Optimising Incremental Generation for Information Presentation of Mobile Search Results

Nina Dethlefs, Helen Hastie, Verena Rieser, Oliver Lemon
Heriot-Watt University
Riccarton
Edinburgh, UK EH14 4AS
n.s.dethlefs, h.hastie, v.t.rieser, o.lemon@hw.ac.uk

1 Introduction

This abstract discusses a proof-of-concept study in incremental Natural Language Generation (NLG) in the domain of Information Presentation for Spoken Dialogue Systems. The work presented is part of the FP7 EC Parlance project (http://www.parlance-project.eu). The goal of Parlance is to develop personalised, mobile, interactive, hyper-local search through speech. Recent trends in Information Retrieval are towards incremental, interactive search and we argue that spoken dialogue systems can provide a truly natural medium for this type of interactive search. This is particularly attractive for people on the move, who have their hands and eyes busy.

Timely and *relevant* presentation of search results is therefore key to the adoption of speech-driven, mobile interfaces as is natural language interaction. We discuss an *incremental* NLG component for Information Presentation (IP), where a reinforcement learning agent faces the trade-off of whether to present information as soon as it is available (for high reactiveness) or else to wait until input hypotheses of user input, coming through Spoken Language Understanding, are more stable (to avoid self-confusions). Results show that the agent learns to avoid long waiting times and self-corrections, sometimes by re-ordering pieces of information based on their confidence. The agent outperforms a number of hand-crafted baselines providing evidence for the feasibility of formulating incremental generation as an optimisation problem.

2 Incremental Dialogue and Generation

NLG modules in interactive systems are typically triggered at the end of a processing cycle when the user's utterance has been fully interpreted. This strict-turn taking model does not match human incremental processing though, where users can process partial input, and plan partial utterances in parallel. Recent work on incremental systems has shown that processing smaller 'chunks' of user input can improve the user experience (Skantze and Schlangen, 2009; Buss et al., 2010; Skantze and Hjalmarsson, 2010; Baumann et al., 2011) by showing higher reactiveness and more natural behaviour.

In addition, incrementality in NLG systems enables the system designer to model several dialogue phenomena that play a vital role in human discourse (Levelt, 1989) but have so far been absent from NLG systems. These include more natural turn-taking through rapid system responses, grounding through the generation of backchannels and feedback, and barge-ins (from both user and system). Furthermore, corrections and self-corrections through constant monitoring of user and system utterances play an important role, enabling the system to recover smoothly from a recognition error or a change in the user's goals (such as first asking for French restaurants, but then changing their preference to Mexican ones). Some examples of the phenomena we are targeting are given in Figure 1.

2.1 Optimising NLG Decision Making

A separate recent line of research investigates new statistical approaches for optimising NLG decisions, e.g. Rieser et al. (2010) or Dethlefs and Cuayáhuitl

(2011). Based on the success of both directions, incrementality and optimisation of NLG decision making, we argue for a combination of the two approaches by implementing the *update*, *revoke* and *purge* operations typically used in incremental approaches (Schlangen and Skantze, 2009) as state transitions in a Markov Decision Process. This allows us to move away from conventional strict turn taking and present information incrementally. At the same time, we make a step forward in incremental processing systems, which have so far relied on deterministic systems without automatic optimisation. We design a model of incremental NLG that generates output based on micro-turn interpretations of the user's utterances and is able to optimise its decisions using statistical machine learning. Figure 2 shows the state representation, the action set and reward function of our learning agent. Please refer to (Dethlefs et al., 2012) for details on this approach.

2.2 Summary of Findings

In a proof-of-concept study in the domain of mobile Information Presentation, we optimised the *timing and order* of Information Presentation. The learning agent optimised the trade-off of whether to present information as soon as it becomes available (for high responsiveness) or else to wait until input hypotheses are more confident (to avoid self-corrections). Results in a simulated environment showed that the agent learnt to tackle the problem by re-ordering information according to its confidence. Figure 1 shows examples of the agent's learnt action policy, where it avoids self-corrections and long waiting times, often by presenting information in order of decreasing confidence. It outperforms three hand-crafted baselines due to its enhanced adaptivity. See Figure 3 for a description of these baselines and example dialogues.

Backchannels (when the user pauses)

USR I want Italian food [500 ms] in the centre of town . . .
SYS uh-huh
SYS OK. I found 24 Italian restaurants in the city centre. The restaurant *Roma* is in the medium price range, but does not have great food. The restaurants *Firenze* and *Verona* . . .

Self-correction (the system made a mistake)

USR I want Italian food in the centre of town . . .
SYS OK. I found 35 Indian restaurants . . .
USR No, I want Italian.
SYS oh sorry . . .
SYS I have 24 Italian restaurants in the city centre . . .

