communication and pragmatics is significant. In addition, the issue of population heterogeneity will be key in considering how to most effectively understand, treat, and support older adults. Wright seeks to paint a picture of how important this knowledge will be not only currently but also in the future as the older population continues to grow.

REFERENCES

Brown, A. S. (2012). *The tip of the tongue state*. New York, NY: Psychology Press.
Cohen, G., & Burke, D. M. (1993). *Memory for proper names*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Cohen, G., & Faulkner, D. (1986). Memory for proper names: age differences in retrieval. *British Journal of Developmental Psychology*, **4**, 187–197.
James, W. (1893). *The principles of psychology* (Vol. 1). New York, NY: Holt.
Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, **85**(5), 363–394.
Levelt, W. J. M. (1989). *Speaking: from intention to articulation*. Cambridge, MA: MIT Press.
Obler, L. K., Fein, D., Nicholas, M., & Albert, M. L. (1991). Auditory comprehension and aging: decline in syntactic processing. *Applied Psycholinguistics*, **12**, 433–452.
Rayner, K., Reichle, E., Stroud, M., Williams, C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging*, **21**, 448–465.
Salthouse, T. (2012). Consequences of age-related cognitive declines. *Annual Review of Psychology*, **63**, 201–226.
World Health Organization (2015). *Ageing and health* [Fact Sheet]. Online: <http://www.who.int/mediacentre/factsheets/fs404/en/>.

Reviewed by RACHEL H. MESSER

Oklahoma State University, Stillwater, Oklahoma, USA

E-mail: rachel.messer@okstate.edu

K. van Deemter, *Computational Models of Referring: A study in cognitive science*. MIT Press, 2016. Pp. 339. ISBN: 978-0-262-03455-5 (hardcover)

Computational Models of Referring is a book that successfully achieves two goals. First, it gets the reader interested in a problem that in a nutshell could be stated as "how do we create expressions in natural language to refer to the things we want to talk about?" Many could consider the problem of reference generation an easy one, but it actually reveals itself as more and more faceted the more it is explored. Second, the book satiates this newly created curiosity by looking at the problem of reference from different directions. As stated clearly in the beginning of the book, van Deemter considers mainly reference in natural language from the point of view of the speaker/author. That is, the book is a broad discussion about the task known as GENERATING REFERRING EXPRESSION. Referring Expression Generation (REG, in short) is a canonical

746

subtask of Natural Language Generation, arguably one of the most important ones among them, and definitely a complex problem, as becomes clear as this book unfolds. What the book leaves out is, on the one hand, the generation of other semantic structures, such as the relations between concepts and the structure of discourse and, on the other hand, the generation subtasks that are typically tackled down the line, such as aggregation ("John and Mary buy a book" vs. "John buys a book and Mary buys a book"), lexical choice, linearization (making sense of word order), and so on. What is left, however, is more than enough, as shown by the vast corpus of work presented in this book.

The book opens with a map (a directed acyclic graph, in fact) of suggested reading strategies. This is not only a nice idea, and a useful tool for the casual reader who may be interested in some part of the work only, but it also gives a hint of the multi-perspective nature of the book itself. Admittedly, one of the aim of *Computational Models of Referring* is to bring together the results from research area that contribute to our understanding of how the generation of referring expressions work. These areas are essentially three: (i) computational linguistics, combining input from both computer science – algorithmics in particular – and linguistics; (ii) psycholinguistics and cognitive science; and (iii) logic and formal semantics. While these three areas are represented in the book, they are in different proportions. Computational linguistics takes the lion's share: the main story of the book is a succession of algorithms to solve increasingly complex variations of the REG problem, starting from simple solutions and later taking into account different aspects to focus on. For instance, in Natural Language Generation, the main goal of REG is to produce expressions that refer CORRECTLY and COMPLETELY to the intended entity, but humans also use reference in their speech for other functions, such as pointing out specific features of the entities, e.g., the feature with the most discriminatory power.

Part II of the book provides a sequential presentation of solutions to REG, focusing on what van Deemter refers to as the classical REG task. Starting from the most basic algorithms, this section works its way up to solutions that take more and more aspects of the problem into account, such as producing referring expressions that are concise, respectful of the preference order of properties, or accounting for salience. The algorithms are often little more than high-level sketches that abstract over many implementation details, hiding their complexity. This is not a criticism though, as the presentation is always very precise, thanks to the extensive use of the notation from predicate logic and set theory.

Formal logic, and predicate logic in particular, indeed plays a central role in *Computational Models of Referring*. Apart from the notation, the whole task of generating a referring expression here is cast as a function that maps

some elements of a given knowledge base to a LOGICAL FORM. The exact characteristics of the input are not really specified, other than being some kind of formally represented (semantic?) information, but there is no loss of generality here. It is a common leitmotiv of Natural Language Generation to have no fixed format for the representation of input, other than being formally defined. These themes are partly discussed in the third part of the book, in particular in relation to REG from knowledge bases.

More interesting than its input, the output of REG, as framed by van Deemter, is a logical form, that is, a structure with little to no linguistic content. This consideration may be counterintuitive at first – after all, the book is about the generation of referring EXPRESSIONS, natural language utterances. However, the REG problem proves itself to be so multifaceted that it makes sense to restrict its boundaries to the content determination part of the generation process. Unfortunately, the topic of full natural language generation pipelines is not discussed extensively, besides the few elements of algorithms presented in Chapter 8 as extensions to the classic REG. Obviously, a full treatment of REG from its abstract input to the natural language surface form would result in a completely different book, but perhaps a discussion of the interaction between the generation of a correct logical form and its translation into a fluent surface form could highlight new issues, but also ways in which the different tasks can inform each other. In fact, some recent approaches to Natural Language Generation explore architectures that are alternative to the classic pipeline by Reiter and Dale (2000), such as the PhD thesis of Basile (2015).

