήθηκε, ώστε να επικοινωνεί με τον Rasa μέσω ενός ειδικού καναλιού, το οποίο είναι ένας custom connector που ονομάζεται `meta_channel`. Περισσότερα για το σήσιμο και τις λειτουργίες αυτού του καναλιού στην ενότητα 5.6. Μέσω αυτού του καναλιού ο Master εκτός από τις απαντήσεις του διαλογικού συστήματος που λάμβανε μέχρι πρότινος, λαμβάνει πλέον και τα σχετικά hints σε μορφή `metadata`. Αυτά τα οργανώνει σε δύο κατηγορίες. Η μία είναι τα hints που αφορούν το ίδιο intent και η άλλη είναι τα hints που αφορούν τρία διαφορετικά intents. Τότε, ανάλογα με την επιλογή που έχει κάνει ο χρήστης στέλνονται στον client τα hints της αντίστοιχης κατηγορίας με το κάθε hint να στέλνεται σε ξεχωριστό μήνυμα. Πριν σταλεί όμως το κάθε hint, προωθείται στον TTS Server, προκειμένου αυτός να συνθέσει το αντίστοιχο φωνητικό μήνυμα. Έτσι, ο χρήστης πέρα από το ότι μπορεί να διαβάσει τα hints, μπορεί και να τα ακούσει. Συνοψίζοντας, για κάθε πρόταση που στέλνεται από τον client προς το River, ο Master Server απαντά στον client εκτός των μηνυμάτων με τα ενδιάμεσα αποτελέσματα του ASR με 4 ακόμη διαδοχικά μηνύματα, την απάντηση του διαλογικού και τα 3 hints. Ακολουθεί το σχετικό διάγραμμα.



Σχήμα 4.14: Η επέκταση της επικοινωνίας στο River για τα hints.

4.5 Διαλογικό σύστημα

4.5.1 Εισαγωγή στο Rasa

Το Rasa¹⁷ είναι ένα σύστημα ανοιχτού κώδικα και μηχανικής μάθησης για αυτοματοποιημένες γραπτές αλλά και προφορικές συζητήσεις. Είναι ένα εργαλείο λοιπόν που βοηθά στο να δημιουργήσεις διαλογικά συστήματα που προσανατολίζονται σε κάποιο σκοπό. Με τον "σκοπό" εννοείται ότι ο χρήστης θέλει να πετύχει έναν στόχο. Με τον όρο "διαλογικό σύστημα" εννοείται ότι πρόκειται για αυτοματοποιημένα συστήματα τα οποία συμμετέχουν σε συζητήσεις με δύο μέρη. Αυτό σημαίνει ότι οι διαλογικοί βοηθοί μπορούν να μιλούν με τον χρήστη, προκειμένου να τον βοηθούν να πετύχει το στόχο τον οποίο έχει. Ο πυρήνας της δημιουργίας ενός διαλογικού βοηθού με το Rasa έχει να κάνει με το να εισάγεις παραδείγματα στο σύστημά σου κι αυτό να μαθαίνει από αυτά. Με αυτόν τον τρόπο, το Rasa επιχειρεί να δημιουργήσει πιο αφηρημένα μοντέλα βασισμένο στα δεδομένα που του δίνεις. Έτσι, πετυχαίνει στο να είναι αρκετά ευέλικτο και να προσαρμόζεται σε οποιοδήποτε θέμα από δραστηριότητες και χόμπι μέχρι την περιγραφή ή υποστήριξη πολύπλοκων συστημάτων.

Σε γενικές γραμμές το Rasa έχει δύο μηχανισμούς που υποστηρίζουν το διαλογικό βοηθό. Ο ένας είναι το **NLU** και ο δεύτερος είναι τα **Dialogue policies**. Με το NLU εννοείται η κατανόηση φυσικής γλώσσας. Στην περίπτωση του Rasa εννοείται το κομμάτι αυτό του συστήματος το οποίο δέχεται απλό κείμενο κι επιστρέφει δεδομένα που είναι αναγνωρίσιμα από τη μηχανή. Με άλλα λόγια μετατρέπει ένα κείμενο σε **intents** και **entities**, των οποίων οι έννοιες θα αποσαφηνιστούν στη συνέχεια. Αυτός ο μηχανισμός μπορεί να λειτουργεί είτε με τη χρήση κανόνων (rule based), όπως για παράδειγμα αν χρησιμοποιείται Regex, είτε με τη χρήση νευρωνικών δικτύων. Η αρχιτεκτονική νευρωνικών δικτύων που χρησιμοποιεί το Rasa ονομάζεται DIET και ουσιαστικά αναλύει τα κείμενα σε intents κι entities βασισμένο σε παραδείγματα που έχει ήδη. Όσον αφορά την πρακτική τους χρήση, η προσέγγιση των κανόνων τείνει να είναι ελαφρύτερη, αλλά ταυτόχρονα απαιτεί και περισσότερα δεδομένα προκειμένου να λειτουργήσει σωστά. Από την άλλη, τα νευρωνικά δίκτυα τείνουν να χρειάζονται περισσότερα δεδομένα για να εκπαιδευτούν καθώς και υπολογιστικούς πόρους, αλλά είναι πολύ καλά στο να χειρίζονται δεδομένα που δεν έχουν ξαναδει ποτέ.

Τα Dialogue policies αναφέρονται στο κομμάτι αυτό του συστήματος που προβλέπει την επόμενη κίνηση του διαλογικού. Η κίνηση δεν καθορίζεται μόνο από το τρέχον intent, αλλά στη γενική περίπτωση επηρεάζεται από ολόκληρη τη μέχρι τότε συζήτηση. Όπως και στο NLU έτσι και στα Dialogue policies, οι πολιτικές με βάση τις οποίες γίνονται οι προβλέψεις μπορούν να βασίζονται είτε σε κανόνες, είτε σε νευρωνικά δίκτυα. Έτσι, το Rasa επιτρέπει την εισαγωγή κανόνων που θα ορίζουν τι χρειάζεται να συμβεί (ελαφρύτερη υλοποίηση), αλλά παρέχει επίσης και ένα νευρωνικό δίκτυο, που ονομάζεται TED και το οποίο επιλέγει την επόμενη καλύτερη κίνηση βασισμένο στο διάλογο μέχρι εκείνο το σημείο, αλλά και όλες τις συζητήσεις με τις οποίες εκπαιδεύτηκε. Όσον αφορά στη χρήση τους, ο τρόπος που προτείνεται είναι να χρησιμοποιούνται παράλληλα και οι δύο πολιτικές.

Ένα project του Rasa πρέπει να έχει μία συγκεκριμένη δομή αρχείων. Τα βασικότερα από αυτά τα αρχεία είναι τα εξής:

¹⁷<https://rasa.com/docs/rasa/>

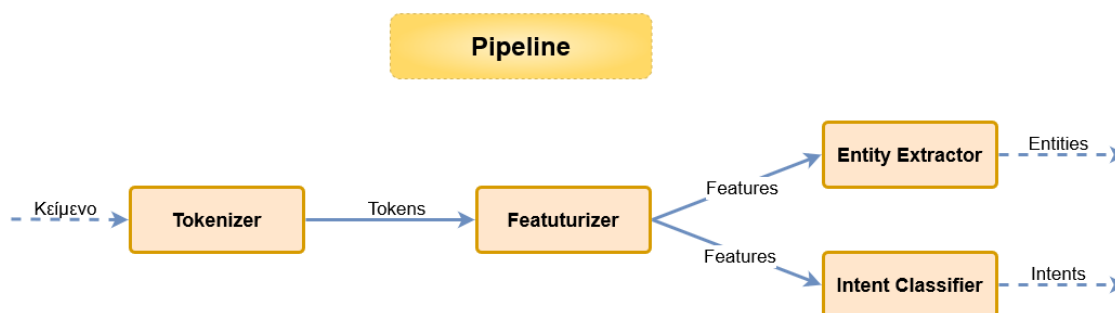
- Το **domain.yml** είναι το αρχείο όπου τα πάντα συνδέονται μεταξύ τους.
- Το **config.yml** είναι το αρχείο που περιέχει τις ρυθμίσεις για τα μοντέλα μηχανικής μάθησης.
- Το **NLU.yml** αρχείο περιέχει παραδείγματα για τα intents και τα entities.
- Το **stories.yml** είναι το αρχείο που περιέχει παραδείγματα ροής των συζητήσεων.
- Το **rules.yml** αρχείο περιέχει τους προκαθορισμένους κανόνες των policies για τις ροές διαλόγων.

Το `domain.yml` είναι το αρχείο που περιέχει όλα όσα ξέρει το διαλογικό σύστημα. Ξέρει, δηλαδή, τις απαντήσεις που ο διαλογικός βοηθός μπορεί να δώσει (**responses**), τις κατηγορίες των πραγμάτων που ο χρήστης μπορεί να πει (**intents**), τις μεταβλητές που μπορούν να χρησιμοποιηθούν κατά τη διάρκεια ενός διαλόγου (**slots**), τις πληροφορίες που μπορούν να εξαχθούν από μια πρόταση του χρήστη (**entities**), και στοιχεία λογικής ακολουθίας και λοιπές επεκτάσεις που ο διαλογικός βοηθός μπορεί να έχει (**forms and actions**). Τα `responses` μπορούν να είναι είτε στατικά, είτε δυναμικά και είτε απλό κείμενο, είτε ένα `custom` μήνυμα. Η διάκριση ανάμεσα στις στατικές και τις δυναμικές απαντήσεις έχει να κάνει με το αν ένα `response` δεν είναι πάντα το ίδιο, ενώ η διάκριση ανάμεσα στις απλές και τις `custom` απαντήσεις έχει να κάνει με το αν αυτή θα είναι απλό κείμενο ή θα περιέχει εικόνες, κουμπιά ή άλλα στοιχεία. Τα `intents` σε αυτό το αρχείο είναι σε μορφή λίστας, ώστε απλά εδώ να ορίζονται. Τα δεδομένα που θα χρησιμοποιηθούν για την εκπαίδευση του διαλογικού που περιέχουν μεταξύ άλλων και τα διάφορα παραδείγματα για τα `intents` βρίσκονται στο αρχείο `NLU.yml`.

Το αρχείο `stories.yml` περιέχει τα δεδομένα που αναπαριστούν τα δεδομένα του μοντέλου που θα το μάθουν για το τι πρέπει να γίνεται σε κάθε επόμενο βήμα. Ουσιαστικά δηλαδή, ο σχεδιαστής καλείται να δημιουργήσει συνηθισμένες ροές συζήτησης, όπως και "χαρούμενα μονοπάτια", αλλά και κάποια `stories` που αποτελούν παρεκβάσεις αυτών. Αυτά τα μονοπάτια έχουν τη μορφή εναλλαγών από `intents` που αναγνωρίζονται σε `actions` που τα ακολουθούν. Τέλος, τα `stories` αυτά μπορούν να εμπλουτίζονται και από άλλα στοιχεία, όπως είναι τα `checkpoints` τα οποία είναι σημεία αναφοράς στη ροή ενός διαλόγου, ώστε να μεταπηδάς από ένα σημείο σε ένα άλλο. Το αρχείο `rules.yml` περιέχει δεδομένα τα οποία μοιάζουν με απλούστερα και μικρότερα σε μήκος `stories` με τη διαφορά πως αυτά τα μονοπάτια πρέπει πάντα να ακολουθούνται με τον ίδιο τρόπο. Η βασική διαφορά ενός `rule` και ενός `story` είναι ότι το διαλογικό βλέπει το `story` ως ένα παράδειγμα από το οποίο μπορεί να μάθει, ενώ ένα `rule` είναι ένα μοτίβο το οποίο πρέπει να ακολουθεί πάντα.

Όπως ειπώθηκε από πάνω το Rasa χρησιμοποιεί μηχανική μάθηση για να προβλέπει αφενός ποιο είναι το `intent` στα λόγια του χρήστη κι αφετέρου για να αποφασίζει ποια θα είναι η επόμενη κίνηση του διαλογικού. Η παραμετροποίηση του Rasa, προκειμένου να ρυθμίζεται ποια ακολουθία εργασιών θα ακολουθείται για αυτές τις εργασίες, γίνεται μέσω του αρχείου `config.yml`. Το αρχείο αυτό αποτελείται από δύο μέρη, το **pipeline** και το **policies**. Το `pipeline` αφορά την πρόβλεψη των `intents` και την εξαγωγή των `entities`, ενώ το `policies` αφορά την επιλογή της επόμενης κίνησης. Το `pipeline`, συγκεκριμένα, μπορεί να

αποτελείται από διάφορα υπομέρη. Τα κυριότερα από αυτά είναι **tokenizers**, **Featurizers**, **Intent Classifiers** και **Entity Extractors**.



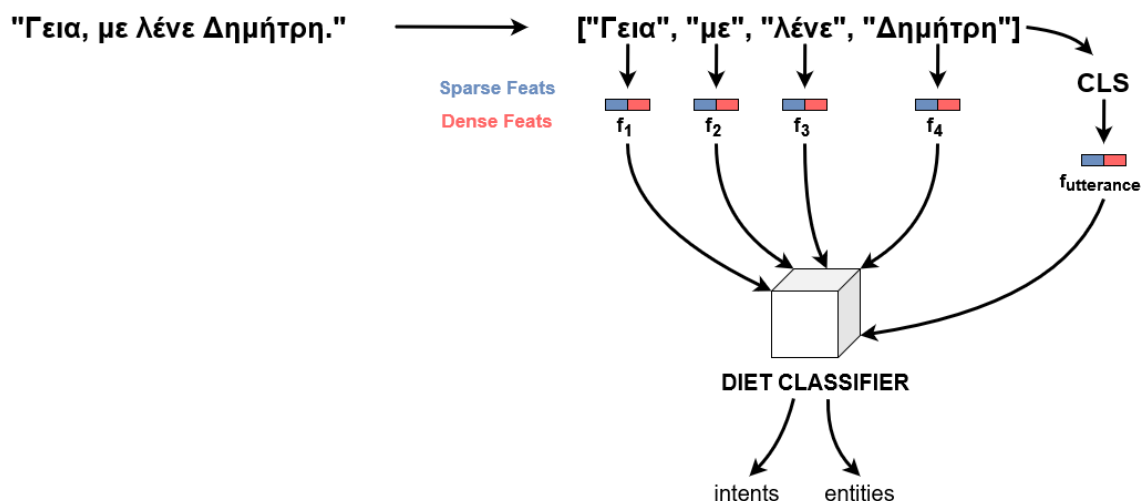
Σχήμα 4.15: Το pipeline του Rasa.

Συνήθως, το πρώτο βήμα είναι ο διαχωρισμός του αρχικού κειμένου σε μικρότερα τμήματα που ονομάζονται tokens. Αυτό πρέπει να γίνει πριν το κείμενο μπει σε διαδικασία χαρακτηρισμού κι επεξεργασίας για το στάδιο της μηχανικής μάθησης. Αυτό το κάνουν οι tokenizers. Αφού το κείμενο χωριστεί σε tokens, τότε περνάει από τη φάση του χαρακτηρισμού, όπου στο κάθε token αποδίδονται αριθμητικές τιμές (machine learning features). Αυτή τη λειτουργία την επιτελούν οι featurizers.

