Project proposals

Michael Franke

last update: July 11, 2018

1 Alternative-dependent interpretation of probability expressions

Recent linguistic theories of verbal probability expressions like *probably*, *certainly* or *possibly* have postulated that their meaning is given by a threshold-semantics (e.g. Yalcin, 2010). For example, *It will probably rain* is true iff the probability of rain according to the speaker is higher than some contextually specified θ_{probably}. A few studies have looked at the different interpretations of probability expressions (e.g. Windschitl and Wells, 1996), i.e., trying to determine where exactly a threshold like θ_{probably} should be. One case has gained some recent coverage in the online media:

- `https://www.vedrashko.com/thisisbig/how-likely-is-probably`
- `https://github.com/zonination/perceptions`

Building on the latter, it would be interesting to see if/how people's judgements depend on the alternative expressions they encounter in the environment. For example, recent studies from psycholinguistics have shown that which alternative descriptions are present in an experiment can influence the interpretation of certain words (e.g. Degen and Tanenhaus, 2015; Franke, 2016). This project would experimentally investigate the interpretation of various probability expressions and manipulate the alternative expressions offered to different subjects. This would help to come to understand if the thresholds postulated in the literature are better thought of as pragmatic or semantic.

[experimental, non-technical, course project, BSc thesis]

2 Pragmatic reasoning about conditionals

Devise, implement and test a model of the use and interpretation of conditional sentences in the style of the Rational Speech Act modelling approach.

[modeling, technical, linguistics, logic, MSc thesis]

3 Measure reaction times for reference games; diffusion drift models

Reference games are the motivating experimental paradigm for the Rational Speech Act model (Goodman and Frank, 2016). Many replications exist, but most attention has focused on accounting for choice probabilities. Much less attention has been paid to the process of decision making. This project would try to shed light on how fast participants come to form a decision. After experimental measurement of choice reaction times we would look at a statistical model that jointly predicts choice probabilities and reaction times, like the diffusion drift model (e.g. Ratcliff and McKoon, 2008) or a form of ACT-R.

[experimental, modeling, technical, BSc or MSc thesis]

4 Combining mouse- and eyetracking

Replicate the study of Roettger and Franke (2017) which used mousetracking, but also record participants eye-gazes as well.

[experimental (difficult) in the lab, MSc thesis]

5 Modeling politeness in language use

Building on recent models of polite language use (Tessler & Yoon), this project would try to incorporate the dimensions of politeness outlined in the seminal work by Brown & Levinson, namely *face threats* and *rate of imposition*, into a Rational Speech Act model of polite language use.

[modeling, linguistics, sociolinguistics, BSc or MSc thesis]

6 Reference games with topic choice

Reference games are a widely studied experimental paradigm. In the production part of these experiments, participants are told to select a description of a referent which is highlighted as the topic of conversation in each experimental trial. This project collects data on a related event, namely where the speaker is allowed to freely choose which referent to talk about first, and will then produce a description to communicate to the listener. The hypothesis is that this topic choice is affected by factors such as visual salience, but might also lead to a strategic choice of topic, namely that participants choose expressions for which they have an efficient utterance to communicate it. Finally, it is interesting to compare choices of topics with the results from listener prior elicitation conditions.

[experimental (easy), project or BSc thesis]

7 Evolution of semantic meaning via pragmatic language use

Many (agent-based) models of the evolution of semantic meaning do not have a particularly sophisticated vision of how agents use and interpret language when they interact. In response, Brochhagen, Franke, and van Rooij (2016); Brochhagen and Franke (2017) have used the Rational Speech Act model to look at simulations addressing the question: which semantic meaning would evolve if agents used language as described by the Rational Speech Act model. In this project, this approach could be applied, for example, to studying the evolution of meaning of gradable adjective, color terms, spatial expressions or a similar interesting domain. This could also use reference games with obects from a multi-dimensional feature space.

[simulation, programming, modeling, language evolution, project, BSc or MSc thesis]

8 Web-app for (non-)cooperative MasterMind

We program a web-application that has a little AI and plays a variant of mastermind with our subjects (taking inspiration from Verbrugge and Mol, 2008). Subjects see, say, 3 colored objects from a pool of 3 shapes in 3 colors, no repetition possible, but order matters, so we have 504 possible "world states". The AI needs to guess the world state. Subjects give feedback on each guess of the AI. Feedback is given by constructing a sentence from a reasonable number of alternative expressions at different positions of the sentence, e.g.:

- quantifier: at least one of the, several of the, some of the, all of the, ... (some theoretically interesting ones)
- object: squares, blue squares, triangles, green circles, ... (all nine possibilities)
- predicate: is/are in the right place, has/have the right color, ...