Part II of *Computational Models of Referring* is also where the cognitive side of the work is introduced, mainly in the form of the TUNA experiment (van Deemter, van der Sluis, & Gatt, 2006). Van Deemter and colleagues at Aberdeen created a corpus of referring expressions paired with the pictures that were shown to the subjects who produced the utterances. During and after the process of collecting the corpus, they were able to test a series of hypotheses about how humans model and produce referring expressions in different situations. The resulting dataset has been used to test subsequent REG systems in shared tasks, effectively becoming the first 'official' benchmark for NLG. Furthermore, the findings from the TUNA experiments help direct the focus of attention for the future development of REG algorithms.

The third part of the book follows naturally from the second part, and presents, one by one, four extension to the classic REG task. As for the previous part, some of the research that serves as source material comes from van Deemter himself, and from other scholars from the Aberdeen school. This includes recent studies on the use of proper names as referring expressions, another case of a seemingly simple yet subtly tricky phenomenon, as well as models

of reference for fuzzy concepts, sets, and so on. Also in this part, Chapter 10 deals with a topic that is perhaps given less space than deserved, that is, the relationship between REG (and NLG in general) and knowledge bases, focusing on those expressed in some Description Logic. The author formalizes the description of a knowledge base that could be used as the starting point for the generation of natural language expressions, and then goes on to show how characteristics of the formalism can be elegantly employed to produce sophisticated referring expressions, e.g., those that include a negation ("The car that is not blue"). There is an important aspect missing from the selection of topics explored in this book, and especially in this section, namely the connection with existing, real-world, large-scale knowledge bases. Since the first increase in popularity of the Web and Web-related technologies, an enormous amount of effort has been made by various communities on formalizing, representing, and computationally reasoning about all kinds of knowledge. Moreover, the de facto standard formalisms of the Semantic Web are based on Description Logic (RDF/OWL and the like). Although it is quite recent, there does exist work on NLG in the framework of the Semantic Web, summarized in Bouayad-Agha, Casamayor, and Wanner (2014), and in the proceedings of the two editions of the WebNLG workshops (https://webnlg2016.sciencesconf.org/). Although this is not a strong complaint, as the book is self-sufficient even without mentioning the Web, completely omitting this discussion from the book feels like a missed opportunity to include relevant recent research, even if it were just an addendum to the chapter on knowledge bases.

The fourth and last part of the boo exhibits a similar structure to the previous one, that is, a number of extensions of the central theme in different directions. As opposed to Part III, here the extensions are about different and problematic situations where the reference could take place. While in the typical manual scenario there is a series of assumptions that REG systems make to simplify the computation, in the real world all kind of scenarios regarding information must be accounted for. This part of the book contains more discussion than algorithms and experiments, which is expected given its nature, yet still provides interesting pointers to more in-depth recent studies.

There are at least a couple of aspects that make this book good reading for students and for readers who are approaching the topic of REG for the first time or from distant research areas. One such aspect is the historical perspective underlying the whole book. Far from being a simple sequential exposition, the book does a good job of introducing new extensions to the generation problem and then motivating each new solution that is introduced, while always maintaining a coherent structure. The writing in general is very consistent, allowing the reader to easily cross the boundaries

749

between the different areas such as psychology and computer science. The reader who comes from a computing or engineering background, in particular, will find the cognitive perspective refreshing, as it really gives a sense of connection between the mechanical algorithms for generating referring expressions and what happens inside the human mind when speaking.

Computational Models of Referring is not a practical book for learning how to program REG systems. Although many examples are provided as pseudo-code algorithms, these are visual hints to aid the comprehension of the approaches that are incrementally showcased, rather than exercises to reproduce, and often at a very high level of abstraction. Regardless of this, I would recommend this book to students, though perhaps more to those interested in the cognitive aspects of the task.

In conclusion, *Computational Models of Referring* is a book I would recommend to scholars interested in the topic of producing references in natural language, either for applicative purposes or to look into the history and the state of the art of such a complex problem. The first set of readers will find many algorithmic approaches, each suitable for a different specification of the REG task at hand, while the second set can still find the computational perspective an interesting companion to more classical cognitive science and psycholinguistic studies. Finally, the book is recommended to whoever may be not aware of the many complexities of reference, as van Deemter's brilliant narrative does a great job at expanding the horizon of the problem one step at a time.

REFERENCES

Basile, Valerio (2015). *From logic to language: natural language generation from logical forms*. PhD thesis, University of Groningen.
Bouayad-Agha, Nadjet, Casamayor, Gerard, & Wanner, Leo (2014). Natural language generation in the context of the semantic web. *Semantic Web*, **5**(6), 493–513.
Reiter, Ehud, & Dale, Robert (2000). *Building natural language generation systems*. New York: Cambridge University Press.
van Deemter, Kees, van der Sluis, Ielka, & Gatt, Albert (2006). Building a semantically transparent corpus for the generation of referring expressions. In *Proceedings of the Fourth International Natural Language Generation Conference, INLG '06* (pp. 130–132). Stroudsburg, PA: Association for Computational Linguistics.

Reviewed by VALERIO BASILE

Université Côte d'Azur, Inria, CNRS, I3S, France

E-mail: valerio.basile@inria.fr