Adaptable Conversational Machines

Nurul Lubis, Michael Heck, Carel van Niekerk, Milica Gašić

■ *In recent years we have witnessed a surge in machine learning methods that provide machines with conversational abilities. Most notably, neural-network–based systems have set the state of the art for difficult tasks such as speech recognition, semantic understanding, dialogue management, language generation, and speech synthesis. Still, unlike for the ancient game of Go for instance, we are far from achieving human-level performance in dialogue. The reasons for this are numerous. One property of human–human dialogue that stands out is the infinite number of possibilities of expressing oneself during the conversation, even when the topic of the conversation is restricted. A typical solution to this problem was scaling-up the data. The most prominent mantra in speech and language technology has been "There is no data like more data." However, the researchers now are focused on building smarter algorithms — algorithms that can learn efficiently from just a few examples. This is an intrinsic property of human behavior: an average human sees during their lifetime a fraction of data that we nowadays present to machines. A human can even have an intuition about a solution before ever experiencing an example solution. The human-inspired ability to adapt may just be one of the keys in pushing dialogue systems toward human performance. This article reviews advancements in dialogue systems research with a focus on the adaptation methods for dialogue modeling, and ventures to have a glance at the future of research on adaptable conversational machines.*

Spoken dialogue systems enable human-computer interaction where primary modes of input are speech and language. It has been a long-standing goal of computer science to build a machine that can communicate with humans using natural language. Alan Turing said that we can only claim to have built true artificial intelligence (AI) once we can talk to it without being able to tell if we are talking to a human being or not (Turing 1950). While this has long been a part of science-fiction literature and films (Rossumovi Univerzální Roboti [Rossum's Universal Robots], or R.U.R., 1920; the Heuristically Programmed Algorithmic or HAL computer in the movie 2001: A Space Odyssey, 1968), due to the surge in the ubiquity of virtual personal assistants such as Apple's Siri, Amazon's Alexa, and Microsoft's Cortana, it seems that this dream is closer to reality than ever. The main driving factors for this development are the impressive results that deep learning has achieved in automatic speech recognition (ASR), exploiting huge amounts of data available for widely-spoken languages. The word error rate on a well-respected conversational English benchmark has dropped from about seventeen percent in 2011 to just about five percent in 2017 — a staggering improvement (Seide, Li, and Yu 2011; Xiong et al. 2017; Saon et al. 2017). While the claim of human parity is still somewhat controversial, it is clear that now the downstream tasks have the chance of making a real impact. Still, it seems that this has only opened a Pandora's box of problems to deal with on our way to achieving intelligent conversation with machines.

  ISSN 0738-4602

The most promising way to tackle the problem of human-computer conversation is to take a machine learning approach: collect data, define the model, and then make predictions. The data are simply dialogues or some part of dialogues, such as annotated user requests or user-provided feedback. The model is the underlying statistical model that we use to explain the data we have. In dialogue, a variety of models are used: supervised, unsupervised, or reinforcement-learning (RL) models. Machine learning methods hinge on data. More crucially, they hinge on this data being good. That is, showing high coverage, being extensively annotated with high inter-annotator agreement labels, as well as being consistent and noise-free. The reality is, however, that only a tiny fraction of available data sets have these properties. Particularly for dialogue modeling, these requirements are difficult to meet due to the dynamic and infinite nature of human conversation.

In contrast, the requirements for humans to achieve good comprehension skills are much lower than for machines, according to studies (Moore 2003). An oft-cited estimation speaks of exposure to up to 10,000 hours of speech for 10-year–olds (Hart and Risley 1995), which amounts to fewer than 100-million spoken words. Today's large-scale conversational ASR systems with acceptable performance are trained on tens of thousands of hours of recorded speech and on billions of written words. Given the struggles to achieve human parity for ASR, one needs little imagination to see how data requirements would be even more demanding for building conversational AI. However, even the largest dialogue corpora, few in number, do not exceed 1,000 hours of speech (Serban et al. 2018).

This realization leads us to another route in dialogue system research. While we can embark on a mission to collect larger and better data sets, which would certainly help research in this area, we also need to address a more fundamental issue: how to adapt to and learn from imperfect conditions. Endowing machines with this human-inspired capability is one of the necessary steps in advancing conversational machines to make our sci-fi visions come true. In this article, we review research efforts toward adaptive conversational AI with a focus on the adaptation methods used in dialogue modeling and cast a glance at possible future developments.

