

## Περίληψη κεφάλαιο 24

*Daniel Jurafsky & James H. Martin*

Αντωνακάκη Ελένη

AM: Lt1200002

Στο κεφάλαιο 24 του βιβλίου *“Speech and Language Processing”*, ο Jurafsky μελετά τον διάλογο μηχανής και χρήστη αντιστοιχίζοντας τον με τον ανθρώπινο διάλογο. Αρχικά, αναφέρεται στα χαρακτηριστικά και τις ιδιότητες του ανθρώπινου διαλόγου ο οποίος χαρακτηρίζεται, από τον ίδιο, αρκετά πολύπλοκος κυρίως μέσω στοιχείων όπως τα **endpoints**. Βέβαια, υπάρχουν στοιχεία όπως τα **adjacency pairs** (ερώτηση-απάντηση, πρόταση-αποδοχή ή απόρριψη κα.) ή **side-sequences** τα οποία βοηθούν το σύστημα να έχει μια ομαλή ροή στην συνομιλία του με τον χρήστη. Τα διαλογικά συστήματα είναι **user-initiative** ή **system-initiative** ανάλογα με το ποιος διαθέτει την πρωτοβουλία στον διάλογο.

Τα chatbots χαρακτηρίζονται ως τα πιο απλά συστήματα διαλόγου και κατηγοριοποιούνται σε **rule based** (ELIZA και PARRY) ή **corpus-based** συστήματα (corpora από crowdsourcing, Wikipedia κα). Τα τελευταία ονομάζονται επίσης και **συστήματα παραγωγής απόκρισης (response generation systems)**. Βέβαια, υπάρχουν και chatbots τα οποία συνδυάζουν rule-based και neural/corpus based αρχιτεκτονικές (Chirpy Cardinal system).

Έπειτα, γίνεται αναφορά στο **GUS**, ένα **simple frame-based dialogue system** το οποίο αποτελεί βάση για τους περισσότερους σύγχρονους εμπορικούς βοηθούς διαλόγου. Το GUS βασίζεται σε **frames** τα οποία αναπαριστούν τις προθέσεις τις οποίες ανιχνεύει το σύστημα από το χρήστη, και απαρτίζονται από μια σειρά από **slots**, κάθε ένα εκ των οποίων μπορεί να λάβει μια σειρά πιθανών τιμών (**values**). Αυτή είναι η έννοια του πλαισίου. Επομένως, το σύστημα στοχεύει στο να εξακριβώσει το **domain** στο οποίο αναφέρεται ο χρήστης. Συστήματα όπως αυτό στοχεύουν σε ενέργειες όπως κράτηση εισιτηρίου, εύρεση μιας ταινίας ή σημείωση ενός ραντεβού πράγμα που επιτυγχάνεται μέσω σημασιακών γραμματικών, **ASR (automatic speech recognition) component** και **natural language generation modules**.

Ένα διαλογικό σύστημα αποτελείται από τα παρακάτω μέρη: **automatic speech recognition, spoken language understanding, dialogue state tracker, dialog policy, natural language generation** και **text to speech synthesis**.

Στην συνέχεια, γίνεται λόγος για τις γλωσσικές πράξεις (**speech acts**) οι οποίες χρησιμοποιούνται από τα Dialogue-state συστήματα με στόχο την διαδραστική λειτουργία μιας ερωταπόκρισης ή μιας πρότασης.

Τα **supervised semantic parsing** στοχεύουν στη συμπλήρωση των slots, καθώς και στη κατάταξη της πρόθεσης του χρήστη βάσει ερωταποκρίσεων. Καταυτόν τον τρόπο, εκπαιδεύεται ένα μοντέλο στο οποίο δίνεται ως input η αναπαράσταση μιας πρότασης και αυτό με τη σειρά του προσπαθεί να την αντιστοιχίσει με τα κατάλληλα slots, domain και intents.

Ο ρόλος του **dialogue-state tracker** είναι να αναγνωρίζει τόσο την τρέχουσα κατάσταση της συμπλήρωσης των slots, όσο και την πιο πρόσφατη γλωσσική πράξη του χρήστη. Η διαδικασία του dialogue-act detection και της συμπλήρωσης των slots πραγματοποιούνται από κοινού στον ίδιο χρόνο.

Σημαντική είναι και η χρήση του **dialogue policy** το οποίο στοχεύει στο να αποφασίσει την κατάλληλη γλωσσική πράξη βάσει της φάσης στην οποία βρίσκεται ο διάλογος κάθε φάρα είτε μέσω ενός neural classifier είτε με reinforcement learning. Το confirmation και το rejection είναι οι δύο μέθοδοι με τις οποίες το σύστημα εξετάζει εάν έχει αντιληφθεί σωστά τον χρήστη. Ακόμα, η διαδικασία παραγωγής απάντησης σε μορφή κειμένου (**NLG**) υλοποιείται και αυτή σε δύο στάδια: προσδιορισμός περιεχομένου και προτασιακής δομής. Η διαδικασίες αυτές αναλύονται διεξοδικά στο εν λόγω κεφάλαιο.

Τέλος, γίνεται λόγος και στην αξιολόγηση των chatbots. Το **evaluation** δεν γίνεται αυτόματα, π.χ. μέσω BLUE ή ROUGE αλλά μέσω ανθρώπων που βαθμολογούν την συζήτηση τους με το σύστημα βάσει διαφόρων κριτηρίων (π.χ. ευφράδεια, αποφυγή επαναλήψεων κ.α.). Ωστόσο, όπως αναφέρεται από τον συγγραφέα γίνεται έρευνα για την ανάπτυξη εξειδικευμένων τεχνικών αυτόματης αξιολόγησης. Η αξιολόγηση των task-oriented συστημάτων μπορεί να γίνει είτε με απόλυτο τρόπο (έλεγχος εάν διεκπεραιώθηκε το task όταν αυτό είναι συγκεκριμένο) είτε μέσω μεθόδων που προσομοιάζουν με την μέτρηση της ικανοποίησης του χρήστη.

Σε κάθε περίπτωση, τονίζεται ότι ο σχεδιασμός των διαλογικών συστημάτων γίνεται με γνώμονα τον χρήστη (**ανθρωποκεντρική/ value sensitive design προσέγγιση**).