Όπως φαίνεται και στο παράδειγμα της εικόνας 4.16, οι τύποι features που μπορούν να παραχθούν είναι δύο. Η πρώτη είναι τα sparse features που συνήθως παράγονται από έναν CountVectorizer, ο οποίος δημιουργεί μια αναπαράσταση ομάδων από λέξεις με τα μηνύματα του χρήστη, απαντήσεις του διαλογικού συστήματος και intents. Ουσιαστικά, μετράει πόσες φορές μια λέξη από τα δεδομένα εκπαίδευσης του διαλογικού εμφανίζεται στο κείμενο που εξετάζεται. Στην περίπτωση του CountVectorizer κι ανάλογα με τις ιδιαιτερότητες των δεδομένων θα μπορούσε αντί για εμφανίσεις λέξεων να μετράει εμφανίσεις συλλαβών. Όπως και να έχει, αυτά τα sparse features που παράγονται δίνονται στη συνέχεια ως είσοδος στον intent classifier. Η δεύτερη κατηγορία features είναι τα dense τα οποία παράγονται από την ενσωμάτωση ήδη εκπαιδευμένων μοντέλων. Το πιο σύνθετο από τα μοντέλα αυτά είναι το SpacyNLP που χρησιμοποιείται από τον SpaCyFeaturizer. Το Rasa εφαρμόζει αυτούς τους χαρακτηρισμούς σε όλα τα tokens, αλλά παράγει και features για ολόκληρες τις προτάσεις. Τα αποτελέσματα της διαδικασίας αυτής αναφέρονται συνήθως ως CLS.

Το επόμενο βήμα αφορά δύο ανεξάρτητες μεταξύ τους διαδικασίες. Αυτές αντιστοιχούν στους intents Classifiers και στους Entity Extractors. Ο classifier που συστήνεται από το ίδιο το Rasa είναι ο DIET. Πρόκειται για μια αρχιτεκτονική που βασισμένη σε έναν μετασχηματιστή μοιράζεται ταυτόχρονα σε δύο σκοπούς, να κάνει την ταξινόμηση των intents, αλλά και να αναγνωρίζει τα entities. Αν και ο DIET είναι δυνατό να εκπαιδευτεί, ώστε να εντοπίζει ο ίδιος τα entities, δεν είναι απαραίτητο να χρησιμοποιείται για όλες τις περιπτώσεις. Για παράδειγμα, τα entities που έχουν δομή που ακολουθεί κάποιο μοτίβο, δε χρειάζονται κάποιο αλγόριθμο για να τα εντοπίζει. Για αυτά αρκεί η χρήση ενός regex.

Τα policies αφορούν τη φάση πρόβλεψης του επόμενου action. Για αυτό το σκοπό το Rasa δεν ελέγχει μόνο τα intents και entities που έχουν αναγνωριστεί αμέσως πριν, αλλά κοιτάει σε όλο το διάλογο που έχει γίνει μέχρι εκείνη τη στιγμή. Το αρχείο config.yml έχει



Σχήμα 4.16: Παράδειγμα ροής ενός pipeline.

αρχικοποιηθεί με τις εξής τρεις πολιτικές :

- Η **RulePolicy** χειρίζεται συζητήσεις που ταιριάζουν με το προκαθορισμένο μοτίβο κάποιου κανόνα. Τις προβλέψεις τις κάνει βασιζόμενη σε οποιοδήποτε κανόνα από το αρχείο rules.yml.
- Η **MemoizationPolicy** ελέγχει αν η τρέχουσα συζήτηση ταιριάζει με κάποια από τα stories των δεδομένων εκπαίδευσης. Αν ναι, τότε προβλέπει την επόμενη κίνηση από το story που αντιστοιχήθηκε.
- Η **TEDPolicy** χρησιμοποιεί μεθόδους μηχανικής μάθησης για να προβλέψει την επόμενη κίνηση.

Υπάρχουν όμως ακόμη δύο policies που είναι διαθέσιμες από το Rasa, οι οποίες είναι οι εξής:

- Η **UnexpectEDIntentPolicy** είναι βοηθητική πολιτική και πρέπει να χρησιμοποιείται σε συνδυασμό με μία τουλάχιστον ακόμη πολιτική, γιατί το μόνο action που μπορεί να προβλέψει είναι το ειδικό. **action_unlikely_intent**, το οποίο γίνεται όταν από την ακολουθία του pipeline έχει αναγνωριστεί κάποιο intent που σε προγενέστερο χρόνο είχε χαρακτηριστεί ως μη αναμενόμενο.
- Η **AugmentedMemoizationPolicy** αποτελεί μια επέκταση της MemoizationPolicy. Η λειτουργία που έχει προστεθεί είναι η δυνατότητα να ξεχνά κάποια βήματα από το ιστορικό της συζήτησης και να προσπαθεί να εντοπίσει κάποια κοινά σημεία της τρέχουσας συζήτησης με το μειωμένο πια ιστορικό.

Ο τρόπος που οι πολιτικές αυτές συνδυάζονται είναι ο εξής. Κάθε μία από αυτές κάνει την πρόβλεψη της για την επόμενη κίνηση, την οποία πρόβλεψη τη συνοδεύει με ένα σκορ που αντιστοιχεί στο πόσο σίγουρη είναι γι' αυτήν. Τελικά, επιλέγεται το action της policy με τη μεγαλύτερη αυτοπεποίθηση. Όταν δύο πολιτικές έχουν την ίδια σιγουριά για την πρόβλεψή τους, τότε το Rasa έχει ένα μηχανισμό προτεραιότητας για να παίρνει την απόφαση του.

Αυτός ο μηχανισμός ορίζει το πώς ο διαλογικός βοηθός παίρνει αποφάσεις όταν περισσότερες από μια πολιτικές προβλέπουν το επόμενο action με την ίδια σιγουριά. Η προεπιλεγμένη σειρά προτεραιότητας στο Rasa είναι η εξής:

- 6 για τη RulePolicy,
- 3 για τη MemoizationPolicy ή τη AugmentedMemoizationPolicy,
- 2 για την UnexpectTEDIntentPolicy και
- 1 για την TEDPolicy.

Ακόμη, δεν προτείνεται να έχεις πάνω από μια πολιτικές με την ίδια προτεραιότητα στην παραμετροποίηση σου. Αν έχεις δύο πολιτικές με την ίδια προτεραιότητα, τότε αν και οι δύο τους προβλέπουν με την ίδια σιγουριά, τότε το action που θα ακολουθήσει το διαλογικό επιλέγεται τυχαία.

Ο λόγος για τον οποίο οι πολιτικές RulePolicy και MemoizationPolicy έχουν υψηλότερη προτεραιότητα είναι γιατί βασίζονται άμεσα στα δεδομένα που έχει εισάγει ο ίδιος ο σχεδιαστής. Τα rules και τα stories που έχουν προστεθεί, δίνονται ως παραδείγματα για το πώς θα έπρεπε να εξελίσσονται οι συζητήσεις. Από την άλλη, η TEDPolicy είναι μια πολιτική που βασίζεται στη μηχανική μάθηση και προσπαθεί να γενικεύσει τα δεδομένα που έχουν εισαχθεί, προκειμένου να χειρίζεται συζητήσεις που δεν έχουν οριστεί στα αρχεία stories.yml και rules.yml. Από αυτή τη σχεδίαση του συστήματος πρόβλεψης του Rasa φαίνεται ότι αν και είναι πολύ θετικό το να μπορούμε να ξεφεύγουμε από τους κανόνες και τα μονοπάτια συζήτησης που έχουμε ορίσει, συνήθως εμπιστευόμαστε καλύτερα τα δικά μας δεδομένα σε σχέση μια στατιστική προσέγγιση ενός συστήματος μηχανικής μάθησης.

4.5.2 Υλοποίηση του διαλογικού μοντέλου

Οι εργασίες που αφορούν το διαλογικό κομμάτι του backend έχουν να κάνουν με τη δημιουργία ενός Rasa μοντέλου και την παραμετροποίηση (fine tuning) των διαφόρων λειτουργιών γύρω από αυτό. Έτσι, με οδηγό τη θεωρητική περιγραφή των λειτουργιών του Rasa, όπως αυτές έγιναν σε προηγούμενο κεφάλαιο, θα σκιαγραφηθούν τα βήματα που έγιναν στο κομμάτι της υλοποίησης αυτού του κομματιού. Πιο συγκεκριμένα, γίνεται αναφορά από τη μία στα δεδομένα εκπαίδευσης που δόθηκαν στο NLU και από την άλλη στο πώς στήθηκαν η αρχιτεκτονική και οι μηχανισμοί των pipeline και policies.

Δεδομένα εκπαίδευσης NLU

Τα δεδομένα που δόθηκαν για την εκπαίδευση του διαλογικού μοντέλου είναι εξειδικευμένα στο σενάριο διαλόγου ανάμεσα στον χρήστη και την υπάλληλο της ρεσεψιόν. Δεν χρησιμοποιήθηκαν, δηλαδή, δεδομένα που έχουν να κάνουν με τη γενική χρήση της γλώσσας από κάποιο προεκπαιδευμένο μοντέλο. Έτσι, όπως θα δούμε και στη συνέχεια, το περιεχόμενο των διαλόγων είναι πολύ συγκεκριμένο και αφορά κυρίως πράγματα που λέγονται στα πλαίσια εξυπηρέτησης ενός ατόμου από ένα ξενοδοχείο. Επειδή, όμως, σκοπός της εφαρμογής είναι και η όσο το δυνατόν πιο έντονη εμπύθιση σε αυτήν με τη συμμετοχή του

χρήστη σε ρεαλιστικά σενάρια καθημερινής χρήσης της γλώσσας, έπρεπε να στηθεί και ένα στοιχειώδες πλαίσιο προφορικής αλληλεπίδρασης, που θα προσομοίαζε την πραγματικότητα.

Έτσι, τα intents που δόθηκαν στα δεδομένα που δεν είναι εξειδικευμένα στο σενάριο της ρεσεψιόν είναι τα εξής:

- χαιρετισμός (έναρξη συζήτησης),
- έκφραση ευχαριστιών κι ανταπόδοση αυτών,
- θετική απάντηση,
- επιβεβαίωση μιας κατάστασης,
- αρνητική απάντηση,
- αίτημα για επανάληψη ή περαιτέρω εξήγηση και
- αποχαιρετισμός (τερματισμός συζήτησης).

Από την άλλη, τα intents που είναι εξειδικευμένα στο σενάριο της ρεσεψιόν είναι αυτά:

- επιβεβαίωση μιας ήδη υπάρχουσας κράτησης,
- κράτηση ενός νέου δωματίου,
- παραχώρηση διάφορων προσωπικών στοιχείων (όνομα, τηλέφωνο, αριθμός ταυτότητας / διαβατηρίου κλπ),
- απαντήσεις σε ερωτήσεις της υπαλλήλου για την κράτηση (έναρξη και διάρκεια της διαμονής, μέγεθος του δωματίου κ.ά.),
- ερωτήματα του χρήστη που αφορούν την κράτηση (κόστος του δωματίου, θέα του δωματίου κ.ά.),
- ερωτήματα για παροχές του ξενοδοχείου (ίντερνετ, πάρκινγκ, εστιατόριο, σάουνα κ.ά.),
- αίτημα για βοήθεια (για την πρόσβαση στο δωμάτιο, για το κουβάλημα των αποσκευών κ.ά.) και
- απαντήσεις σε γενικές / συμβουλευτικές ερωτήσεις της υπαλλήλου (π.χ. αν επισκέπτεται πρώτη φορά την πόλη, αν θέλει συμβουλές για διάφορα αξιοθέατα κ.ά.).

Ακόμη, δημιουργήθηκαν και τα intents με τα responses για να περιγράψουν την περίπτωση κατά την οποία η επιβεβαίωση μιας κράτησης δεν μπορεί να ολοκληρωθεί. Σε αυτά περιλαμβάνονται η έκφραση του προβληματισμού από την πλευρά του χρήστη, μια σειρά ερωταπαντήσεων γύρω από τη διαδικασία που ακολουθήθηκε για να πραγματοποιηθεί η κράτηση και τέλος η αντιμετώπιση του κολλήματος με εξηγήσεις για το τι έφταιξε, το ποια λύση βρέθηκε και το καινούργιο κόστος της κράτησης.

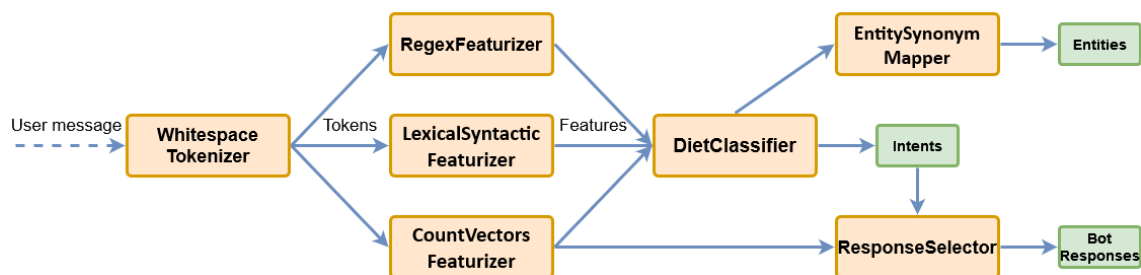
Εκτός από το να οριστούν τα intents αυτά στο αρχείο domain.yml, δόθηκαν διάφορα παραδείγματα για το καθένα από αυτά στο αρχείο NLU.yml. Στη συνέχεια, πάλι στο αρχείο

domain.yml, για κάθε ένα από αυτά τα intents ορίστηκαν όλες οι προτάσεις που θα μπορούσε να πει η υπάλληλος, είτε αυτές αφορούν απαντήσεις σε ερωτήματα του χρήστη, είτε πρόκειται για ερωτήσεις προς τον χρήστη. Μετά από αυτό, στο αρχείο stories.yml δόθηκαν αρκετές διαφορετικές περιπτώσεις ακολουθιών από intents και responses, σκιαγραφώντας ποια θεωρούνται τα “**χαρούμενα μονοπάτια**”, ώστε το διαλογικό να εκπαιδευτεί κατάλληλα. Οι ακολουθίες αυτές διακρίνονται σε δύο βασικές θεματικές, την κράτηση ενός δωματίου και τις παροχές που προσφέρει το ξενοδοχείο. Και για τις δύο θεματικές δόθηκαν αρκετές παραλλαγές, ώστε το μοντέλο που θα εκπαιδευόταν να είναι ευέλικτο και να δίνει τη αίσθηση πως υπάρχει ποικιλία στο πώς εξελίσσεται ένας διάλογος ανάμεσα στο χρήστη και την υπάλληλο. Επίσης, δημιουργήθηκε μια σχέση εξάρτησης ανάμεσα σε αυτά τα δύο, τοποθετώντας ένα checkpoint στο τέλος των μονοπατιών που εμπίπτουν στη θεματική της κράτησης και στην αρχή των μονοπατιών που εμπίπτουν στη θεματική των παροχών του ξενοδοχείου. Ουσιαστικά, τα **checkpoints** λειτουργούν ως δείκτες που ενώνουν ξεχωριστά μεταξύ τους μονοπάτια. Με αυτόν τον τρόπο, ο χρήστης αναγκάζεται να ολοκληρώσει τη διαδικασία κράτησης ή επιβεβαίωσης μιας κράτησης ενός δωματίου, για να μπορεί να ξεκινήσει κάποιο διάλογο που αφορά τις παροχές του ξενοδοχείου.