Correction (the user changed their mind)

USR I want Indian food in the centre of town . . .
SYS There are 35 Indian . . .
USR Oh sorry, I meant Italian.
SYS OK, Italian . . .
SYS I have 24 Italian restaurants . . .

Re-ordering (from high to low confidence)

USR I want Italian food . . .
SYS [waits]
USR in the city centre.
SYS I have 120 places in the city centre . . .
USR I love Italian . . .
SYS 24 of them Italian . . .

Holding the floor

USR I want cheap Italian food . . .
SYS Ok, let me see
SYS I have 3 cheap Italian places . . .

Figure 1: Example discourse phenomena generated with the trained policy that become possible through incremental NLG.

3 Influencing User's Decision Making through Incremental, Optimised NLG

Automatic optimisations of NLG decision making have been shown to lead to systems that are more robust to unseen circumstances, more adaptable to new situation and users and more flexible in terms of its interactive NLG behaviour than non-optimised systems (Rieser et al., 2010; Janarthanam and Lemon, 2010; Dethlefs et al., 2011). Similarly, incremental processing system have been shown to be perceived as more reactive and natural by human users than their non-incremental counterparts (Skantze and Schlangen, 2009; Buss et al., 2010; Skantze and Hjalmarsson, 2010; Baumann et al., 2011). We believe that a synergy between both approaches can lead to information being presented in a more reliable, efficient and natural way and will therefore help the user to make decisions faster and build trust in the system's recommendations. We will test this hypothesis in future work by conducting a task-based evaluation study with real users.

States
incrementalStatus {0=none,1=holdFloor,2=correct,3=selfCorrect}
presStrategy {0=unfilled,1=filled}
statusCuisine {0=unfilled,1=low,2=medium,3=high,4=realised}
statusFood {0=unfilled,1=low,2=medium,3=high,4=realised}
statusLocation {0=unfilled,1=low,2=medium,3=high,4=realised}
statusPrice {0=unfilled,1=low,2=medium,3=high,4=realised}
statusService {0=unfilled,1=low,2=medium,3=high,4=realised}
userReaction {0=none,1=select,2=askMore,3=other}
userSilence={0=false,1=true}
Actions
IP: compare, recommend, summarise, summariseCompare, summariseRecommend, summariseCompareRecommend,
Slot-ordering: presentCuisine, presentFood, presentLocation, presentPrice, presentService,
Incremental: backchannel, correct, selfCorrect, holdFloor, waitMore
Rewards
$+100$ if the user selects an item
$+0$ if the user adds more constraints to the search
-100 if the user does something else
-0.5 for holding the turn
$-waiting\_time^2$ for waiting
-1 otherwise

Figure 2: The state and action space and reward function of the learning agent that define its learning environment. The goal state is reached when all items (that the user may be interested in) have been presented.

4 Future Directions

Future work can take several directions. Currently we learn the agent's behaviour offline, before the interaction, and then execute it statistically. More adaptivity towards individual users and situations could be achieved if the agent was able to learn from ongoing interactions by estimating user types online (Janarthanam and Lemon, 2010; Cuayáhuitl and Dethlefs, 2011). Current NLG systems tend to assume that the user's goals and situational circumstances are known with certainty. This is often an unrealistic assumption which future work could address using POMDPs (Williams and Young, 2007). Finally, using flat Reinforcement Learning, we cannot address complex problems with large state-action spaces so that future work may apply *hierarchical reinforcement learning* (Barto and Mahadevan, 2003; Cuayáhuitl, 2009; Dethlefs and Cuayáhuitl, 2011) to incremental NLG.

Baseline 1 (present information as soon as it becomes available)
USR Please show me Italian places
($cuisine = Indian$;$conf\_score = 0.2$)
SYS I have found 35 Indian ...
USR No, *ITALIAN*
($cuisine = Italian$; $conf\_score = 0.8$)
SYS Oh, you meant Italian. I have 104 Italian places in Edinburgh. 24 in the city centre, 21 in the New Town, 14 in ...
USR Show me some in the city centre
($location = centre$; $conf\_score = 0.6$)
SYS OK. I found 24 Italian restaurants in the city centre. The restaurant *Roma* is in the medium price range, but does not have great food. The restaurants *Firenze* and *Verona* ...

Baseline 2 (always wait until confidence is high)
USR Do you have Italian restaurants in the centre of town?
$cuisine = Italian$; $conf\_score = 0.4$
$location = centre$; $conf\_score = 0.2$
SYS *waits*
USR Italian in the centre.
$cuisine = Italian, conf\_score = 0.7$
$location = centre, conf\_score = 0.5$
SYS I have 104 Italian restaurants in the city.
USR *waits*
SYS *waits*
USR city centre please
$location = centre, conf\_score = 0.7$
SYS I have 24 Italian restaurants in the city centre ...

