Natural Language Generation enhances human decision-making with uncertain information

Dimitra Gkatzia, Oliver Lemon and Verena Rieser
Contact: D.Gkatzia@napier.ac.uk
INTRODUCTION

• This work explores the effect that different information presentation strategies have on human decision making

• Decision-making is often dependent on uncertain data

• e.g. There is 70% chance of rain

• We investigate whether NLG approaches can assist in decision making when the input to the generation system is uncertain data

• The Extended Weather Game (Gkatzia et al., 2015) is used for task-based evaluation
BACKGROUND

- **Data-to-text** generation is concerned with the task of generating text from non-linguistic data.

- Previous work has shown that regarding decision making, NLG can lead to comparable results to commonly used data visualisation techniques (Gatt et al., 2009).

- Current data-to-text approaches assume that data is precise and correct, an assumption which is unrealistic in real-world applications.
MOTIVATION

• Presenting numerical expressions of risk and uncertainty is not enough, as a high percentage of people do not understand and cannot act upon numerical uncertainty (e.g. Cokely et al., 2012).

• “Which of the following numbers represents the biggest risk of getting a disease: 1 in 100, 1 in 1000, 1 in 10?”

• Only 28% of Germans and 25% of Americans are able to answer this question (Galesic and Garcia-Retamero, 2010)

• Therefore, we explore how to convert numerical representations of uncertainty into Natural Language so as to maximise confidence and effective outcomes of human decision-making
1. Graphics only (Stephens et al., 2011)

2. NLG only
   - **WMO-based**: Rule-based, uses the guidelines recommended by the World Meteorological Organisation (WMO) for reporting uncertainty, e.g. \( p > 0.99 \) -> “extremely likely”
   - **NATURAL**: Rule-based, imitates media forecasters, e.g. \( p = 0.3 \) -> “mainly dry”

3. Multi-modal
   - WMO-based + graphics
   - NATURAL + graphics

Example graph from (Stephens et al., 2011)
THE EXTENDED WEATHER GAME

• Ice cream seller scenario

1. Given the forecasts of two locations on rainfall and temperature levels for four weeks, the player has to suggest where to send the ice-cream vendor

2. The participants had to declare how confident they were with their decisions

3. Based on their decisions and confidence they were presented with a projected “monetary gain”

• The probability of whether rain occurred is estimated by sampling the probability distribution
EVALUATION SETUP

• 442 unique players via social media
  • 197 females
  • 241 males
  • 4 non-disclosed
• 450 unique games (a few participants played more than once)
• We analysed both the players’ scores (in terms of monetary gain) and their confidence levels
## RESULTS

| Presentation  | Monetary gains (score) | Confidence |
|---------------|------------------------|------------|
| Graphs only   | 81.15                  | 78.5%      |
| NLG only      | 101.33                 | 66%        |
| Multi-modal   | 117.15                 | 83.7%      |
DEMOGRAPHIC FACTORS

• Predictors of the players’ understanding of uncertainty:
  - prior experience on making decisions based on risk
  - familiarity with weather models

• Factors which are not relevant in terms of understanding uncertainty:
  - education level
  - gender
  - English native (or not)
GENDER-BASED RESULTS

• Female participants:
  - females score significantly higher at the game when exposed to either of the NLG output presentations, when compared to the graphics-only presentation
  - the multi-modal presentation adds little more in effectiveness of decision-making than the NLG-only condition, but the multi-modal presentations do enhance their confidence (+15%)

• Male participants:
  - males score similarly for all types of representations
  - however, males feel significantly more confident when exposed to multi-modal and graphs only rather than NLG only
CONCLUSIONS

• Confidence does not necessarily correlate with effective decision making

• Visualisation techniques enhance participants’ confidence, however they don’t enhance decision making

• Multi-modal representation of data can result in effective decision making and also contribute to participants’ confidence.

• Gender, prior experience with uncertain data and risk taking are predictors of high score in this task
THANK YOU!
• Dimitra Gkatzia, Amanda Cercas Curry, Verena Rieser, and Oliver Lemon. 2015. A game-based setup for data collection and task-based evaluation of uncertain information presentation. In Proceedings of the 15th European Workshop on Natural Language Generation (ENLG).

• Albert Gatt, Francois Portet, Ehud Reiter, James Hunter, Saad Mahamood, Wendy Moncur, and Somayajulu Sripada. 2009. From Data to Text in the Neonatal Intensive Care Unit: Using NLG Technology for Decision Support and Information Management. AI Communications, 22: 153-186.

• Edward T. Cokely, Mirta Galesic, Eric Schulz, Saima Ghazal, and Rocio Garcia-Retamero. 2012. Measuring risk literacy: The berlin numeracy test. Judgment and Decision Making, 7(1):25–47.

• Mirta Galesic and Rocio Garcia-Retamero. 2010. Statistical numeracy for health: A cross-cultural comparison with probabilistic national samples. Archives of Internal Medicine, 170(462–468).

• Liz Stephens, Ken Mylne, and David Spiegelhalter. 2011. Using an online game to evaluate effective methods of communicating ensemble model output to different audiences. In American Geophysical Union, Fall Meeting.