Pipeline

Η δομή του pipeline που χρησιμοποιείται είναι η ακόλουθη 4.17. Σε κάθε pipeline ένας tokenizer είναι αρκετός. Αυτός που αποφασίστηκε να μπει στην περίπτωση του συγκεκριμένου μοντέλου του Rasa είναι ο **WhitespaceTokenizer**. Πρόκειται για έναν tokenizer που για κάθε χαρακτήρα κενού (white space) που συναντά, δημιουργεί ένα νέο token. Είναι ο απλούστερος tokenizer που θα μπορούσαμε να χρησιμοποιήσουμε, αλλά μιας και δεν υπάρχει ανάγκη για περαιτέρω έλεγχο προσαρμοσμένο σε κάποια γλώσσα, κρίθηκε ως ο καλύτερος. Στη συνέχεια ακολουθούν τρεις featurizers, οι οποίοι είναι ο **RegexFeaturizer**, ο **LexicalSyntacticFeaturizer** και ο **CountVectorsFeaturizer**. Σε αυτό το σημείο να σημειωθεί ότι δεν επιλέχθηκε να γίνει καμία ενσωμάτωση κάποιου προεκπαιδευμένου μοντέλου. Συνεπώς, δεν υπάρχει ανάγκη να γίνει η εκπαίδευση σε μια συγκεκριμένη γλώσσα. Αντίθετα, υπάρχει η δυνατότητα το μοντέλο που θα εκπαιδευτεί να στοχεύει στον τομέα του σεναρίου. Για αυτό το λόγο και οι τρεις featurizers που εντάχθηκαν στο pipeline είναι sparse. Ο **RegexFeaturizer** είναι ένα εργαλείο που εφόσον προστεθεί στο pipeline, θα παράγει ένα feature το οποίο θα χρησιμοποιήσει ο intent classifier για να δημιουργεί μοτίβο (patterns) στην αναγνώριση intents. Είναι σημαντικό να διευκρινιστεί πως ο **RegexFeaturizer** αν και παρέχει στον intent classifier δεδομένα, δεν προβλέπει το intent άμεσα. Αντίθετα, πρέπει να συμπεριληφθούν αρκετά παραδείγματα που περιέχουν κανονικές εκφράσεις (regex), έτσι ώστε ο classifier να μάθει να χρησιμοποιεί τα features που παράγονται. Ο **LexicalSyntacticFeaturizer**, λαμβάνοντας ως είσοδο μια ακολουθία από tokens, παράγει μια ακολουθία από features, όπου το -οστό feature αντιστοιχεί στο ν-οστό token. Τα features αυτά που παράγονται περιέχουν πληροφορίες από τη λεξιλογική και συντακτική ανάλυση των tokens. Ο τελευταίος featurizer είναι ο **CountVectorsFeaturizer**, ο οποίος περιγράφηκε σε προηγούμενη ενότητα στα πλαίσια εισαγωγής για το pipeline του Rasa. Να σημειωθεί όμως ότι παραμετροποιήθηκε, ώστε να μετράει από μονογράμματα μέχρι και τετραγράμματα, τα οποία μάλιστα θα πρέπει να

βρίσκονται ολόκληρα μέσα στα όρια μιας λέξης. Αυτή η παραμετροποίηση προέκυψε ύστερα από διαδικασία βελτιστοποίησης, ώστε στο λιγότερο χρόνο εκπαίδευσης να επιτυγχάνονται τα καλύτερα αποτελέσματα. Μετά από τους featurizers βρίσκεται ο **DietClassifier**. Αυτός προβλέπει τα intents και τα κατηγοριοποιεί, ενώ εξάγει και τα entities, εφόσον υπάρχουν. Ακολουθεί ο **EntitySynonymMapper** ο οποίος είναι ένα εργαλείο που εντοπίζει τα συνώνυμα που υπάρχουν στις τιμές των entities, ανάλογα και με το πώς και αν αυτά έχουν οριστεί στα δεδομένα εκπαίδευσης. Να σημειωθεί πως αυτό το μέρος του pipeline πρέπει πάντα να συμβαίνει μετά από οποιαδήποτε εξαγωγή των entities, γι' αυτό και τοποθετήθηκε σε αυτό το σημείο. Στο τέλος αυτού του pipeline έχει τοποθετηθεί ο **responseselector**. Πρόκειται για ένα εργαλείο που μπορεί να χρησιμοποιηθεί ώστε να γίνεται άμεση πρόβλεψη και ανάκτηση της απάντησης του διαλογικού, εφόσον τα intents που αναγνωρίστηκαν επιτρέπουν κάτι τέτοιο. Να επισημανθεί ότι ο response selector μπορεί να αξιοποιεί τα αποτελέσματα που προκύπτουν και από τον CountVectorsFeaturizer (sparse features), αλλά και από τον DietClassifier (intents). Στο διάγραμμα που ακολουθεί φαίνεται η αρχιτεκτονική του υλοποιημένου pipeline.



Σχήμα 4.17: Το pipeline που υλοποιήθηκε.

Policies

Όπως ειπώθηκε και σε προηγούμενο κεφάλαιο, ο διαλογικός βοηθός χρησιμοποιεί policies για να αποφασίσει ποια actions θα ακολουθήσει σε κάθε βήμα στη συζήτηση. Επίσης, υπάρχουν διαφόρων ειδών policies, όπως για παράδειγμα οι πολιτικές που βασίζονται σε κανόνες, αλλά και οι πολιτικές που βασίζονται στη μηχανική μάθηση. Στην περίπτωση αυτού του διαλογικού μοντέλου επιλέχθηκαν τρεις πολιτικές, οι οποίες είναι η **MemoizationPolicy**, η **TEDPolicy** και η **RulePolicy**. Επιπλέον, η προτεραιότητα ανάμεσα στις πολιτικές αυτές επιλέχθηκε να μείνει όπως ήταν από προεπιλογή. Σε μια ιεραρχία όπου ο μεγαλύτερος αριθμός αντιστοιχεί και σε υψηλότερη προτεραιότητα, η RulePolicy έχει 6, η MemoizationPolicy έχει 3 και η TEDPolicy έχει 1.

4.6 Παραγωγή των hints

Το κομμάτι παραγωγής των hints είναι αρκετά ενδιαφέρον μιας και είναι ένα πεδίο σχετικά ανεξερευνήτο στο χώρο των chatbots. Αυτό που ουσιαστικά είναι ο σκοπός της λειτουργικότητας των hints είναι για κάθε απάντηση που παράγεται από το Rasa για τον χρήστη, να υπάρχουν και κάποιες προτάσεις ως ενδεικνυόμενες που θα μπορούσε να χρησιμοποιήσει ο χρήστης για να συνεχίσει η συζήτηση. Στη διαδικασία να βρεθεί ένας τρόπος να υλοποιηθεί αυτή η λειτουργία, εμφανίστηκαν διάφορες λύσεις, από τις οποίες αξίζει να αναφερθούν μερικές.

Μια ιδέα που φαίνεται αρκετά τετριμμένη, αλλά θα μπορούσε να είναι σε κάποιο βαθμό λειτουργική θα ήταν από την πλευρά ανάπτυξης του εικονικού περιβάλλοντος (Unity), να έχει στηθεί ένα πρόγραμμα που θα είχε στη διάθεσή του κάποιες λέξεις-κλειδιά και αντιστοιχίες αυτών των λέξεων με μία ή περισσότερες προτάσεις. Αυτές οι λέξεις-κλειδιά, όπως και οι αντίστοιχες απαντήσεις, θα έπρεπε να έχουν σχέση με το αντικείμενο του κάθε σεναρίου. Διαφορετικά, στις περισσότερες των περιπτώσεων αν ο χρήστης ακολουθούσε τα hints αυτά, η ροή του διαλόγου δε συνέχιζε σε ένα καλό μονοπάτι (happy path). Έτσι, φαίνεται ότι ο προγραμματιστής του εικονικού περιβάλλοντος της εφαρμογής θα έπρεπε να ανατρέξει στα δεδομένα που χρησιμοποιήθηκαν για την εκπαίδευση του διαλογικού μοντέλου, να εντοπίσει τα παραδείγματα των intents και των σχετικών utterances και στη συνέχεια να δημιουργήσει κάτι σαν βάση δεδομένων στο πρόγραμμα που θα έχει τη μορφή λεξιλογίου (λέξεις-κλειδιά - προτάσεις). Μια παραλλαγή αυτής της προσέγγισης θα μπορούσε να έχει ακριβώς το ίδιο σκεπτικό, αλλά το σχετικό πρόγραμμα θα ήταν υλοποιημένο σε ένα πιο πίσω επίπεδο, παραδείγματος χάριν στον Master Server του backend.

Αυτή η λύση δεν είναι καλή για διάφορους λόγους. Ο κυριότερος από αυτούς είναι πως τα hints που θα παράγονται θα είναι αρκετά περιορισμένα και στο πλήθος και στην ποικιλία μιας και θα πρόκειται απλά για μια απλή αναπαραγωγή των δεδομένων που συγκεντρώθηκαν όταν γινόταν η υλοποίηση. Δεν θα είναι δηλαδή αυτό που αρχικά υπήρχε σαν στόχος, τα smart hints. Ο δεύτερος, αλλά εξίσου σημαντικός, λόγος είναι ότι κάθε φορά που θα υπάρχει ανάγκη για ένα διαφορετικό διάλογο στο ίδιο ή σε κάποιο καινούργιο σενάριο, θα πρέπει να γίνεται η ίδια διαδικασία. Από τις παραπάνω χρήσιμες παρατηρήσεις, προκύπτουν δύο πολύ βασικά κριτήρια για την υλοποίηση. Θέλουμε μια λύση που είναι **portable**, δηλαδή η υλοποίηση δεν πρέπει να εξαρτάται από τον κάθε ξεχωριστό διάλογο. Και θέλουμε μια λύση που να παράγει “έξυπνα” hints, ώστε αυτά να είναι πραγματικά χρήσιμα για τη συνέχεια του διαλόγου, αλλά να επιτυγχάνεται κι αρκετή ποικιλία, ώστε να ευνοείται ο εκπαιδευτικός χαρακτήρας των hints.

Με αυτά στο μυαλό, καταλήξαμε πως η καλύτερη λύση από άποψη φορητότητας, αλλά και παραγωγής εύστοχων κι έξυπνων hints, θα ήταν μια υλοποίηση που θα είναι όσο το δυνατόν “πιο κοντά” στον πυρήνα του Rasa και το διαλογικό μοντέλο που χρησιμοποιείται. Επίσης, ο Rasa Server παρέχει κάποιες διευθύνσεις (**endpoints**) μέσω των οποίων είναι δυνατόν να αποκτηθεί πρόσβαση στο αρχείο των συζητήσεων (**trackers of conversations**), αλλά και να τροποποιηθεί το περιεχόμενό του. Πιο συγκεκριμένα, υπάρχουν endpoints μέσω των οποίων μεταξύ άλλων εκτελούνται ενέργειες σχετικές με τους trackers (ανάκτηση ολόκληρου του αρχείου μιας συζήτησης - **tracker**, πρόσθεση, αφαίρεση ή αντικατάσταση

γεγονότων στο αρχείο - **append event list**, πρόβλεψη του επόμενου action του διαλογικού κ.λπ.), ενέργειες σχετικές με το μοντέλο που είναι "φορτωμένο" από το Rasa (εκπαίδευση ενός μοντέλου, αξιολόγηση ενός intent - **evaluate**, πρόβλεψη ενός action με βάση μια προσωρινή κατάσταση του tracker - **predict**, ανάλυση μιας πρότασης - **parse** κ.λπ.).

Αυτά τα δεδομένα είναι πολύ χρήσιμα στην περίπτωση μας και φαίνεται ότι από αυτά προκύπτουν δύο δυνατές υλοποιήσεις. Η μία είναι με χρήση των **custom actions**¹⁸ και η δεύτερη είναι με χρήση των **custom connectors**¹⁹. Το μοντέλο που χρησιμοποιεί το Rasa λειτουργεί ως εξής. Μετά από κάθε μήνυμα του χρήστη, προβλέπει ένα Action που στη συνέχεια το διαλογικό σύστημα θα εκτελέσει. Αυτά τα Actions μπορούν να αντιστοιχούν είτε σε κάποιες προεπιλεγμένες λειτουργίες, όπως το να δοθεί μια απάντηση πίσω στο χρήστη, να δημιουργηθεί μια φόρμα που ο χρήστης καλείται να συμπληρώσει κ.ά., είτε σε custom actions που μπορούν να τρέξουν όποιον κώδικα θέλει ο προγραμματιστής, όπως για παράδειγμα να γίνει μια κλήση σε κάποιο API²⁰, να γίνει ένα ερώτημα σε κάποια βάση δεδομένων κ.λπ.. Από την άλλη, ο Rasa Server επιτυγχάνει την επικοινωνία του με άλλους Servers και πλατφόρμες με τη χρήση connectors που διαχειρίζονται τα εισερχόμενα και εξερχόμενα μηνύματα. Για παράδειγμα, η επικοινωνία με τον Master Server επιτυγχάνεται μέσω ενός προεπιλεγμένου REST καναλιού. Εκτός από αυτό όμως, διαθέτει και πολλούς έτοιμους connectors για επικοινωνία με διάφορες πλατφόρμες, όπως είναι το Facebook Messenger, το Slack, το Telegram, το Twilio κ.ά.²¹. Τέλος, όπως και στα Actions, έτσι και στην περίπτωση των connectors, είναι δυνατόν να υλοποιηθεί ένας custom connector που εκτός από τη διαχείριση των μηνυμάτων από και προς το Rasa μπορεί να τρέξει κι οποιοδήποτε άλλο κώδικα.

Από τις δύο αυτές δυνατές υλοποιήσεις επιλέχθηκε αυτή με τους custom connectors. Η συγκεκριμένη επιλογή κρίθηκε ως η καλύτερη, γιατί θεωρείται και η πιο portable ως λύση. Αυτό οφείλεται στο ότι η λειτουργία ενός τέτοιου connector, αλλά και ολόκληρη η διαδικασία παραγωγής κι αποστολής των hints είναι ανεξάρτητη από τον πυρήνα εργασιών που γίνονται σε ένα chatbot. Έτσι, για την υλοποίηση του connector, ο οποίος ονομάστηκε **meta_channel** δε χρειάστηκε να γίνει καμία αλλαγή στο μοντέλο του Rasa ή στα Actions αυτού. Για να γίνει πιο σαφές, αν επιλεγόταν η χρήση των custom actions, τότε θα έπρεπε να παρέμβουμε στο μοντέλο του Rasa και πιο συγκεκριμένα στα stories του, κάτι που σημαίνει πως για κάθε αλλαγή ή προσθήκη στις λειτουργίες της υλοποίησης θα έπρεπε να το εκπαιδεύουμε. Εκτός την εργασία που θα χρειάζεται κάθε φορά και το έξτρα κόστος χρόνου για την εκάστοτε εκπαίδευση, αυτή η επιλογή κρίνεται καλύτερη κι από άποψη διάκρισης των διαφορετικών στοιχείων ενός συστήματος, κάτι που κάνει τον κώδικά του πιο ευανάγνωστο και τη χρήση του πιο λειτουργική. Για παράδειγμα, στην περίπτωση των custom actions, αν η λειτουργία των hints δεν είναι επιθυμητή και ο χρήστης ήθελε απλά να επικοινωνήσει με το chatbot, θα έπρεπε να εκπαιδευτεί ένα νέο μοντέλο που δεν την εμπεριέχει. Ενώ τώρα, στην περίπτωση του custom connector, αν δε θες να κάνεις χρήση της συγκεκριμένης λειτουργίας, απλά επικοινωνείς με το chatbot μέσω του προεπιλεγμένου REST καναλιού και ταυτόχρονα το

¹⁸<https://rasa.com/docs/rasa/custom-actions/>

¹⁹<https://rasa.com/docs/rasa/connectors/custom-connectors>

²⁰<https://rasa.com/docs/rasa/http-API/>

²¹<https://rasa.com/docs/rasa/messaging-and-voice-channels/>

διαλογικό μοντέλο που χρησιμοποιείται και στη μία, αλλά και στην άλλη περίπτωση είναι το ίδιο.