This can be realized with simple dropdown menus. Feedback must be true (easy and fast to check with a precomputed lookup table). In the cooperative version subjects win if the AI correctly guesses the sequence after, say, 7 rounds. In the uncooperative version subjects win if the AI fails to guess the correct sequence for, say, 10 rounds (the precise numbers need to be chosen so that the game is fun to play, not too easy not too hard, which depends on how smart the AI is). We can measure the informativity of a feedback description quite simply in the program in terms of the number of contextually live possibilities ruled out by the given description at the current state of play. Cooperative language users should tend to choose more informative messages; uncooperative language users should ideally send maximally uninformative, yet true messages. It would be interesting to look at (i) differences in language use between cooperative and uncooperative plays, and possibly (ii) the amount of uncooperativity in choices (measured quantitatively (!) in terms of contextual informativity) as they develop over time.

[programming, web-app, group project or MSc thesis]

9 Inference of speaker competence

This project is about how different expressions might lead to different inferences about how knowledgeable the speaker is. For example, hearing Jones say *I ate all of the cake* will not usually have us believe that Jones is in any way uncertain about the amount of cake he ate. This gets more interesting when we come to comparing expressions like *up to 10* to *no more then 10* and *close to 10 but not more* etc. This project would try to collect data using web-based experiments.

[experimental, single project, possibly BSc thesis]

10 Argumentative language use

Much linguistic theories on language use assume that the speaker wants to ideally inform the listener, i.e., transfer a maximum of relevant and true information for the sake of the listener's knowledge. However, other forms of discourse exist. An example is *argumentative discourse* where interlocutors hold different opinions about some topic and give arguments in order to persuade one another of their own view (e.g. Anscombre and Ducrot, 1983; Merin, 1999; Glazer and Rubinstein, 2006). This project has two parts which can be independently executed or done in conjunction:

- In the **experimental part**, we implement an experiment in which a speaker chooses expressions to influence the hearer of a particular conclusion, e.g., report on the exam results of a class of students (using similar stimuli as Cummins, 2012).
- In the **modeling part**, we extend a model of goal-directed choice of expressions to incorporate a notion of argument strength (e.g., the notion of argumentative relevance of Merin, 1999).

[experimental and/or modeling, project BSc and MSc]

11 Bayesian data analysis for a model of hyperbolic language use

Kao et al. (2014) suggested an interesting model of how people can interpret non-literal language use. The model is intuitive and backed up by empirical data in the paper. However, the data analysis in the paper leaves some room for exploration. This project would take the data from the original study, and use an implementation of the model to practice a Bayesian data analysis. The goal of this project is to scrutinize the model, i.e., check which aspects of the data the model explains well, and which it might have trouble with.

[Bayesian data analysis, programming, cognitive modeling, individual project or BSc thesis]

12 Wason selection task and optimal data selection

The Wason selection task is a classic from the psychology of reasoning. Oaksford and Chater (1994) offer a rationalization of participants' choice behavior in terms of *optimal data selection*. This model makes some crucial assumptions which influence the predicted outcome. This project would explore the model by implementing it and changing a number of assumptions. For example, the original model assumes that participants' seek to distinguish between two hypotheses (p and q are probabilistically independent vs. p makes q likely). There are other possible hypotheses relevant to the interpretation of conditional sentences which arise if we interpret a conditional as an answer to a particular *question under discussion*, such as *What happens if p?* vs. *Under which circumstances q?*. The project could implement a model and check its predictions with these extended set of participant hypotheses. It could also gather data or use existing data to compare model variants.

[psychology of reasoning, conditionals, linguistics, cognitive modeling, BSc or MSc thesis]

13 Statistical model comparison for models of vague gradable expressions

The interpretation of gradable adjectives, like *tall* or *short*, but also of vague quantifiers like *many* and *few*, seems to depend on statistical world knowledge. At least three formal models of how a *prior expectation*, e.g., about the tallness of a person, can lead to choices of adequacy of a description like *John is tall*: a fixed-threshold hypothesis (Fernando and Kamp, 1996), an evolutionary account (Qing and Franke, 2014), and a pragmatic reasoning account (Lassiter and

Goodman, online first). This project would implement a (Bayesian) model comparison of these approaches based on existing experimental data.

[Bayesian data analysis, programming, linguistics, cognitive modeling, project or BSc/MSc thesis]

14 Preferences for objective meaning in referential expressions

Not every word's semantic meaning is equally uncontroversial. What the word *flat* means is less subjective than what *beautiful* means. Speakers of a language should know (at least roughly) which words are more subjective than others. A theory of rational expression choice would predict that speakers would therefore preferably describe a given referent using words with a less subjective meaning (because that minimizes the chance of misunderstanding, when compared to the use of words that have a more subjective meaning; after all, the listener might not agree which of the dogs in front of us is *beautiful*). In this project, we will set up an experiment in which participants choose expressions to refer to objects, such that they choice options are more or less subjective. We will investigate whether the prediction that less subjective expressions are preferred is supported by the data.

[experimental, some modeling, project or BSc thesis (MSc thesis if extended)]

15 Goodness-of-fit & model comparison

There are many measures for a model's goodness-of-fit and several ways of comparing (cognitive) models. This project would use simulations (based on generating hypothetical empirical data) to compare different theoretical notions against each other. In particular, this project could compare correlation scores against likelihood-based criteria.