Spoken Dialogue System 101

For a long time, the research in conversational AI has been divided into two tracks: task-oriented-dialogue systems and chat-based systems. Task-oriented dialogue systems are conversational systems that provide information based on a particular user goal — for example, finding an Italian restaurant in the east of town. On the other hand, chat-based systems could talk about anything the user possibly wants and focus on simulating a human-like chit-chat. This article will mainly focus on task-oriented dialogue systems, drawing parallels with chat-based systems where necessary.

The most widely adopted approach to dialogue systems within the research community has followed the divide-and-conquer paradigm. The idea is to divide the system into smaller modules such that each module can be trained with well-defined and well-labeled data sets. Figure 1 illustrates an interaction within the typical pipeline of a statistical spoken dialogue system.

The front-end of a dialogue system allows spoken interaction between the user and system. It consists of a speech recognizer, which turns user's speech into text, and a speech synthesizer, which maps natural language in text form to speech. In recent years, both tasks have hugely benefited from the advancement of neural network models, showing continuous improvements in performance (Seide, Li, and Yu 2011; Hinton et al. 2012; van den Oord et al. 2016). In a dialogue system, the text-to-speech synthesizer can make use of dialogue context to produce more natural and expressive speech (Yu et al. 2010).

Underlying all the dialogue system modules is the ontology, a structural representation of concepts that can occur in conversations with the system. In information-seeking dialogue systems, they contain categories of interest, called *slots,* such as *food*, *area*, *price*; their possible instances, called *values*, such as *Italian*, *east*, *cheap*; and actions that the system can perform, called *dialogue acts,* such as *request*, *inform*, *confirm*. The amount of work that flows into building good ontologies by hand is substantial, which inevitably renders them finite and limited to specific domains, such as *restaurant*.

Given a user utterance, the natural language understanding unit decodes the concepts from the ontology that occur in the input using a dialogue act formalism. In its simplest form, a dialogue act formalism represents each user utterance by a dialogue act type and a list of slot values (Traum 2000). Subsequently, the belief tracker is tasked with keeping track of user goals throughout the dialogue, producing the belief state. While natural language understanding only considers the current user utterance, the belief tracker considers the whole dialogue context up to that point in time. More recently, these modules have merged into a single module, alleviating the need for labeled data to train a natural language understanding component and to avoid information loss between natural language understanding and the belief tracker. However, belief trackers are still imperfect and rely on the availability of sufficient labeled data.

Given a belief state, the policy decides on the system's next action. Because fulfilling user goals typically requires multiple dialogue turns, it is not sufficient to estimate the optimal immediate action, but to behave in such a way that the overall conversation will be successful. Consequently, the dialogue is treated as a sequential decision-making task. The policy is optimized using an RL algorithm to maximize the reward of the dialogue. The reward function