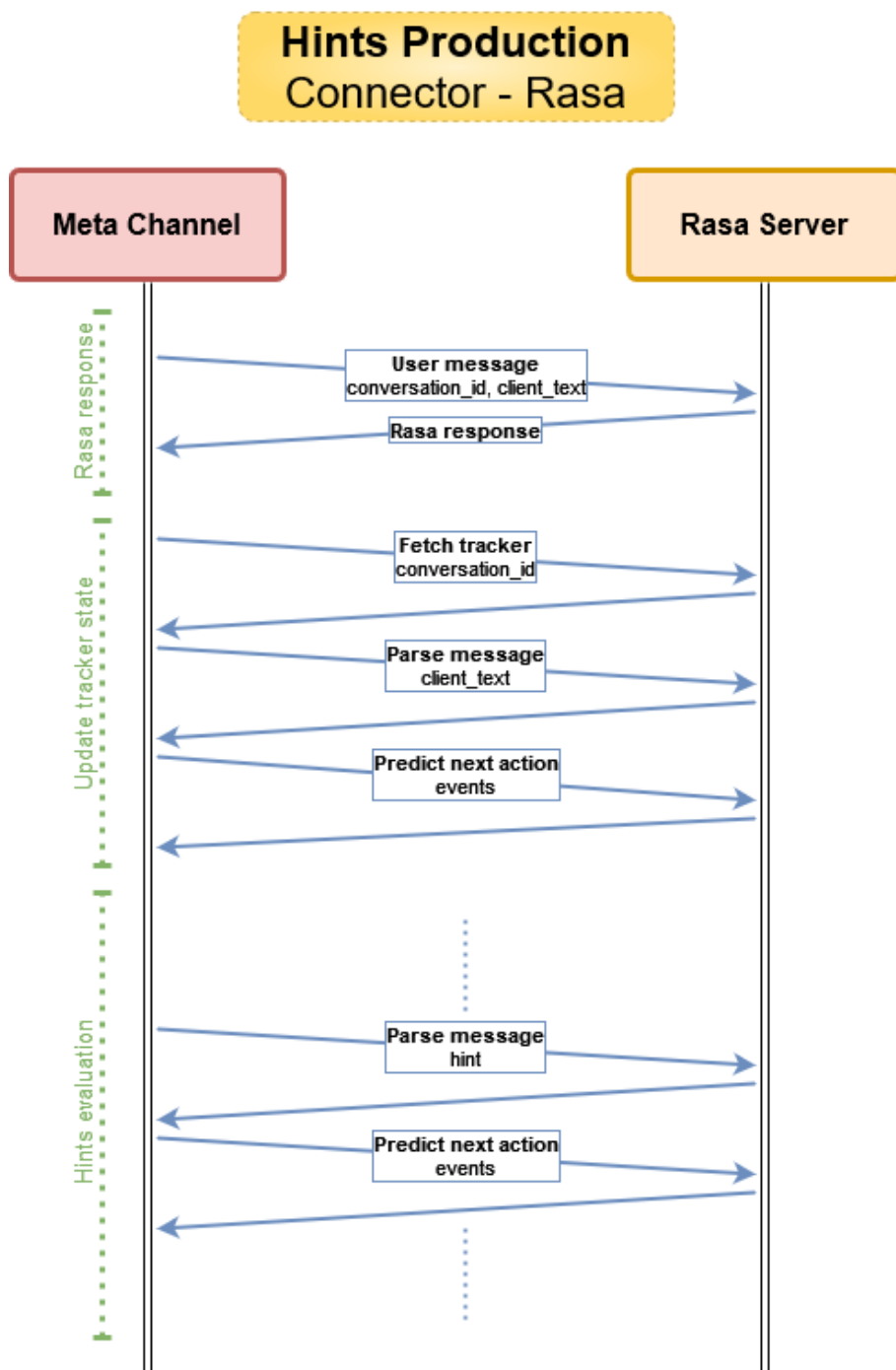
Αφού περιγράφηκαν τα κριτήρια με βάση τα οποία επιλέχθηκε το πώς θα υλοποιηθεί η λειτουργία των smart hints, ας περιγραφούν οι ενέργειες που εκτελούνται στο meta_channel. Το meta_channel, όπως και οι περισσότεροι connectors του Rasa, αποτελούνται από δύο κλάσεις. Η μία διαχειρίζεται τη ροή εισερχομένων μηνυμάτων στο Rasa Server (MetaInputChannel) και η άλλη την αντίστοιχη ροή των εξερχομένων μηνυμάτων (MetaOutputChannel). Η ροή των συμβάντων έχει ως εξής. Αρχικά, με το που λαμβάνεται ένα εισερχόμενο μήνυμα, δημιουργείται ένα αντικείμενο που είναι στιγμιότυπο της κλάσης MetaInputChannel. Αμέσως μετά, εξάγονται όλα τα στοιχεία του εισερχόμενου μηνύματος (αποστολέας, περιεχόμενο μηνύματος και metadata) και δημιουργείται μια σύνδεση του μηνύματος αυτού με το MetaOutputChannel. Τότε, φεύγει ο έλεγχος ροής από το MetaInputChannel και μεταφέρεται στον πυρήνα του Rasa, όπου γίνονται όλες οι λειτουργίες για την παραγωγή της απάντησης από το διαλογικό σύστημα. Όταν αυτές ολοκληρωθούν, αρχικοποιείται ένα αντικείμενο που είναι στιγμιότυπο της κλάσης MetaOutputChannel.

Η παραγωγή των hints γίνεται με τα ακόλουθα βήματα :

1. Ανακτάται το στιγμιότυπο του tracker της συζήτησης, όπως αυτή έχει διαμορφωθεί μέχρι εκείνη τη στιγμή. Το σχετικό endpoint του Rasa Server API είναι το `"/conversations/<conversation_id:path>/tracker"`.
2. Ενημερώνεται ο tracker, προσθέτοντας σε αυτόν τα events που αφορούν το τελευταίο μήνυμα του χρήστη. Αυτό συμβαίνει, γιατί η προσθήκη των καινούργιων events στον tracker γίνεται μετά την ολοκλήρωση των εργασιών του MetaOutputChannel, δηλαδή μετά και την αποστολή της απάντησης πίσω στο χρήστη.
 - (α) Γίνεται ανάλυση του μηνύματος του χρήστη, ώστε να εξαχθούν οι χρήσιμες πληροφορίες (intents, entities κλπ). Το σχετικό endpoint του Rasa Server API είναι το `"/model/parse"`.
 - (β) Προστίθεται το event του μηνύματος στον tracker με τις πληροφορίες που εξήχθησαν στο προηγούμενο βήμα.
 - (γ) Γίνεται πρόβλεψη του επόμενου Action του chatbot με βάση τον ενημερωμένο tracker. Το σχετικό endpoint του Rasa Server API είναι το `"/model/predict"`.
 - (δ) Προσθήκη του event του επόμενου Action του chatbot στον tracker.
 - (ε) Τα βήματα c και d επαναλαμβάνονται διαδοχικά, μέχρι το επόμενο Action που προβλέπεται για το chatbot να είναι η μετάβαση σε κατάσταση αναμονής μηνύματος από το χρήστη. Αυτό συμβαίνει, γιατί σε ένα μοντέλο του Rasa είναι αρκετά συνηθισμένο να εκτελούνται περισσότερα από ένα διαδοχικά Actions.
 - (ς) Προσθήκη του τελευταίου event του chatbot στον tracker.
3. Ανακτάται, μέσω του κατάλληλου αρχείου εκπαίδευσης του διαλογικού μοντέλου, η λίστα των διαφόρων intents και τριών παραδειγμάτων για το καθένα.
4. Σε αυτό το στάδιο γίνεται αξιολόγηση των intents, που αποτελούν τα ενδεχόμενα hints. Για κάθε intent:

- (α) Επιλέγεται τυχαία ένα από τα τρία παραδείγματα του intent και γίνεται ανάλυση αυτού, ώστε να εξαχθούν οι χρήσιμες πληροφορίες (intents, entities κ.λπ.). Το σχετικό endpoint του Rasa Server API είναι το `"/model/parse"`.
 - (β) Γίνεται προσθήκη του event του παραδείγματος στον tracker με τις πληροφορίες που εξήχθησαν στο προηγούμενο βήμα.
 - (γ) Γίνεται πρόβλεψη του επόμενου Action του chatbot με βάση τον ενημερωμένο tracker. Το σχετικό endpoint του Rasa Server API είναι το `"/model/predict"`.
 - (δ) Αν το Action που προβλέπεται ως απάντηση του chatbot με τον υψηλότερο βαθμό δεν είναι το προεπιλεγμένο Action, που χρησιμοποιείται σε περίπτωση που το chatbot δεν ξέρει τι να απαντήσει, τότε αυτό σημαίνει ότι το hint είναι χρήσιμο για να συνεχιστεί ο διάλογος. Τότε, το hint με τα τρία παραδείγματά του και τον καλύτερο βαθμό που πήρε το επόμενο προβλεπόμενο Action αποθηκεύονται σε μια κατάλληλη λίστα.
5. Η λίστα με τα hints που κρίθηκαν χρήσιμα ταξινομείται σε αύξουσα σειρά με βάση το βαθμό που πήραν τα αντίστοιχα προβλεπόμενα Actions και επιλέγονται τα τρία καλύτερα και εισάγονται στα metadata του μηνύματος που θα σταλεί ως απάντηση του chatbot.

Τότε η απάντηση στέλνεται πίσω ακριβώς στη διεύθυνση από την οποία στάλθηκε το μήνυμα του χρήστη. Συγκεκριμένα, η απάντηση φτάνει πρώτα στον Master Server και στη συνέχεια με τον τρόπο που περιγράφεται σε προηγούμενη ενότητα προωθείται μαζί με τα hints στον client του εικονικού περιβάλλοντος. Ακολουθεί διάγραμμα στο οποίο οπτικοποιούνται τα παραπάνω βήματα για την παραγωγή των hints [4.18](#).



Σχήμα 4.18: Επικοινωνία meta_channel και rasa server για την παραγωγή των hints.

4.7 Στήσιμο του client - συντονισμός του frontend

Αφού δημιουργήθηκαν όλα τα στοιχεία που έχουν να κάνουν με το εικονικό περιβάλλον, όπως είναι οι χώροι, οι ανθρωπίνοι χαρακτήρες, η κινησιολογία αυτών και οι διεπαφές της εφαρμογής με το χρήστη κι ενσωματώθηκε η λειτουργία του "River - VA backend" με την επέκταση για την παραγωγή των smart hints έπρεπε να υλοποιηθούν και οι εξής δύο βασικές λειτουργίες. Η πρώτη είναι η επικοινωνία με τον Server του backend της εφαρμογής και η δεύτερη ο συντονισμός όλων των στοιχείων του εικονικού περιβάλλοντος που περιγράφονται σε αυτό το κεφάλαιο. Αυτές οι δύο λειτουργίες υλοποιήθηκαν μαζί στα πλαίσια ενός Script που προσδόθηκε ως Component στο GameObject OVRPlayerController. Η επιλογή το Script αυτό να ανατεθεί στο GameObject που είναι υπεύθυνο για την κίνηση στον χώρο και τον έλεγχο των διεπαφών (controllers) του χρήστη δεν ήταν μονόδρομος. Αντιθέτως, έγινε αφενός για λόγους νοηματικής συνάφειας κι αφετέρου γιατί εξυπηρετούσε, όπως θα φανεί και στη συνέχεια, στην άμεση προσπέλαση διαφόρων στοιχείων του Scene. Κατά τα άλλα, θα μπορούσε κάλλιστα το Script να ανατεθεί σε άλλο GameObject, ακόμα κι αν αυτό είχε δημιουργηθεί μόνο για το συγκεκριμένο σκοπό, με τον πυρήνα της υλοποίησης να είναι ίδιος με αυτόν που τελικά έγινε.

Οι λειτουργίες που έχει αυτό το Component λοιπόν είναι οι ακόλουθες:

- Επικοινωνία με τον Master Server του River μέσω WebSockets.
- Χρήση του μικροφώνου του headset και εκπομπή του stream αυτού προς τον Master Server.
- Αναπαραγωγή των ηχητικών μηνυμάτων από το διαλογικό σύστημα, είτε πρόκειται για τις απαντήσεις, είτε για τα hints.
- Ενημέρωση των στοιχείων διεπαφής με τον χρήστη.

Αρχικοποιήσεις

Πριν πραγματοποιηθεί οποιαδήποτε ενέργεια, το Script είναι υπεύθυνο για την αρχικοποίηση όλων των εργαλείων και στοιχείων που θα χρησιμοποιηθούν στη συνέχεια, όπως είναι ο εντοπισμός του μικροφώνου του headset, η ρύθμιση του websocket να "βλέπει" τον Master Server, η σύνδεση με όλα τα στοιχεία του UI (από ολόκληρα τα Canvas μέχρι το κάθε εικονικό κουμπί ή στοιχείο κειμένου), η σύνδεση με τα στοιχεία που είναι απαραίτητα για το lip sync κλπ. Επίσης, γίνονται κάποιοι απαραίτητοι έλεγχοι για να είναι βέβαιος ο χρήστης ότι η εφαρμογή μπορεί να χρησιμοποιηθεί κανονικά. Ένας από αυτούς, είναι ότι το headset είναι συνδεδεμένο στο internet. Στην περίπτωση που δεν υπάρχει πρόσβαση στο internet εμφανίζεται σχετική ενημέρωση.

Ρόλος των Colliders

Σε αυτό το σημείο αξίζει να γίνει αναφορά στη χρήση που γίνεται από το συγκεκριμένο Script στους Colliders. Σε αυτήν την περίπτωση, δε χρησιμοποιούνται με σκοπό να οριοθετήσουν κάποια περιοχή, κάνοντας χρήση των ιδιοτήτων τους με βάση τους κανόνες της

φυσικής, αλλά να ενεργοποιήσουν μια προκαθορισμένη ενέργεια, μέσω της ιδιότητας των Colliders να κάνουν trigger κάποια γεγονότα. Ορίστηκε, λοιπόν, ένας Collider με γεωμετρικό σχήμα ορθογώνιου παραλληλεπίπεδου και προσδόθηκε ως Component στο GameObject που αντιστοιχεί στο avatar της υπαλλήλου της ρεσεψιόν, με την κατάλληλη παραμετροποίηση, έτσι ώστε όταν εισέρχεται και εξέρχεται από αυτό ένας άλλος Collider να προκαλείται trigger. Έτσι, τα triggers αυτά που αφορούν τον Collider του GameObject OVRPlayerController τα πιάνει το Script, μέσω των συναρτήσεων OnTriggerEnter και OnTriggerExit, και εκτελούνται οι εξής ενέργειες.

