Hierarchical Dialogue Management
Francesca Giordaniello¹ (fg350@cam.ac.uk), Thomas Voice² (tvoice@apple.com), Milica Gašić¹ (mg436@cam.ac.uk)
¹Cambridge University Engineering Department, ²Apple

Spoken Dialogue Systems

Spoken Dialogue Systems (SDSs) offer an easy and intuitive way for the user-machine interaction. The user speech is interpreted through Spoken Language Understanding and mapped to an abstract representation $u_t$. The Dialogue Manager updates the belief state $b_t$ of the system and selects an action $a_t$ via a decision rule (policy) $\pi$, then converting the response into speech through Natural Language Generation (Figure 1).

Figure 1: Components of a Spoken Dialogue System.

Policy Optimisation

At each turn, the policy chooses the action that maximises the expected cumulative reward $Q$:

$$\pi(b) = \arg\max_a \{Q(b, a) : a \in A\}$$

(1)

The GPSARSA algorithm is used, which models the Q-function as a Gaussian Process (GP):

$$Q(b, a) \sim GP(m(b, a), k((b, a), (b, a)))$$

(2)

Bayesian Committee Machine

The Bayesian Committee Machine (BCM) approach combines estimators trained on different datasets [1] (Figure 2), such as multiple estimates of the policy from different domains [2]. In general, it guarantees higher performance with respect to the correspondent in-domain policy (Figure 3).

Figure 2: Configuration of the BCM.

Figure 3: Performance of in-domain and BCM-based policies.

Hierarchical BCM

The same approach can be performed in a hierarchical fashion (Figure 4) by specifying a BCM for each of the $n$ subsets of $M$ domains, which compose the upper-level committee [3].

$$Q_n(b, a) = \sum_{i=1}^{M} \sum_{j=1}^{M} Q_{n,i}(b, a) - 1 Q_{n,i}(b, a)$$

$$\sum_{i=1}^{M} Q_{n,i}(b, a) - 1 = (1 - M) \cdot k((b, a), (b, a)) - 1 + \sum_{i=1}^{M} \sum_{j=1}^{M} Q_{n,i}(b, a) - 1$$

(3)

Figure 4: Configuration of the HBCM.

Preliminary Results

The hierarchical configuration allows more efficient scaling and guarantees a simpler parallelisation comparing to the BCM, especially when a larger number of domain is used. (Figure 5).

Figure 5: Comparison of the BCM and HBCM for a set of six domains.

Further Experiments

The generalisation capability of the policy optimisation algorithm in a HBCM setup could be highlighted by:

- using a larger domain database
- exploring different hierarchical configurations
- evaluating the performance of a policy on an unseen domain

References