



Reprezentace stavu hlasových dialogových systémů

Filip Polák¹

1 Úvod

Cílem této práce je vytvoření softwarové knihovny v jazyku Python, která by zprostředkovala reprezentaci stavu hlasového dialogového systému bez ohledu na to, kde by byl systém použit. Stav hlasového dialogového systému představuje všechnu informaci potřebnou k úspěšnému pokračování dialogu, tj. zejména cíl uživatele a to, co bylo během konverzace řečeno a potvrzeno. Proto je reprezentace stavu jedním z hlavních problémů, který je třeba řešit při návrhu hlasových dialogových systémů.

2 Analýza knihovny

Hlavními požadavky kladenými na knihovnu byly modularita, dynamičnost a udržování historie dialogu. Modularita jak ve smyslu možnosti využití knihovny pro jiné domény, tak v různých platformách pro vývoj dialogových systémů. Dynamičnost ve smyslu umožnění přidávat nebo odebírat jednotlivé sloty a jejich hodnoty. Pro splnění požadavku na udržování historie byla pro reprezentaci historie, slotů a jejich hodnot využita softwarová knihovna Pysri-stent (Gustafsson (2021)), která nabízí možnost implementace immutable Python objektů.

3 Návrh řešení

Knihovna pro splnění požadavků představených v analýze knihovna obsahuje tři třídy:

- **Value** - reprezentuje hodnoty a jejich confidence score (důvěra systému)
- **Slot** - reprezentuje sloty a obaluje třídu **Value**
- **State** - reprezentuje kompozitní stav dialogu a obaluje třídy **Slot**

kde je každá z tříd přenositelná do jakékoliv domény, čímž je splněn požadavek na modularitu, a dynamičnost je zajištěna samotnou implementací třídy **State**, kde je možné přidávat a odebírat jednotlivé sloty libovolně během konverzace. Naplňování historie je umožněno využitím callback funkcí, které obsluhují budoucí vstup od uživatele a jimiž je samotná historie plněna. Historie pak obsahuje očekávání určité uživatelské odpovědi a také to, co se v dialogu reálně stalo a jaké sloty či hodnoty byly případně přidány nebo odebrány.

¹ student navazujícího studijního programu Aplikované vědy a informatika, obor Kybernetika, specializace Umělá inteligence a biokybernetika, e-mail: polakf@students.zcu.cz

```
def push(self, *args, **kwargs):
    expectations, self.expectations = self.expectations, v()
    if len(expectations) != 0: # condition so as not to have an empty spaces in
        history
        self.History = self.History.append(expectations)
        for cb, slot_names in expectations:
            cbkwargs = {key: kwargs.get(key) for key in slot_names}

            # Call the expectation callback and pass myself as first argument
            handled = cb(self, **cbkwargs)
            if handled:
                # The arguments were handled by the callback
                for key in slot_names:
                    del kwargs[key]
            if kwargs:
                # Only change state if there is something to push
                self.extend(**kwargs)

            if args and cb.__name__ == "confirm_unconfirmed":
                self.confirm_slots(*args)
            if args and cb.__name__ == "emptying_slots":
                self.empty_slots_fcn(*args)
```

Obrázek 1: Ukázka zdrojového kódu knihovny

4 Závěr

Softwarová knihovna byla otestována ve třech dialogích s dvěma doménami a na dvou dialogových platformách. První doménou byly pokerové karty, kde dialog řekl uživateli, jestli jím předložené dvě karty tvoří dohromady pár. Druhou doménou byl systém poskytující informace o vlakových spojení, kde dialog vyhledával vlakový spoj ze synteticky vytvořené databáze. Jeden dialog pro každou doménu byl vytvořen pro platformu SpeechCloud (Švec, (2021)), která operuje v češtině, a jeden dialog pro karetní doménu byl vytvořen pro platformu Mycroft (Lorensen (2021)), fungující v angličtině.

Knihovna ve všech případech pracovala správně, což ukazuje její versatilitu, co se týče platformy i jazyku. Největší výhodou implementace knihovny je možnost využít historie dialogů jako data k trénování neuronových sítí, které by na základě uživatelova vstupu mohly generovat nejlepší možnou reakci systému.

Literatura

- Švec, J. (2021) *Knihovna SpeechCloud.dialog*. Katedra kybernetiky ZČU. Interní dokument.
- Lorensen, A. et al. (2021) Mycroft. *Mycroft AI*. Available from: <https://github.com/MycroftAI> [Accessed 2nd February 2021]
- Gustafsson, T. et al. (2021) *pyrsistent*. Available from: <https://pypi.org/project/pyrsistent/> [Accessed 1st January 2021]