Ενέργειες που δρομολογούνται

Αρχικά, με χρήση του websocket API²² πραγματοποιείται σύνδεση του client με τον Server που έχει σηκωθεί από τον Master Server και ανταλλάσσονται τα κατάλληλα στοιχεία (όνομα, id κλπ). Παράλληλα, ενεργοποιείται το reception_canvas, το οποίο μέχρι τότε ήταν αόρατο, πάνω από τη ρεσεψιόν. Στη συνέχεια, ενεργοποιείται η καταγραφή του μικροφώνου και το stream αυτής προς το backend, προκειμένου να γίνει η αναγνώριση φωνής από τον ASR Server. Το sample rate με βάση το οποίο λειτουργεί ο ASR είναι στα 16000 Hz. Για αυτό το λόγο και η καταγραφή του μικροφώνου ρυθμίστηκε να είναι τόσο. Επίσης, η αποστολή των πακέτων ήχου σε τεμάχια των 100ms, τα οποία αντιστοιχούν σε 1600 bytes ανά τεμάχιο. Αφού ολοκληρώνεται το κάθε τεμάχιο, τότε μετατρέπεται μέσω του εργαλείου OpenWavParser²³ που προσφέρει η βιβλιοθήκη της Unity σε bytes και στη συνέχεια μετατρέπεται με βάση την κωδικοποίηση MP3 base64. Από εδώ και πέρα, ξεκινάει η παραλαβή από τον Master Server μιας σειράς από μηνύματα, τα οποία αντιστοιχούν στα αποτελέσματα που δίνει ο ASR. Τα μηνύματα αυτά εμφανίζονται στο σχετικό παράθυρο του reception_canvas. Στη συνέχεια, στέλνονται κι άλλα μηνύματα που αφορούν στην απάντηση του διαλογικού συστήματος, αλλά και στα hints. Όταν ληφθούν αυτά, τότε ενημερώνονται και τα υπόλοιπα παράθυρα του reception_canvas, ενημερώνεται το αρχείο της συζήτησης του menu_canvas και αποθηκεύονται τα αρχεία ήχου της απάντησης και των τριών hints, αφού πρώτα έχουν μετατραπεί από τη μορφή MP3 base64 σε σειρά από bytes. Τότε, ενεργοποιείται αυτόματα και η αναπαραγωγή του ήχου της απάντησης, ώστε να την ακούσει ο χρήστης με το που εμφανιστεί και στο σχετικό παράθυρο. Επίσης, αφού αποθηκεύτηκαν τα αρχεία ήχου, ο χρήστης πατώντας τα αντίστοιχα εικονικά κουμπιά μπορεί να τα ακούσει. Όταν ολοκληρωθεί η αναπαραγωγή της απάντησης, ξεκινάει πάλι η καταγραφή και η αποστολή της ομιλίας του χρήστη από το μικρόφωνο.

Εναλλακτικές ροές εργασιών

Εδώ να επισημανθεί πως η ροή εργασιών έτσι όπως αυτή μόλις περιγράφηκε αντιστοιχεί χάριν συντομίας στην περιγραφή της πιο συνηθισμένης ροής που μπορεί να προκύψει. Κατά τη διάρκεια υλοποίησης όμως του συγκεκριμένου Script προβλέφθηκαν και ετοιμάστηκαν οι λειτουργίες για όλα τα ενδεχόμενα που μια ενέργεια μπορεί να διακόψει ή να παύσει προσωρινά μια άλλη που ήδη εκτελείται ή που περισσότερες από μία ενέργειες χρειάζεται

²²<https://github.com/endel/NativeWebSocket>

²³<https://assetstore.unity.com/packages/tools/audio/open-wav-parser-90832>

να εκτελεστούν παράλληλα. Για παράδειγμα, ο χρήστης μπορεί ανά πάσα στιγμή να απομακρυνθεί από το χώρο της ρεσεψιόν κι αυτό να σημαίνει πως θα πρέπει να τερματιστεί οποιαδήποτε εργασία γίνεται που αφορά την επικοινωνία (stream μικροφώνου, αναμονή για κάποια απάντηση, αναπαραγωγή κάποιας απάντησης κλπ). Ή ο χρήστης μπορεί να πατήσει το κατάλληλο κουμπί για να ακούσει ένα από τα hints, ενώ το μικρόφωνο του είναι σε διαδικασία καταγραφής κι αποστολής πακέτων ήχου. Ο τρόπος να δοθεί αυτή η δυνατότητα στο Script ήταν με χρήση συγκεκριμένων μεταβλητών (flags) που σημαίνουν την επίτρεψη εκτέλεσης μιας συγκεκριμένης περιοχής κώδικα ή μη. Έπρεπε δηλαδή να εντοπιστούν όλα τα ενδεχόμενα race conditions [53], ώστε να αποφευχθεί το να προκύψουν καταστάσεις deadlock [54].

Κεφάλαιο 5

Αξιολόγηση και Συμπεράσματα

Το κεφάλαιο αυτό αποτελεί ουσιαστικά τον επίλογο της εργασίας. Σε αυτό παρουσιάζονται η αξιολόγηση που ακολουθήθηκε κατά τη διάρκεια ανάπτυξης της εφαρμογής με τις επιδράσεις που είχε στις επιλογές που έγιναν και η προετοιμασία και οι παρατηρήσεις της πιλοτικής αξιολόγησης. Στη συνέχεια, γίνεται μια αποτίμηση της προσπάθειας να δημιουργηθεί η βάση για μια καινοτόμα εφαρμογή εκμάθησης γλωσσών, χρησιμοποιώντας ως σημεία αναφοράς το τι βελτιώσεις επιτεύχθηκαν και που αυτές εφαρμόστηκαν και τα state-of-the-art χαρακτηριστικά που έχει το prototype. Τέλος, γίνεται μια σκιαγράφηση των επόμενων βημάτων που πρέπει να γίνουν. Αυτά αφορούν αφενός την εξέλιξη του prototype σε μια ολοκληρωμένη εκπαιδευτική εφαρμογή κι αφετέρου την προσθήκη νέων λειτουργιών που θα βελτιώνουν στο state-of-the-art και σε άλλους τομείς. Οι ενότητες που περιέχει το κεφάλαιο είναι οι εξής:

1. Φάσεις αξιολόγησης.
2. Αποτίμηση και μελλοντικά βήματα.

5.1 Φάσεις αξιολόγησης

Κατά τις φάσεις του σχεδιασμού και της υλοποίησης του prototype έγιναν κάποιες αξιολογήσεις από διάφορους χρήστες, έτσι ώστε να βρεθούν και να λυθούν όποια προβλήματα ευχρηστίας εντοπιζόνταν. Αυτή η μορφή αξιολόγησης ονομάζεται διαμορφωτική και παρόλο που πραγματοποιήθηκε από απλούς χρήστες που δεν αντιπροσώπευαν τους τελικούς χρήστες της εφαρμογής κατάφεραν και ανέδειξαν κάποια προβλήματα. Οι αδυναμίες ή παραλείψεις που αναδείχθηκαν είχαν να κάνουν με τα εξής θέματα.

5.1.1 Διαμορφωτική αξιολόγηση

Η κίνηση του χρήστη είχε κάποιες αδυναμίες. Μάλιστα, αυτό το στοιχείο είναι από τα βασικότερα για την εμπύθιση του χρήστη και για αυτό είναι από τα πρώτα στα οποία δίνει την περισσότερη σημασία. Έτσι, είναι κι από τα πιο εύκολα για να εντοπιστούν όποια προβλήματα έχει. Συγκεκριμένα, το βασικότερο ζήτημα είχε να κάνει με την ομαλότητα της μετακίνησης και της περιστροφής στην οπτική του χρήστη. Η κίνηση είχε υλοποιηθεί, ώστε να ξεκινά και να σταματά απότομα με το που ο χρήστης έδινε την αντίστοιχη εντολή από τα

controllers. Αυτό, όμως, πέρα από το ότι προσέθετε δυσκολία στον έλεγχο της, κούραζε τον χρήστη και του προκαλούσε ναυτία. Ο τρόπος με τον οποίο αντιμετωπίστηκε αυτό το πρόβλημα ήταν να οριστεί ένα χρονικό μεταβατικό διάστημα ανάμεσα στις καταστάσεις ακινησίας και κίνησης όπου η ταχύτητα μεταβολής της θέσης και του προσανατολισμού αυξανόταν ή μειωνόταν σταδιακά. Ένα επίσης σημαντικό ζήτημα που επιλύθηκε μέσω της φάσης αυτής της αξιολόγησης, ήταν ότι βρέθηκαν όλα τα σημεία που η κίνηση του χρήστη εμποδιζόταν επειδή υπήρχε κάποιο αντικείμενο του εικονικού χώρου σε σημείο που δεν έπρεπε ή διακοπτόταν απότομα με το χαρακτήρα του να μεταφέρεται εκτός της οριοθετημένης περιοχής του σεναρίου.

Αδυναμίες εντοπίστηκαν και στις λειτουργίες του UI. Πιο συγκεκριμένα, σχεδόν όλα τα κουμπιά αποδείχθηκαν δύσκολα στο να πατηθούν λόγω του μικρού μεγέθους τους, αλλά και την αυξημένη ευαισθησία στον έλεγχο του pointer από το δεξί χειριστήριο. Αυτή η παρατήρηση είχε ως αποτέλεσμα αφενός να αυξηθεί το μέγεθος των κουμπιών κι αφετέρου να γίνει η κίνηση του pointer πιο ομαλή μειώνοντας την ευαισθησία στις κινήσεις του χρήστη. Επιπλέον, μέσω των παρατηρήσεων των χρηστών σε αυτό το στάδιο αξιολόγησης, βρέθηκε η ιδανική σχέση ανάμεσα στο μέγεθος της γραμματοσειράς των κειμένων και την απόσταση των παραθύρων του UI από την οπτική της κάμερας.

Οι οδηγίες που δίνονται στον χρήστη στην αρχή του σεναρίου έπρεπε να αλλάξουν με τον εξής τρόπο. Το κομμάτι που καλούσε τον χρήστη να πλησιάσει την υπάλληλο της ρεσεψιόν, για να ξεκινήσει το διάλογο μαζί της, βρισκόταν στο τέλος του κεντρικού παραθύρου που παρουσιάζονταν οι στόχοι και τα controls των χειριστηρίων. Αυτό είχε ως αποτέλεσμα οι χρήστες να μη δίνουν τόση σημασία σε αυτό και να το ξεχνούν ενώ πολλές φορές παρέλειπαν μέχρι και να το διαβάσουν. Η λύση η οποία βρέθηκε ήταν να εξατομικευτούν οι οδηγίες που δίνονται στον χρήστη ανάλογα με το σκοπό τους, αλλά και το αν ο χρήστης θα χρειαστεί να τις ξαναδιαβάσει. Για παράδειγμα, το μήνυμα που καλούσε τον χρήστη να πλησιάσει τη ρεσεψιόν, ορίστηκε να εμφανίζεται μέσω ενός popup παραθύρου το οποίο άνοιγε μόνο μία φορά μετά το πρώτο κλείσιμο του κεντρικού παραθύρου οδηγιών. Μετά από αυτήν την αλλαγή κανένας χρήστης δεν ξεχνούσε το βήμα αυτό και ο επιτηρητής δε χρειάστηκε να το επισημάνει ξανά.

5.1.2 Πιλοτική αξιολόγηση

Αφού η ανάπτυξη του prototype ολοκληρώθηκε και η εφαρμογή έφτασε στη μορφή που περιγράφεται από πάνω, ετοιμάστηκε ένα **demo**. Αυτό ετοιμάστηκε για χρήση από χρήστες που αντιστοιχούν στο προφίλ τελικών χρηστών της εφαρμογής κι έγινε ένας προγραμματισμός για τον τρόπο που θα έπρεπε να δοκιμαστεί. Πιο συγκεκριμένα, ορίστηκε ότι η δοκιμή αυτή θα πρέπει να γίνεται πάντα υπό την επίβλεψη ενός υπεύθυνου. Ο ρόλος του υπεύθυνου είναι καταρχάς να προετοιμάσει τον χρήστη για την εμπειρία που θα έχει, δίνοντας του οδηγίες για το πώς χρησιμοποιείται ο εξοπλισμός της Oculus (headset και controllers), περιγράφοντάς του τη φιλοσοφία και το σκοπό που έχει η εφαρμογή, εισάγοντας τον στο συγκεκριμένο περιεχόμενο του σεναρίου του demo και απαντώντας του σε όποιες ερωτήσεις έχει. Στη συνέχεια και κατά τη διάρκεια χρήσης της εφαρμογής, ο υπεύθυνος στέκεται δίπλα στον χρήστη και τον καθοδηγεί όπου χρειάζεται. Μάλιστα, αν και η φύση του εξο-

πλισμού και της εφαρμογής είναι τέτοια που μόνο ο χρήστης μπορεί να φοράει το headset και να βρίσκεται στον εικονικό κόσμο, ο υπεύθυνος έχει τη δυνατότητα να παρακολουθεί τη διαδικασία στην οθόνη του υπολογιστή ή του κινητού του, μέσω της screencast υπηρεσίας που προσφέρει το λογισμικό της Oculus. Έτσι, όσο βρίσκεται δίπλα στον χρήστη μπορεί να παρέμβει, για να του εξηγήσει οτιδήποτε χρειαστεί και να τον βοηθήσει να ολοκληρώσει τη χρήση της εφαρμογής. Επίσης, οι χρήστες μπορεί να μην έχουν εμπειρία από σχετικές εφαρμογές εικονικής πραγματικότητας, αλλά κι επειδή στα πλαίσια του συγκεκριμένου δοκιμαστικού δεν υπάρχει πάντα το περιθώριο να εξοικειωθούν με το περιβάλλον του λογισμικού της Oculus ή του demo μιας και αυτό είναι μια διαδικασία που συνήθως απαιτεί αρκετό χρόνο. Για αυτό το λόγο, ο υπεύθυνος έχει τον πολύ σημαντικό ρόλο του να επιβλέπει τον χρήστη στο πως κινείται στον χώρο του πραγματικού κόσμου, ώστε να μην κινδυνεύσει να χτυπήσει είτε γιατί δεν έχει οπτική επαφή με αυτόν, είτε γιατί μπορεί να ζαλιστεί. Από τα παραπάνω προκύπτει ότι ο επιβλέπων πρέπει να μπορεί να συνεννοηθεί καλά με τον χρήστη στη μητρική του ή σε μια τρίτη γλώσσα, όπως είναι τα αγγλικά.

Ερωτηματολόγια

Για τη φάση των δοκιμών αυτών συντάχθηκαν και δύο ερωτηματολόγια που προορίζονταν για συμπλήρωση μετά την εκτέλεση του demo. Το ένα είναι για να απαντηθεί από τον χρήστη και το δεύτερο από τον υπεύθυνο που τον επέβλεπε. Το ερωτηματολόγιο του χρήστη επικεντρώνεται στους παρακάτω άξονες και ζητάει την αξιολόγησή τους:

- Το γλωσσικό, πολιτισμικό και εκπαιδευτικό υπόβαθρο του χρήστη.
- Γενικά χαρακτηριστικά της εφαρμογής που έχουν να κάνουν με την εμπυθιστικότητα του εικονικού κόσμου, την αποκρισιμότητα στην κίνηση στο χώρο, τις διεπαφές επικοινωνίας με τον χρήστη κ.ά..
- Συγκεκριμένες λειτουργίες της εφαρμογής, όπως είναι το διαλογικό σύστημα, τα hints, η υπηρεσία αναγνώρισης φωνής, η υπηρεσία σύνθεσης φωνής κ.ά..
- Τα διάφορα στοιχεία του σεναρίου.
- Την εκπαιδευτική αξία της εφαρμογής.