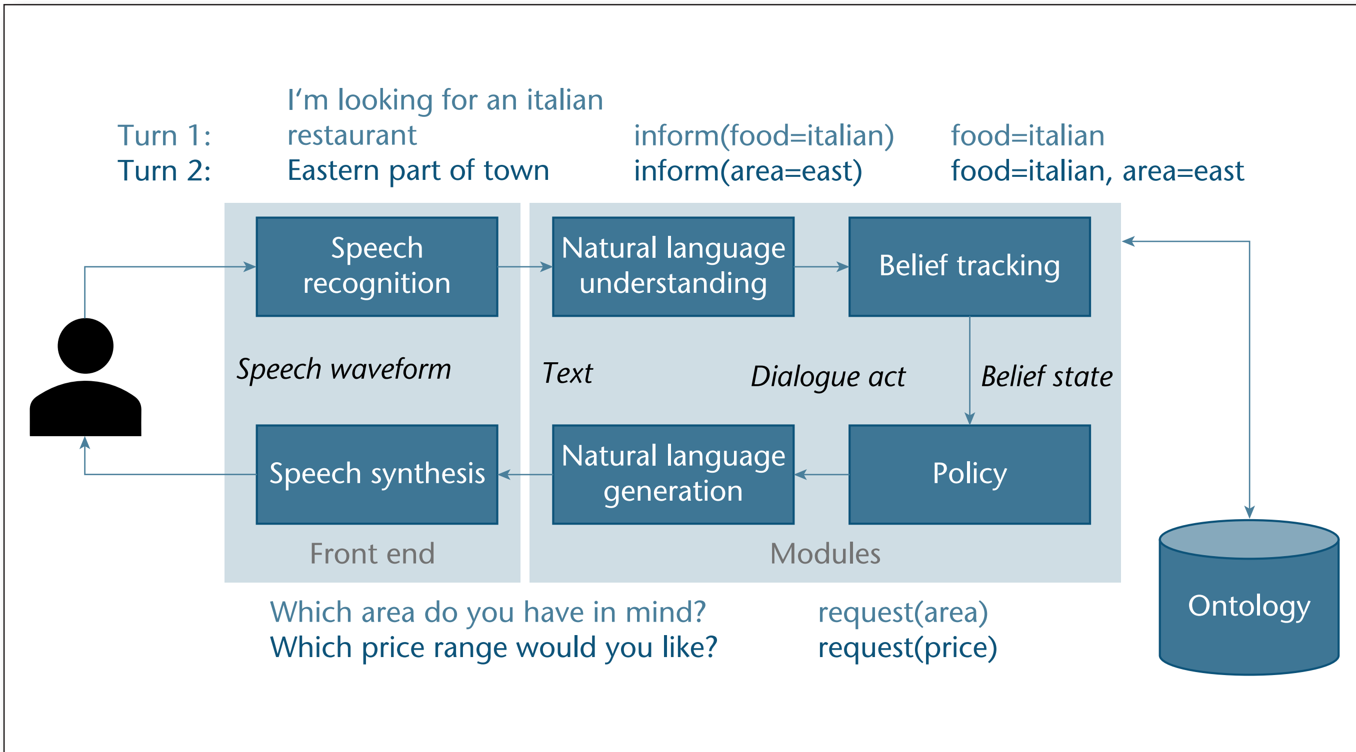

Figure 1. Modular Approach to Statistical Dialogue Systems.

defines the objective of the dialogue. Task success is still commonly used as a dialogue-level reward; however, there are also competing approaches to learn the reward function from data (Walker et al. 1997; Yang, Levow, and Meng 2012; El Asri, Laroche, and Pietquin 2012; Pietquin 2013; Su et al. 2015). Figure 2 illustrates the difference between belief tracking and policy optimization in a dialogue. Policy is essential in ensuring a goal-directed behavior, as governed by the reward function. However, the need for a substantial amount of training interactions is often limiting.

The natural language generator (NLG) component deals with mapping between the system action, selected by the policy, to its corresponding natural language response. A traditional approach is to use hand-coded rules (Walker, Rambow, and Rogati 2002). Class-based (Oh and Rudnicky 2000) and phrase-based (Mairesse et al. 2010) NLGs have also been explored. State-of-the-art systems use long short-term memory networks to generate a sequence of words as the system response, conditioned on the dialogue act representation (Wen et al. 2015). Despite the progress, producing long and coherent sentences remains a difficult challenge in NLG.

We have witnessed progress in each of these tasks and in developing task-oriented dialogue systems as a whole. However, there are issues arising as a consequence of this design. We highlight these issues, and later show how adaptation is a key in minimizing or perhaps eliminating them. The first issue is error propagation down the pipeline. To some extent, each of these modules will produce an error, causing the next module in line to start with incorrect input, accumulating the error over time and severely degrading the final output. Second is the ontology-dependent nature of each module. Having to predefine the concepts that can appear in dialogue poses a serious limitation, especially because ontology coverage is minuscule relative to the number of concepts that humans can talk about. It is desirable to have a dialogue system that is able to learn from its experience, compound knowledge through interaction, and evolve over time — as humans do.

End-to-End Neural Network Dialogue Modeling

In recent years, research on chat-bots and dialogue modeling has started to intertwine, driven by advancements in deep learning. This is owed to the fact that in both settings the basic problem can be treated as a sequence-to-sequence learning task: the user input can be treated as the input sequence and the system output can be treated as the output sequence. For such tasks, models based on neural networks are generally most effective (Sutskever, Vinyals, and Le 2014). Instead of the modular pipeline, a dialogue system can be modeled using a neural network that maps user input sequence directly to system output sequence in an end-to-end fashion. A range of mechanisms are used, including memory networks and additional supervision signals, in an attempt to make the conversation task-oriented (Zhao and Eskenazi