Baseline 3 (present information in the following order: cuisine, food, location, price, service)
USR I want Italian food in the centre of town.
$cuisine = Indian, conf\_score = 0.1$
$location = centre, conf\_score = 0.9$
SYS There are 189 Indian restaurants ...
USR No, Italian please. In the city centre.
$cuisine = Japanese, conf\_score = 0.2$
$location = centre, conf\_score = 0.9$
SYS OK, there are 76 Japanese restaurants. ...
USR *ITALIAN. CITY CENTRE.*
$cuisine = Italian, conf\_score = 0.7$
$location = centre, conf\_score = 0.9$
SYS I have 24 Italian restaurants in the city centre ...

Figure 3: Example dialogues generated with the baseline policies for a user who wants Italian food in the city centre. Confidence scores for cuisine and location variables for the restaurants are shown as updated.

5 Acknowledgements

The research leading to this work has received funding from EC's FP7 programmes: (FP7/2011-14) under grant agreement no. 287615 (PARLANCE); (FP7/2007-13) under grant agreement no. 216594

(CLASSiC); (FP7/2011-14) under grant agreement no. 270019 (SPACEBOOK); (FP7/2011-16) under grant agreement no. 269427 (STAC).

References

Andrew G. Barto and Sridhar Mahadevan. 2003. Recent Advances in Hierarchical Reinforcement Learning. *Discrete Event Dynamic Systems*, 13.

Timo Baumann, Okko Buss, and David Schlangen. 2011. Evaluation and Optimisation of Incremental Processors. *Dialogue and Discourse*, 2(1).

Okko Buss, Timo Baumann, and David Schlangen. 2010. Collaborating on Utterances with a Spoken Dialogue Systen Using an ISU-based Approach to Incremental Dialogue Management. In *Proc. of 11th Annual SIGdial Meeting on Discourse and Dialogue*.

Heriberto Cuayáhuitl and Nina Dethlefs. 2011. Spatially-aware Dialogue Control Using Hierarchical Reinforcement Learning. *ACM Transactions on Speech and Language Processing (Special Issue on Machine Learning for Robust and Adaptive Spoken Dialogue System)*, 7(3).

Heriberto Cuayáhuitl. 2009. *Hierarchical Reinforcement Learning for Spoken Dialogue Systems*. PhD Thesis, University of Edinburgh, School of Informatics.

Nina Dethlefs and Heriberto Cuayáhuitl. 2011. Combining Hierarchical Reinforcement Learning and Bayesian Networks for Natural Language Generation in Situated Dialogue. In *Proc. of the 13th European Workshop on Natural Language Generation (ENLG)*, Nancy, France.

Nina Dethlefs, Heriberto Cuayáhuitl, and Jette Viethen. 2011. Optimising Natural Language Generation Decision Making for Situated Dialogue. In *Proceedings of the 12th Annual Meeting on Discourse and Dialogue (SIGdial)*, Portland, Oregon, USA.

Nina Dethlefs, Helen Hastie, Verena Rieser, and Oliver Lemon. 2012. Optimising Incremental Generation for Spoken Dialogue Systems: Reducing the Need for Fillers by Optimising Waiting Time and Content Reordering. In *Proc. of INLG*, Chicago, IL, USA.

Srini Janarthanam and Oliver Lemon. 2010. Learning to Adapt to Unknown Users: Referring Expression Generation in Spoken Dialogue Systems. In *Proc. of the 48th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 69–78, July.

Willem Levelt. 1989. *Speaking: From Intenion to Articulation*. MIT Press.

Verena Rieser, Oliver Lemon, and Xingkun Liu. 2010. Optimising Information Presentation for Spoken Dialogue Systems. In *Proc. of the 48th Annual Meeting of the Association for Computational Linguistics (ACL)*, Uppsala, Sweden.

David Schlangen and Gabriel Skantze. 2009. A General, Abstract Model of Incremental Dialogue Processing. In *Proc. of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, Athens, Greece.

Gabriel Skantze and Anna Hjalmarsson. 2010. Towards Incremental Speech Generation in Dialogue Systems. In *Proc. of the 11th Annual SigDial Meeting on Discourse and Dialogue*, Tokyo, Japan.

Gabriel Skantze and David Schlangen. 2009. Incremental Dialogue Processing in a Micro-Domain. In *Proc. of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, Athens, Greece.

Jason Williams and Steve Young. 2007. Partially Observable Markov Decision Processes for Spoken Dialog Systems. *Computer Speech and Language*, 21(2):393–422.