Το ερωτηματολόγιο του επιτηρητή από την άλλη προσπαθεί να καταγράψει το ρόλο που είχε στη διαδικασία, οπότε οι ερωτήσεις έχουν να κάνουν με τα εξής πράγματα:

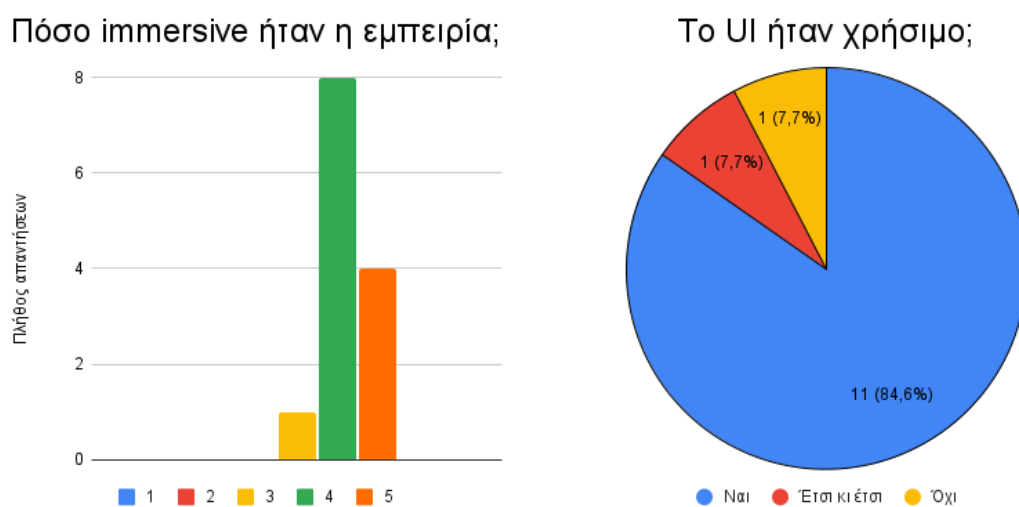
- Τι οδηγίες δόθηκαν στον χρήστη πριν την χρήση της εφαρμογής.
- Γενικές ερωτήσεις γύρω από το αντίστοιχο session που αφορούν τη διάρκεια, ενδεχόμενες παύσεις κ.ά..
- Αν και με ποιον τρόπο παρενέβη κατά τη διάρκεια εκτέλεσης του demo.
- Επιδόσεις του χρήστη, όπως ποιους από τους στόχους που είχαν τεθεί πέτυχε, πόσες φορές ολοκλήρωσε πλήρεις συζητήσεις κ.λπ..

Αρχικά υπήρχε η σκέψη η συγκεκριμένη φάση της αξιολόγησης του prototype να μείνει σε θεωρητικό επίπεδο και να αφηθεί ως μέρος των επόμενων βημάτων στην ανάπτυξη της εφαρμογής. Αυτό συνέβη για δύο κυρίως λόγους. Ο πρώτος ήταν ότι αυτό που έχουμε στα χέρια μας ως έτοιμο προϊόν είναι ουσιαστικά ένα prototype και όχι μια εφαρμογή που έχει τόσο νόημα να τεσταριστεί από κανονικούς χρήστες. Ο δεύτερος είναι ότι το να βρεθούν αρκετοί (για να έχουμε αξιόπιστα αποτελέσματα) άνθρωποι που ανταποκρίνονται στο προφίλ του χρήστη για τον οποίο προορίζεται η εφαρμογή δεν είναι μια εύκολη διαδικασία. Εν τέλει όμως, επειδή μια τέτοια δοκιμή δεν κρίνεται ως ασήμαντη ακόμα και σε αυτό το στάδιο που βρισκόμαστε τώρα, αποφασίσαμε να την κάνουμε και να δούμε τι συμπεράσματα μπορούμε να εξάγουμε από αυτήν.

5.1.3 Αποτελέσματα

Τα συμπεράσματα τα οποία βγήκαν αφορούν διάφορες πτυχές της εφαρμογής κι έχουν να κάνουν με την αξιολόγηση της ως μια εφαρμογή που ανήκει στο φάσμα του εικονικού συνεχούς, τα χαρακτηριστικά του διαλογικού συστήματος που εμπεριέχει καθώς και κάποια από τα στοιχεία αλληλεπίδρασης της με τους χρήστες.

Τα πρώτα αποτελέσματα αφορούν μετρικές που χρησιμοποιούνται γενικά στις VR εφαρμογές. Αν και στο στάδιο που είμαστε τώρα η υλοποίηση μιας εφαρμογής έτοιμης με τελειοποιημένα τα στοιχεία που αφορούν τα γραφικά του εικονικού χώρου δεν ήταν το βασικό μέλημα, τα αποτελέσματα τα οποία είχαμε είναι αρκετά ικανοποιητικά. Συγκεκριμένα, όπως φαίνεται και στα διαγράμματα που ακολουθούν, η εμπειρία χρήσης της εφαρμογής κρίθηκε από τους χρήστες που τη δοκίμασαν ως αρκετά εμπυθιστική, ενώ τα στοιχεία UI, τα οποία παίζουν πολύ σημαντικό ρόλο στον τρόπο με τον οποίο ο χρήστης σχετίζεται και αλληλεπιδρά με τον εικονικό κόσμο που τον περιβάλλει ήταν χρήσιμα. Σε αυτό το σημείο πρέπει να σημειώσουμε και τον βοηθητικό ρόλο που έπαιξε η φάση της διαμορφωτικής αξιολόγησης στην βελτίωση των συγκεκριμένων μετρικών για το βαθμό εμπύθισης των χρηστών.

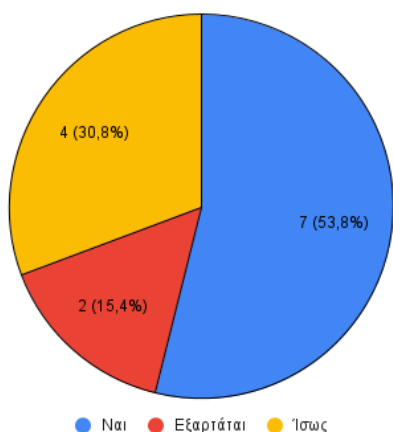


Σχήμα 5.1: Αποτελέσματα σχετικά με τη VR εμπειρία

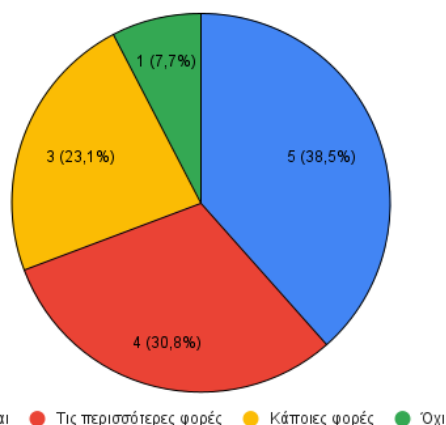
Οι επόμενες παρατηρήσεις έχουν να κάνουν πάλι με την αλληλεπίδραση του χρήστη με την εφαρμογή, αλλά αυτή τη φορά με στοιχεία που είναι πιο εξειδικευμένα στην εκμάθηση γλωσσών. Τα συμπεράσματα που βγαίνουν είναι τα εξής:

- Η μετάφραση σαν λειτουργία θα είναι χρήσιμη.
- Οι επιδόσεις της αναγνώρισης φωνής κρίνονται θετικά με περιθώρια βελτίωσης.
- Ενδιαφέρον έχει το ζήτημα με το αν η καταγραφή της φωνής του χρήστη γίνεται αυτόματα ή με μια λειτουργία τύπου push-to-talk (όπως για παράδειγμα γίνεται στο Mondly VR). Από τη μία το να γίνεται αυτόματα προσφέρει μια φυσικότητα στο διάλογο και κάνει την εμπειρία πιο ρεαλιστική. Από την άλλη το push-to-talk μπορεί να περιορίσει το θόρυβο και να βοηθήσει τον χρήστη να απομονώσει τις φράσεις που τον ενδιαφέρουν περισσότερο. Εν τέλει, όσο βελτιώνονται τα συστήματα αναγνώρισης φωνής αυτό το δίλημμα τόσο λιγότερο νόημα θα έχει.

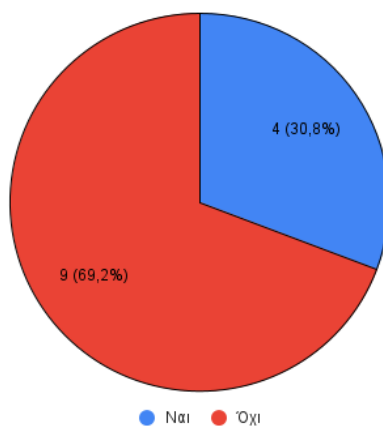
Η μετάφραση θα ήταν χρήσιμη;



Το ASR δούλεψε καλά;

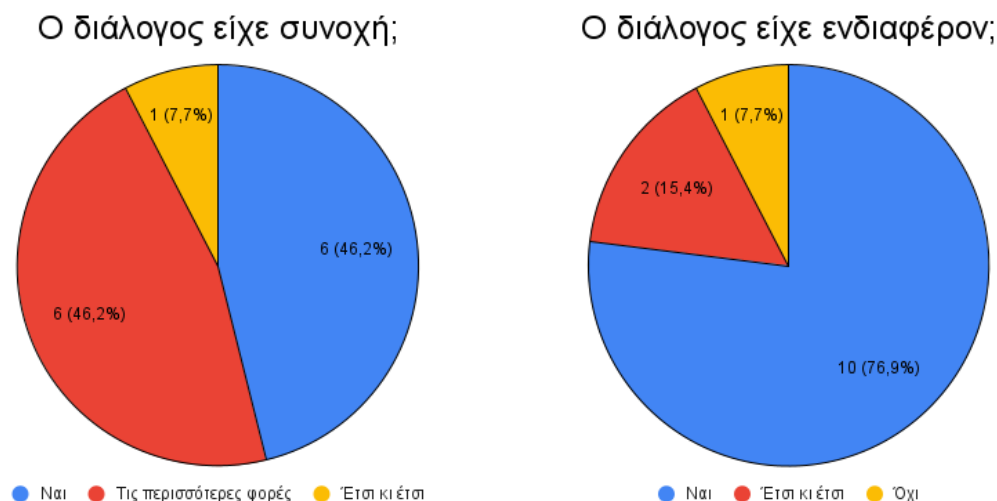


Θα προτιμούσατε push-to-talk;



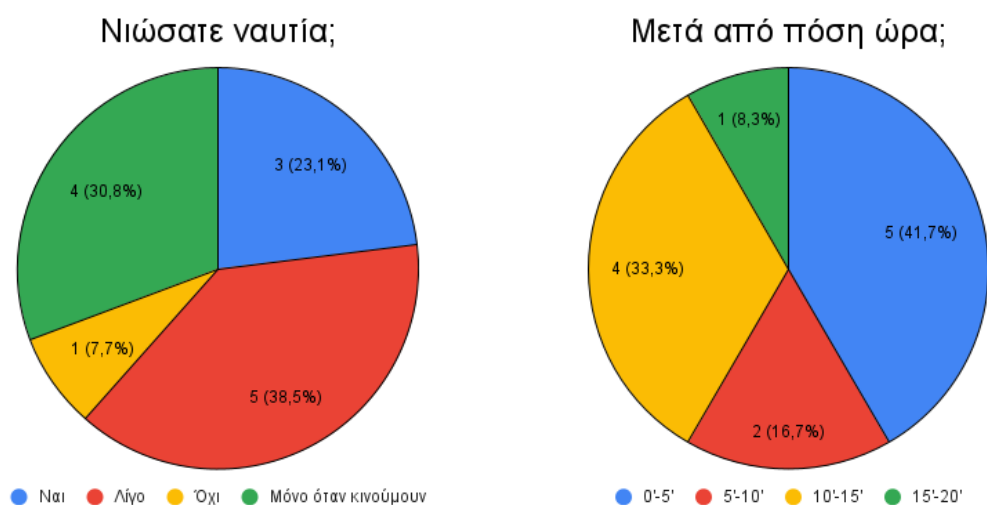
Σχήμα 5.2: Στοιχεία UI σχετικά με την εκμάθηση γλωσσών

Στη συνέχεια έχουμε παρατηρήσεις σχετικές με το διαλογικό σύστημα και το σενάριο το οποίο υλοποιήθηκε. Αν και το διαλογικό σύστημα ήταν φτωχό και η εκπαίδευση του Rasa μοντέλου ήταν κατά κάποιον τρόπο υποτυπώδης, τα αποτελέσματα σχετικά με τη συνοχή που παρουσίαζε το ενδιαφέρον που κρατούσε τον χρήστη ήταν ενθαρρυντικά.



Σχήμα 5.3: Αποτελέσματα σχετικά με το διαλογικό σύστημα

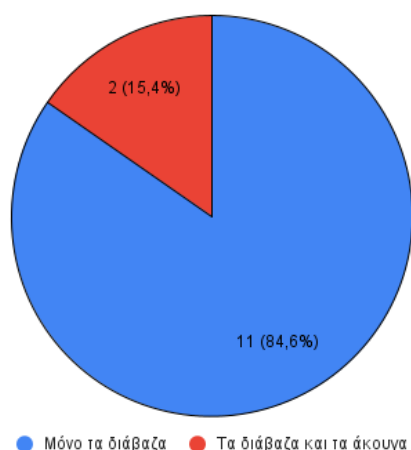
Τα επόμενα αποτελέσματα έχουν να κάνουν με το ζήτημα του motion sickness. Το πρόβλημα αυτό οφείλεται στο ότι πως η κίνηση έχει υλοποιηθεί και στο ότι υπάρχουν ακόμη πολλά περιθώρια βελτίωσης του VR εξοπλισμού. Πιθανές λύσεις σε αυτό το πρόβλημα μπορούν να είναι οι εξής δύο. Αφενός θα μπορούσε να υπάρχει η δυνατότητα τηλεμεταφοράς από σημείο σε σημείο για τους χρήστες που αντιμετωπίζουν έντονο πρόβλημα. Αφετέρου, εφόσον αυτό είναι δυνατόν, θα μπορούσε να υλοποιηθεί και κίνηση με τον βηματισμό του χρήστη, όπου η μετακίνηση του χρήστη στον πραγματικό κόσμο, προκαλεί αντίστοιχη μετακίνηση και στον εικονικό.



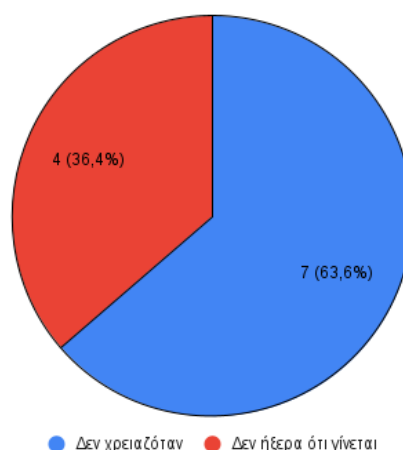
Σχήμα 5.4: Αποτελέσματα σχετικά με το motion sickness

Τέλος, συγκεντρώθηκαν και κάποια αποτελέσματα που αφορούν τη λειτουργία των hints. Πιο συγκεκριμένα, καταγράφηκε η προτίμηση των χρηστών για το περιεχόμενο τους κι αν αυτό θα ήταν καλύτερο να περιλαμβάνει προτάσεις με διαφορετικό νόημα ή τις ίδιες προτάσεις που έχουν παραφραστεί. Επίσης, έγιναν μετρήσεις για τη μορφή στην οποία τα hints χρησιμοποιήθηκαν και παρατηρήσαμε ότι αυτή ήταν κυρίως η γραπτή. Το συγκεκριμένο αποτέλεσμα πρέπει να φιλτραριστεί από το γεγονός ότι οι 12 από τους 13 χρήστες που δοκίμασαν την εφαρμογή είχαν τα ελληνικά ως μητρική γλώσσα.