[statistics, simulation, theory; can be anything: project, thesis etc. depending on scope and depth]

16 Optimal stopping (secretary problem)

The *secretary problem* is a famous instance of a general class of optimal stopping problems: `https://en.wikipedia.org/wiki/Secretary_problem`. While optimal stopping problems are abstract puzzles which mostly concern theoretical computer science and theoretical economics, this project would explore optimal stopping from an experimental perspective, a topic for which substantial and very interesting literature also exists. We would try to cover novel ground, for instance by having participants learn implicitly the distribution of an attribute which needs to be maximized. For example, participants are shown 50 boxes with black and white marbles in sequence. For each box, they can choose it or inspect the next one. They can never go back but can only choose the current box. The goal is to select the box which contains with the highest number of marbles from the set of all 50 boxes. (Alternatively, the goal could be to choose a box with as large a number of marbles as possible.) To play optimally in this case, participants must screen a number of boxes in order to form beliefs about the distribution of the number of marbles in the offered boxes. We can model that as rational belief learning under uncertainty. We will compare such a rational belief learning + rational choice model to data from an experiment.

[experimental, cognitive modeling, rationality; BSc thesis (ambitious!) or MSc thesis]

References

Anscombre, Jean-Claude and Oswald Ducrot (1983). *L'argumentation dans la langue*. Brussels: Mardaga.

Brochhagen, Thomas and Michael Franke (2017). "Effects of transmission perturbation in the cultural evolution of language". In: *Proceedings of CogSci 39*. Ed. by Glenn Gunzelmann et al. Austin, TX: Cognitive Science Society, pp. 1678–1683.

Brochhagen, Thomas, Michael Franke, and Robert van Rooij (2016). "Learning biases may prevent lexicalization of pragmatic inferences: a case study combining iterated (Bayesian) learning and functional selection". In: *Proceedings of the 38th Annual Conference of the Cognitive Science Society*. Ed. by Anna Papafragou et al. Austin, TX: Cognitive Science Society, pp. 2081–2086.

Cummins, Chris (2012). "Using embedded quantifiers". Talk held at the University of Konstanz, October 13 2012.

Degen, Judith and Michael K. Tanenhaus (2015). "Processing Scalar Implicatures: A Constraint-Based Approach". In: *Cognitive Science* 39, pp. 667–710.

Fernando, Tim and Hans Kamp (1996). "Expecting Many". In: *Proceedings of SALT 6*. Ed. by Teresa Galloway and Justin Spence. Ithaca, NY: Cornell University, pp. 53–68.

Franke, Michael (2016). "Task types, link functions & probabilistic modeling in experimental pragmatics". In: *Proceedings of Trends in Experimental Pragmatics*. Ed. by Fabienne Salfner and Uli Sauerland, pp. 56–63.

Glazer, Jacob and Ariel Rubinstein (2006). "A Game Theoretic Approach to the Pragmatics of Debate". In: *Game Theory and Pragmatics*. Ed. by Anton Benz, Gerhard Jäger, and Robert van Rooij. Palgrave MacMillan, pp. 248–262.

Goodman, Noah D. and Michael C. Frank (2016). "Pragmatic Language Interpretation as Probabilistic Inference". In: *Trends in Cognitive Sciences* 20.11, pp. 818–829.

Kao, Justine T. et al. (2014). "Nonliteral Understanding of Number Words". In: *PNAS* 111.33, pp. 12002–12007.

Lassiter, Daniel and Noah D. Goodman (online first). "Adjectival vagueness in a Bayesian model of interpretation". In: *Synthese*.

Merin, Arthur (1999). "Information, Relevance, and Social Decisionmaking: Some Principles and Results of Decision-Theoretic Semantics". In: *Logic, Language and Computation*. Ed. by Lawrence C. Moss, Jonathan Ginzburg, and Maarten de Rijke. Vol. 2. CSLI Publications, pp. 179–221.

Oaksford, Michael and Nick Chater (1994). "A Rational Analysis of the Selection Task as Optimal Data Selection". In: *Psychological Review* 101.4, pp. 608–631.

Qing, Ciyang and Michael Franke (2014). "Gradable Adjectives, Vagueness, and Optimal Language Use: A Speaker-Oriented Model". In: *Proceedings of SALT 44*. Ed. by Jessi Grieser et al. elanguage.net, pp. 23–41.

Ratcliff, Roger and Gail McKoon (2008). "The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks". In: *Neural Computation* 20, pp. 873–922.

Roettger, Timo B. and Michael Franke (2017). "Information integration and adaptation during intonation-based intention recognition". Manuscript, Cologne, Chicago & Tübingen.

Verbrugge, Rineke and Lisette Mol (2008). "Learning to Apply Theory of Mind". In: *Journal of Logic, Language and Information* 17, pp. 489–511.

Windschitl, P. D. and G. L. Wells (1996). "Measuring psychological uncertainty: Verbal versus numeric methods". In: *Journal of Experimental Psychology: Applied* 2.4, p. 343.

Yalcin, Seth (2010). "Probability Operators". In: *Philosophy Compass* 5.11, pp. 916–937.