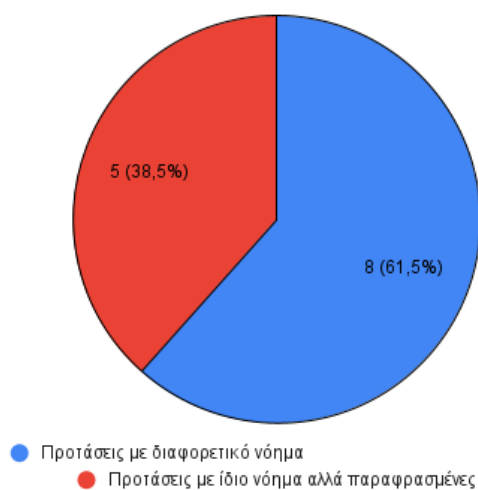
Πως χρησιμοποιήσατε τα hints;



Γιατί δεν ακούγατε τα hints;



Περιεχόμενο των hints



Σχήμα 5.5: Αποτελέσματα σχετικά με τη λειτουργία των hints

5.2 Αποτίμηση και μελλοντικά βήματα

Αποτίμηση

Όπως ειπώθηκε και στην προηγούμενη ενότητα, το προϊόν της παρούσας εργασίας είναι περισσότερο ένα prototype παρά μια ολοκληρωμένη εφαρμογή έτοιμη για χρήση από πραγματικούς χρήστες. Τι είναι όμως αυτό το οποίο επιτεύχθηκε και ποιο είναι το αποτέλεσμα της παρούσας υλοποίησης; Αυτό είναι η σχεδίαση της αρχιτεκτονικής μιας state-of-the-art VRALL εφαρμογής και μια πρώτη υλοποίηση του αντίστοιχου infrastructure. Η κύρια συνεισφορά, δηλαδή, που επιτελείται με τη συγκεκριμένη δουλειά είναι στο τεχνικό κομμάτι και αφορά την υποδομή που χρειάζεται για να συνδυαστούν οι τεχνολογίες αιχμής που χρησιμοποιούνται. Οι τεχνολογίες αυτές είναι συνοπτικά οι εξής:

- η ενσωμάτωση ενός διαλογικού μοντέλου που δε θα είναι rule-based, αλλά θα έχει τη δυνατότητα να αναπτυχθεί σε ένα εκτενές και δυναμικό σύστημα διαλογικής αλληλεπίδρασης, όπως είναι το **Rasa**,
- η υποστήριξη του μοντέλου αυτού από δύο ικανά και γενικού σκοπού εργαλεία αναγνώρισης και σύνθεσης φωνής, όπως είναι οι υπηρεσίες **ASR** και **TTS** της Google,
- η παραγωγή “έξυπνων προτάσεων” (**smart hints**) που θα κατευθύνουν τον χρήστη στην αλληλεπίδραση με το διαλογικό σύστημα και
- ο συνδυασμός όλων αυτών των εργαλείων σε ένα **immersive** πλαίσιο εφαρμογής **εικονικής πραγματικότητας**.

Αντίστοιχα, οι εργασίες που έπρεπε να γίνουν για να επιτευχθούν τα παραπάνω συνοψίζονται στις δύο κατηγορίες που ακολουθούν:

- Την προσαρμογή λειτουργίας του River - VA Backend, ώστε να παράγει, να επεξεργάζεται και να στέλνει με την κατάλληλη μορφοποίηση στον client του frontend του Unity τα έξυπνα hints που παίζουν δομικό ρόλο στον εκπαιδευτικό χαρακτήρα της εφαρμογής που σχεδιάστηκε. Για την προσαρμογή αυτή, έπρεπε:
 1. να γίνουν αλλαγές και προσθήκες στον Master_server και συγκεκριμένα στον τρόπο που διαχειρίζεται τα εισερχόμενα από τον Rasa_server μηνύματα,
 2. να στηθεί εκ νέου ένα κανάλι επικοινωνίας με τον Rasa_server, το meta_channel. Στο κανάλι αυτό ενσωματώθηκε όλη η διαδικασία παραγωγής κι αξιολόγησης των smart hints και
 3. να εκπαιδευτεί και να συνδεθεί με το υπόλοιπο Backend ένα λειτουργικό μοντέλο του Rasa που ανταποκρίνεται στο περιεχόμενο του σεναρίου στο ξενοδοχείο.
- Την υλοποίηση στο frontend της εφαρμογής του client ο οποίος αναλαμβάνει την επικοινωνία με το River, όπου επιτελούνται όλες οι λειτουργίες του διαλογικού συστήματος. Επίσης, στην πλευρά του εικονικού περιβάλλοντος έγινε η κατάλληλη διαμόρφωση όλων των στοιχείων διεπαφής με τον χρήστη, ώστε αφενός να είναι εύχρηστα και

Λειτουργικά για τους τελικούς χρήστες κι αφετέρου να ανταποκρίνονται στις προδιαγραφές λειτουργίας και συμβατότητας των επί μέρους υπηρεσιών του River (βλ. ASR, Rasa και TTS).

Μελλοντικά βήματα

Οι προσθήκες στην εφαρμογή και η επέκτασή της σε μια σειρά από θέματα φαίνονται ως απαραίτητα βήματα, ώστε να περάσει από το στάδιο του prototype που βρίσκεται τώρα σε αυτό της ολοκληρωμένης εφαρμογής που πετυχαίνει τους εκπαιδευτικούς στόχους που έχουν τεθεί. Πιο συγκεκριμένα, οι μελλοντικές αυτές κινήσεις για την περαιτέρω βελτίωση της εφαρμογής διακρίνονται σε δύο βασικές κατηγορίες. Η πρώτη έχει να κάνει με τη βελτίωση ήδη υπάρχουσών λειτουργιών και η δεύτερη με προσθήκες λειτουργιών και επεκτάσεις σε τομείς που δεν υπάρχουν στη παρούσα δουλειά.

Βελτιώσεις σε ήδη υπάρχουσες λειτουργίες

Καταρχήν, απαραίτητο βήμα στην επέκταση της εφαρμογής είναι ο εμπλουτισμός της με περισσότερα σενάρια που το κάθε ένα από αυτά θα σχετίζεται με διαφορετικό περιεχόμενο. Μάλιστα, όπως ισχύει στην περίπτωση της Mondly VR και άλλων εφαρμογών, τα σενάρια αυτά θα μπορούν να καλύπτουν ολόκληρες θεματικές ενότητες της καθημερινότητας. Εκτός από το θεματικό διαχωρισμό, κατηγοριοποίηση σε ενότητες μπορεί να γίνει και στο επίπεδο δυσκολίας των σεναρίων. Με αυτόν τον τρόπο, η εφαρμογή θα μπορούσε να προσφέρει μια πιο ολοκληρωμένη εικόνα στον χρήστη για τον τρόπο που χρησιμοποιείται η γλώσσα.

Μια άλλη επέκταση που θα έχει πολύ βοηθητικό χαρακτήρα είναι η δυνατότητα στον χρήστη να παρακολουθεί σε ποιο βαθμό έχει πετύχει τους επιμέρους στόχους του διαλόγου στον οποίο βρίσκεται με σκοπό να έχει καλύτερη εικόνα των πραγμάτων που έχει ήδη πει ή πρέπει ακόμη να πει. Αφού γίνει αυτό το βήμα, το επόμενο θα είναι η δυνατότητα να μετακινείται μεταξύ των διαφόρων σταδίων σε ένα διάλογο. Θα μπορεί, δηλαδή, να παρακάμψει κάποιο σημείο στο διάλογο ή να επιστρέψει σε κάποιο άλλο όποτε το επιθυμεί. Έτσι, θα εξασκεύεται σε ότι θεωρεί ότι χρειάζεται βελτίωση.

Τέλος, μια σημαντική βελτίωση που αναδείχθηκε ως ανάγκη μέσω των δοκιμών που έγιναν στο demo έχει να κάνει με το πώς πραγματοποιείται η αξιολόγηση των χρηστών. Πρέπει να χρησιμοποιηθούν μέθοδοι που με πιο σαφή και κατατοπιστικό τρόπο καταγράφουν την επίτευξη των στόχων και τη συνολική πρόοδο των χρηστών. Με την παρούσα υλοποίηση όλη αυτή η διαδικασία δεν έχει καμία τυπική μορφή και έχει αφεθεί εξ ολοκλήρου στη διακριτική ευχέρεια του χρήστη ή στον επιβλέποντα αυτού (π.χ. δάσκαλος) αν υπάρχει. Μια τέτοια προσθήκη θα μπορούσε να ανοίξει το δρόμο για τη χρήση της μεθόδου expressing (βλ. κεφάλαιο με τη θεωρία) για την εκπαιδευτική διαδικασία. Σύμφωνα με αυτήν τη μέθοδο, τα επιτεύγματα και οι επιδόσεις που πετυχαίνουν οι χρήστες στις δραστηριότητες της εφαρμογής, θα μπορούν να προβληθούν και στον πραγματικό κόσμο κι έτσι να ανανεώνουν το ενδιαφέρον τους για να βελτιώνονται συνεχώς.

Αποφάσεις στη δημιουργία μοντέλου Rasa και μελλοντικές βελτιώσεις

Όπως φαίνεται στη σχετική ενότητα του κεφαλαίου υλοποίησης, δεν κρίθηκε απαραίτητο να χρησιμοποιηθεί η λειτουργία των rules. Αυτό συνέβη παρόλο που κατά τη διαδικασία ανάπτυξης του διαλογικού δοκιμάστηκε η χρήση τους. Δημιουργήθηκαν, δηλαδή, διάφορα rules για μικρά κομμάτια διαλόγων που πρέπει ή συνηθίζουν να ακολουθούν το ίδιο μονοπάτι, αλλά εν τέλει αποφασίστηκε πως δεν ήταν αναγκαία, μιας και χωρίς την παρουσία τους τα αποτελέσματα ήταν ίδια. Βέβαια, αναγνωρίζεται πως μελλοντικά και με δεδομένο ότι το μοντέλο θα μεγαλώσει και τα ενδεχόμενα διαφορετικών μονοπατιών θα γίνουν αρκετά περισσότερα, η χρήση των rules θα έχει περισσότερο νόημα. Να σημειωθεί πως στο ίδιο μήκος κύματος με τη χρήση των rules μια ακόμη αλλαγή όσον αφορά τα stories του μοντέλου, που θα είναι σίγουρα χρήσιμη καθώς αυτό μεγαλώνει, είναι ο τεμαχισμός των μονοπατιών σε άλλα μικρότερα, ορίζοντας κατά κάποιον τρόπο και περισσότερες θεματικές.

Αντίστοιχα, αποφασίστηκε να μην χρησιμοποιηθούν ούτε τα actions. Ο λόγος και σε αυτήν την περίπτωση ήταν κυρίως το ότι το διαλογικό μοντέλο και το μικρό του μέγεθος δεν απαιτούσαν κάποια πιο εξειδικευμένη λειτουργία. Βέβαια, και σε αυτήν την περίπτωση υπάρχουν σημεία που με τη χρήση των actions θα βελτιώνονταν. Για παράδειγμα, στο κομμάτι του διαλόγου που ολοκληρώνεται μια κράτηση θα μπορούσε να προστεθεί μια φόρμα (form) που θα ζητάει τα δεδομένα από το χρήστη και θα τα αποθηκεύει κατευθείαν ως entities. Μάλιστα, αυτή η υλοποίηση θα μπορούσε να γίνει με τέτοιο τρόπο, ώστε ακόμη κι αν η διαδικασία συμπλήρωσης της φόρμας διακοπτόταν από κάποια άσχετη ερώτηση του χρήστη, η ροή του διαλόγου να επείστρεφε στο σημείο που είχε μείνει αμέσως μετά. Ένα άλλο παράδειγμα χρήσης actions θα μπορούσε να είναι η κλήση σε κάποιο API, για να ανακτηθούν χρήσιμες για το διάλογο πληροφορίες (π.χ. για την ενημέρωση της ώρα) και αυτές να αξιοποιηθούν ώστε να καθοριστεί η σωστή απάντηση του μοντέλου (π.χ. ενημέρωση για την τρέχουσα ώρα, ώστε να αποφασιστεί αν ο χαιρετισμός του χρήστη θα απαντηθεί με "καλημέρα" ή "καλησπέρα").

Ένα εργαλείο που επίσης δεν χρησιμοποιήθηκε, αλλά θα μπορούσε να είναι πολύ χρήσιμο, είναι τα slots. Τα slots είναι στην ουσία μεταβλητές που καθορίζοντας και ελέγχοντας την τιμή τους (π.χ. True / False) να μπορείς να πραγματοποιήσεις και ελέγχους στη ροή του διαλόγου επιτρέποντας έτσι τη δημιουργία παρακλαδιών στο αρχικό μονοπάτι.

Μια σημαντική βελτίωση που δεν έχει να κάνει τόσο με τα δεδομένα εκπαίδευσης, αλλά την ίδια την εκπαίδευση είναι η εξής. Ένας κίνδυνος που εγκυμονεί στα μοντέλα με πολλά δεδομένα εκπαίδευσης είναι αυτός του overfitting. Overfitting είναι η κατάσταση στην οποία ένα μοντέλο έχει εκπαιδευτεί με τέτοιο τρόπο που πλέον τα μονοπάτια που κανονικά θα έπρεπε να χρησιμοποιούνται ως παραδείγματα, καταλήγουν να αποτελούν τις μοναδικές αποδεκτές εξελίξεις ενός διαλόγου. Μάλιστα, αυτή η κατάσταση μπορεί να γίνει τόσο έντονη που τα λεγόμενα "χαρούμενα μονοπάτια" να ακολουθούνται ανεξάρτητα από αυτά που λέει ο χρήστης και τα intents που αναγνωρίζονται. Μια λύση σε αυτό, λοιπόν, είναι όσο αυξάνονται τα δεδομένα εκπαίδευσης, όσο δηλαδή μεγαλώνει το chatbot, να μειώνεται ο αριθμός των εποχών στη διαδικασία εκπαίδευσης των πολιτικών πρόβλεψης του επόμενου action.

Νέες προσθήκες ή επεκτάσεις της εφαρμογής

Όσον αφορά τα εντελώς καινούργια στοιχεία που μπορούν να προστεθούν στις λειτουργικότητες της εφαρμογής και να προωθήσουν τον εκπαιδευτικό σκοπό της υπάρχουν οι εξής ιδέες. Μπορούν να προστεθούν δραστηριότητες που πριν την εμπειρία κάθε σεναρίου, θα εισάγουν τον χρήστη στο αντίστοιχο λεξιλόγιο. Αυτές οι δραστηριότητες μπορούν είτε να αποτελούν μέρος της εμπειρίας στο περιβάλλον του συγκεκριμένου σεναρίου, είτε να είναι πλήρως ανεξάρτητες από αυτό και τη φάση επικοινωνίας με τους χαρακτήρες που περιέχει. Με αυτόν τον τρόπο, η χρήση της εφαρμογής θα μπορεί να αποκτήσει έναν πιο αυτόνομο χαρακτήρα και να μην εξαρτάται από κάποια ειδική προετοιμασία, όπως παραδείγματος χάριν είναι η εξοικείωση με το σχετικό με το σενάριο λεξιλόγιο στα πλαίσια μιας διδακτικής διαδικασίας.

Μια επέκταση που κρίνεται ήδη από πολλές σχετικές έρευνες ως πραγματικά ενδιαφέρουσα είναι η υποστήριξη ταυτόχρονης χρήσης της από περισσότερους του ενός χρήστες ταυτόχρονα (multiplayer mode). Ουσιαστικά, σε αυτή την περίπτωση χρήσης αντί να επικοινωνούν μόνο με έναν φανταστικό χαρακτήρα σε ρεαλιστικές καταστάσεις που προσομοιάζουν την πραγματική ζωή, οι χρήστες θα μπορούν να επικοινωνούν και με άλλους χρήστες που μαθαίνουν ελληνικά, δασκάλους ή και απλούς φυσικούς ομιλητές της γλώσσας.

Ένα ακόμη νέο στοιχείο είναι ο εμπλουτισμός της παρουσίας του χρήστη στον εικονικό κόσμο με πιο φυσικό τρόπο. Δηλαδή, η αλληλεπίδραση του χρήστη με το περιβάλλον του θα μπορούσε να γίνει πλουσιότερη, π.χ. μέσω της μετακίνησης αντικειμένων, της διαμόρφωση του χώρου γύρω του, της ανάγνωσης γραμμάτων ή μηνυμάτων ή μενού κλπ. Αυτό το είδος αλληλεπίδρασης αποτελεί ήδη αναπόσπαστο μέρος πολλών παιχνιδιών εικονικής πραγματικότητας και αξιολογείται ως πολύ θετικό στοιχείο εμπύθισης των χρηστών σε αυτά.

Όσον αφορά διάφορες ιδέες προς αναβάθμιση της εφαρμογής από τη μεριά του ASR κάποιες από αυτές είναι το να δοθεί η δυνατότητα να αναγνωρίζει περισσότερες από μια γλώσσες ταυτόχρονα (code-switching) ή να εντοπίζει ευκολότερα pronunciation errors ανάλογα με το προφίλ του χρήστη (π.χ. γλώσσα, διάλεκτο). Αν συνυπολογιστεί το γεγονός ότι ο σκοπός της εφαρμογής αυτής είναι η εκμάθηση γλωσσών, το code-switching μπορεί να αποδειχθεί πολύ βοηθητικό ειδικά στην περίπτωση των αρχάριων χρηστών.

Τέλος, μια ακόμη προσθήκη που θα μπορούσε να γίνει είναι η εξής. Ταυτόχρονα με τη διαλογική αλληλεπίδραση ανάμεσα στον χρήστη και τον χαρακτήρα του σεναρίου, θα μπορούσε να γίνεται και μια εκτίμηση της συναισθηματικής κατάστασης του χρήστη, μέσω της φωνητικής ανάλυσης της ομιλίας του. Αυτή η λειτουργικότητα είναι ακόμα σε πειραματικό στάδιο, αλλά είναι προφανές πως εφόσον ξεκινήσει να χρησιμοποιείται θα προσφέρει πολλές δυνατότητες. Αρκεί να φανταστούμε ένα διαλογικό σύστημα που όχι μόνο μπορεί και συνομιλεί, αλλά ταυτόχρονα αναγνωρίζει τα συναισθήματα του συνομιλητή του και προσαρμόζει τις απαντήσεις και το ύφος του σε αυτά.

Παραρτήματα

Ερωτηματολόγια

A'.1 Ερωτηματολόγιο χρήστη

Background

1. Age
2. Education (degree obtained or school level attended)
3. Experience / fluency in greek language
4. County of origin
5. Country of residence
6. If country of origin and country of residence are different, how long have you been in the country of your current residence?
7. What is your native language or languages?
8. Have you ever had any virtual reality experience before?
9. Have you ever used any other language learning app before and in what media? (ex. mobile, pc, VR, etc)
10. Is there anything that you feel is interesting or important about your background that you'd like us to know?

App related questions

1. Did you find the virtual environment immersive?
2. How would you evaluate the movement in the virtual environment (ex. smoothness, responsiveness, ease of use, etc)?
3. How would you evaluate the responsiveness of the hand controls (pointer)?
4. User Interface
 - (a) Was the menu with the instructions useful?

- (b) Did you find the dialogue UI useful?
- 5. Communication with the app
 - (a) Was it easy to read?
 - (b) Was it easy to hear the answer / hints?
 - (c) Did you find any other difficulty in communicating?
- 6. Did the app recognize what you said correctly?
- 7. Did you get tired or did you feel motion sickness while using the app? If yes, after how much time of usage did you feel that way?
- 8. Do you think that a “translation functionality” would be useful?
- 9. Hints
 - (a) Would you prefer the hints to be sentences with the same meaning but rephrased or sentences with different meaning?
 - (b) Would you prefer it if you could make the hints window appear or disappear whenever you want?
- 10. What other comments relevant to the virtual environment would you like to make?

Scenario related questions

- 1. Did you find the dialogic interaction interesting?
- 2. Did you find the flow of the dialogue consistent?
- 3. Did you make use of the hints functionality?
- 4. If yes, did you listen to the hint too or did you just read it?
- 5. If not, why didn't you? (ex. didn't need to, found it difficult to use it, wasn't aware of this functionality)
- 6. Do you have any comments about the duration of the scenario?
- 7. What other comments relevant to the scenario would you like to make?

Language learning related questions

- 1. Were you familiar with the scenario-specific vocabulary prior to the usage?
- 2. If yes, then did you find difficulty in using your knowledge?
- 3. If not, then did the scenario help you gain familiarity with the related vocabulary?
- 4. Do you think that this app could be useful as part of a language learning course?

Α.2 Ερωτηματολόγιο Υπευθύνου

1. What instructions did you give to the user prior to starting the app?
2. Did you have to intervene during the session and if yes, then in what way?
3. Did the user achieve the goals as they are set in the instructions menu?
 - Ξεκινήστε και ολοκληρώστε διάλογο με τη ρεσεψιόν.
 - Επιβεβαιώστε μια ήδη υπάρχουσα κράτηση στο ξενοδοχείο.
 - Κάντε μια νέα κράτηση στο ξενοδοχείο.
 - Ζητήστε πληροφορίες για το wifi, τη σάουνα κι άλλες παροχές του ξενοδοχείου.
4. How much time did the user use the app?
5. Was it without a pause or did the user have to restart the app? If yes, then why?
6. How many times did the user have a full conversation with the reception?

Βιβλιογραφία

- [1] Paul Milgram και Fumio Kishino. *A taxonomy of Mixed Reality Visual Displays*. *IEICE Transactions on Information and Systems*, E77(Δ):12, 1994.
- [2] Jonathan Steuer. *Defining Virtual Reality: Dimensions Determining Telepresence*. *Journal of Communication*, 42(4):73-93, 1993.
- [3] *virtual reality*. <https://www.britannica.com/technology/virtual-reality>. Ημερομηνία πρόσβασης: 20-9-2022.
- [4] Lourdes Miró Amarante Luis Muñoz-Saavedra και Manuel Domínguez-Morales. *Augmented and Virtual Reality Evolution and Future Tendency*. *Applied Sciences*, 10(1):322, 2020.
- [5] Cheng G. Huang X., Zou D. και Xie H. *A Systematic Review of AR and VR Enhanced Language Learning*. *Sustainability*, 13(9):4639, 2021.
- [6] Barbara L. Ludlow. *Virtual Reality: Emerging Applications and Future Directions*. *Rural Special Education Quarterly*, 34(3):3-10, 2015.
- [7] Mehrasa Alizadeh. *Virtual Reality in the Language Classroom: Theory and Practice*. *CALL-EJ*, 20(3):21-30, 2019.
- [8] Tomasz Mazuryk και Michael Gervautz. *Virtual Reality - History, Applications, Technology and Future*. 1999.
- [9] Serkan Yildirim Gürkan Yildirim, Mehmet Elban. *Analysis of Use of Virtual Reality Technologies in History Education: A Case Study*. *Asian Journal of Education and Training*, 4(2):62-69, 2018.
- [10] Veronica S. Pantelidis. *Reasons to Use Virtual Reality in Education and Training Courses and a Model to Determine When to Use Virtual Reality*. *THEMES IN SCIENCE AND TECHNOLOGY EDUCATION Special Issue*, σελίδες 59-70, 2009.
- [11] Carl Nätterdal Mustafa Hussein. *The Benefits of Using Virtual Reality in Education*. Πτυχιακή εργασία, University of Gothenburg, Chalmers University of Technology, Department of Computer Science and Engineering, 2015.
- [12] Sofia Mitsigkola. *Virtual Reality Language Learning*. Μεταπτυχιακή διπλωματική εργασία, Malmo University, 2022.

- [13] K. Schwienhorst. *Why Virtual, Why Environments? Implementing Virtual Reality Concepts in Computer-Assisted Language Learning*. *Simulation and Gaming*, 33(2):196–209, 2002.
- [14] E. Bialystok. *A Theoretical Model of Second Language Learning*. *Language Learning*, 28(1):69–83, 1978.
- [15] Haimingyue Xu. *A Future Direction for Virtual Reality Language Learning Using Dialogue Generation Systems*. *Journal of Education, Humanities and Social Sciences*, 2, 2022.
- [16] Y. J. Lan. *Immersion into virtual reality for language learning*. *Psychology of Learning and Motivation*, 72, 2020.
- [17] Wannaporn Chujitarom και Pallop Piriyasurawong. *STEAM-GAAR Field Learning Model to Enhance Grit*. *International Education Studies*, 11(11), 2018.
- [18] Xinjie Deng και Zhonggen Yu. *A Systematic Review on the Influence of Virtual Reality on Language Learning Outcomes*. *International Journal of Online Pedagogy and Course Design*, 12(83):18, 2022.
- [19] Michael L. Mauldin. *CHATTERBOTs, TINYMUDs, and the Turing Test Entering the Loebner Prize Competition*. *The Twelfth National Conference on Artificial Intelligence*, Washington State Convention and Trade Center Seattle, Washington, 1994.
- [20] Bayan Abu Shawar Eric Atwell. *Using dialogue corpora to train a chatbot*. *Proceedings of the Corpus Linguistics*, σελίδες 681–690, Lancaster University, 2003.
- [21] Imre Rudas Peter Toth. *Web-based Learning and Web Mining*. *The Asian Conference on Technology in the Classroom: The Impact of Innovation: Technology and You*, σελίδες 101–113, Lancaster University, 2013.
- [22] *Deep Learning for Chatbots, Part 1 - Introduction*. <https://dennybritz.com/posts/wildml/deep-learning-for-chatbots-part-1/>. Ημερομηνία πρόσβασης: 29-9-2022.
- [23] TURING A. M. *COMPUTING MACHINERY AND INTELLIGENCE*. *Mind*, 119(236):433–460, 1950.
- [24] Weizenbaum J. *ELIZA—a computer program for the study of natural language communication between man and machine*. *Communications of the ACM*, 9(1):36–45, 1966.
- [25] Weber S. Colby K. M. και Hilf F. D. *Artificial Paranoia*. *Artificial Intelligence*, 2(1):1–25, 1971.
- [26] R. S. Wallace. *The Anatomy of A.L.I.C.E. Parsing the Turing Test*, σελίδες 181–210, 2009.
- [27] Molnar G. και Szuts Z. *The Role of Chatbots in Formal Education*. *2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY)*, 2018.

- [28] Card W. I. και Lucas R. W. *Computer interrogation in medical practice*. *International Journal of Man-Machine Studies*, 14(1):49–57, 1981.
- [29] Rafael Lahoz-Beltra και Claudia Corona López. *LENNA (Learning Emotions Neural Network Assisted): An Empathic Chatbot Designed to Study the Simulation of Emotions in a Bot and Their Analysis in a Conversation*. *Social and Semantic Models, Tools and Applications in Science and Technology*, 2021.
- [30] G. Pirani. *Advanced Algorithms and Architectures for Speech Understanding*. Research Reports Esprit. Springer Berlin Heidelberg, 2013.
- [31] Alberto Ciaramella. *A prototype performance evaluation report*. *Sundial workpackage*, 8000, 1993.
- [32] Dan Jurafsky και James H. Martin. *Speech and language processing : an introduction to natural language processing, computational linguistics, and speech recognition*. Pearson Prentice Hall, Upper Saddle River, N.J., 2009.
- [33] Patil R. & Kalbande D. R. Dsouza R., Sahu S. *Chat with Bots Intelligently: A Critical Review & Analysis*. *2019 International Conference on Advances in Computing*. *International Conference on Advances in Computing, Communication and Control (ICAC3)*, San Sebastian, Spain, 2019.
- [34] AbejideAde Ibijola Chinedu WilfredOkonkwo. *Chatbots applications in education: A systematic review*. *Computers and Education: Artificial Intelligence*, 2, 2021.
- [35] & Romenti S. Murtarelli G., Gregory A. *A conversation-based perspective for shaping ethical human-machine interactions: The particular challenge of chatbots*. *Journal of Business Research*, 2020.
- [36] Krotov Batacharia B., Levy D. A. R. C. και Wilks Y. *CONVERSE: a conversational companion*. In Wilks Y. editor, *Machine conversations*, σελίδες 205–2015, 1999.
- [37] Chai et al. *Natural language sales assistant - a web-based dialog system for online sales*. In *Proceedings of thirteenth annual conference on innovative applications of artificial intelligence*, 2000.
- [38] Zadrozny et al. *Natural language dialogue for personalized interaction*. 2000.
- [39] Rollo Fryer, Luke & Carpenter. *Bots as language learning tools*. *Language Learning & Technology*, 10(3):8–14, 2006.
- [40] Dr. Richard S. WALLACE. *The Elements of AIML Style*. ALICE A. I. Foundation, Inc, 2003.
- [41] Eric Atwell Bayan Abu Shawar. *Chatbots: Are they Really Useful*. *Ldv Forum*, 2007.
- [42] Bayan Abu Shawar. *Integrating CALL Systems with Chatbots as Conversational Partners*. *Computación y Sistemas*, 2017.

- [43] Benner Dennis; Schöbel Sofia; Süess Christoph; Baechle Vera; και Janson Andreas. *Level-Up your Learning - Introducing a Framework for Gamified Educational Conversational Agents*. *Wirtschaftsinformatik 2022 Proceedings*, 2022.
- [44] Patrick R Lowenthal. *The evolution and influence of social presence theory on online learning*. *Social computing: Concepts, methodologies, tools, and applications*, σελίδες 113-128, 2010.
- [45] & Krathwohl D. R. Anderson L. W. *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. 2001.
- [46] Csikszentmihalyi M. *Beyond boredom and anxiety*. Jossey-Bass, 2000.
- [47] Panagiotis Panagiotidis. *Virtual Reality Applications and Language Learning*. *International Journal for Cross-Disciplinary Subjects in Education (IJCDSE)*, 12(2), 2021.
- [48] Reyes Sánchez Salvador. *Virtual Reality and foreign language learning: Analysis of the 6 most popular mobile apps from Google Play Store and App Store to learn English*. Πτυχιακή εργασία, Filología Francesa e Inglesa, Universidad de Cádiz, 2019.
- [49] *Review of Mondly VR Language Learning for Oculus 2*. <https://www.icls.edu/review-of-mondly-vr-language-learning-for-oculus-2/>, note = Ημερομηνία πρόσβασης: 12-10-2022.
- [50] *Unity (game engine)*. [https://en.wikipedia.org/wiki/Unity_\(game_engine\)](https://en.wikipedia.org/wiki/Unity_(game_engine)). Ημερομηνία πρόσβασης: 5-10-2022.
- [51] *AAA (video game industry)*. [https://en.wikipedia.org/wiki/AAA_\(video_game_industry\)](https://en.wikipedia.org/wiki/AAA_(video_game_industry)). Ημερομηνία πρόσβασης: 5-10-2022.
- [52] *MPEG-4 Face and Body Animation (MPEG-4 FBA) - An overview*. Visage Technologies AB, Tegskiftesgatan 291.
- [53] *Race condition wiki*. https://en.wikipedia.org/wiki/Race_condition. Ημερομηνία πρόσβασης: 8-10-2022.
- [54] *Deadlock wiki*. <https://en.wikipedia.org/wiki/Deadlock>. Ημερομηνία πρόσβασης: 8-10-2022.

Συντομογραφίες - Αρκτικόλεξα - Ακρωνύμια

| | |
|-------|--|
| VR | Virtual Reality |
| AR | Artificial Reality |
| ML | Machine Learning |
| NLU | Natural Language Understanding |
| NLG | Natural Language Generation |
| NLP | Natural Language Processing |
| TTS | Text-To-Speech |
| HMD | Head Mounted Display |
| CALL | Computer Assisted Language Learning |
| VRALL | Virtual Reality Assisted Language Learning |
| CA | Conversational Agent |
| CRT | Cathode-Ray Tube |
| LCD | Liquid-crystal display |
| TCN | Temporal Convolutional Networks |
| βλ. | βλέπε |
| π.χ. | παραδείγματος χάριν |
| κλπ | και τα λοιπά |
| κ.ά. | και άλλα |