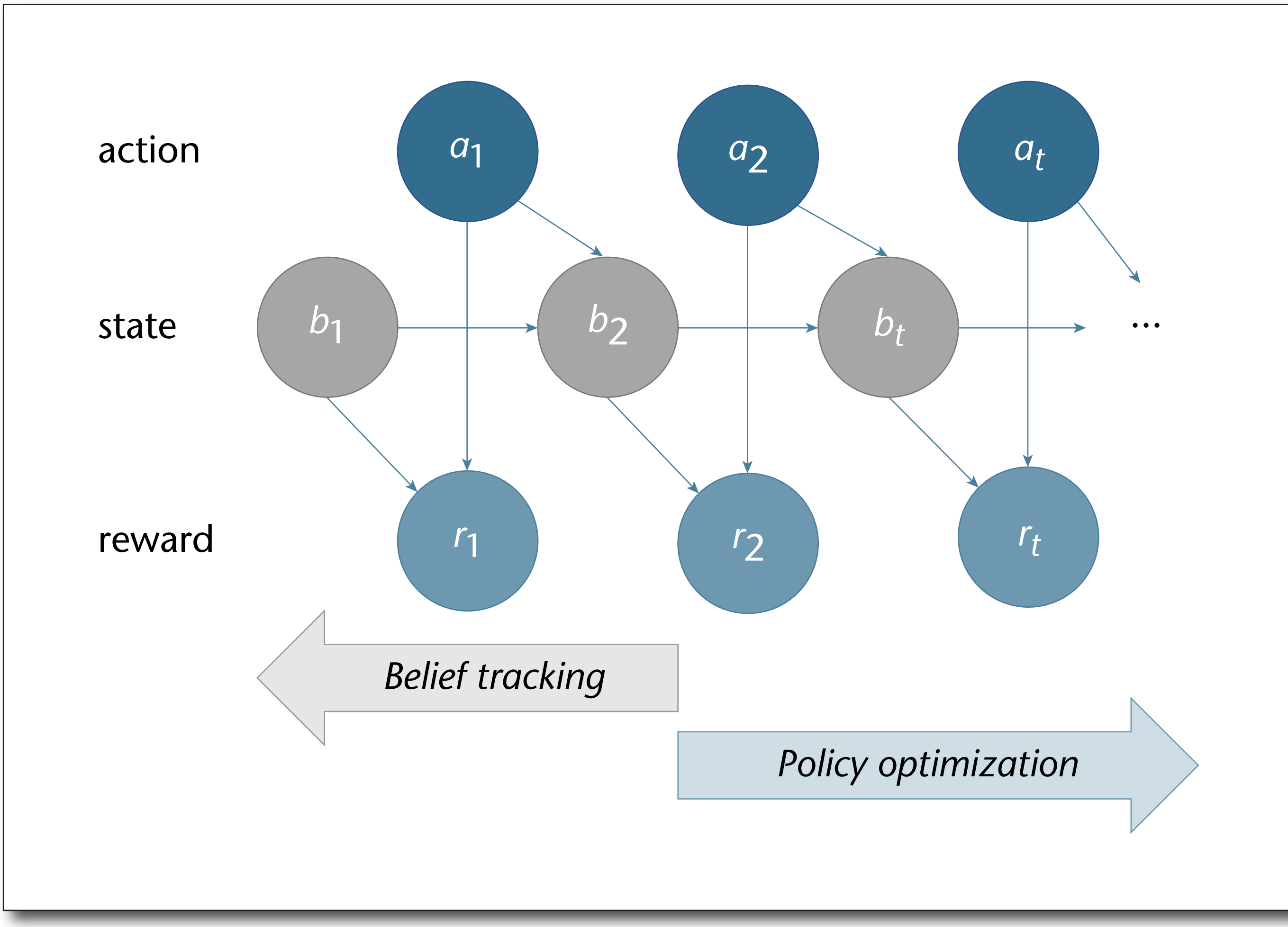

Figure 2. The Temporal Contrast between Belief Tracking and Policy Optimization.

At each dialogue turn, belief tracking aims to encapsulate the past while policy focuses on maximizing future rewards.

2016; Bordes, Boureau, and Weston 2017; Dhingra et al. 2017; Li et al. 2017; Serban et al. 2015; Yang et al. 2017; Wen et al. 2017b-). On the whole, results score highly on automatic translation measures such as bilingual evaluation understudy (BLEU), which measures the similarity of the generated output with the data (Papineni et al. 2002), that is, the output indeed appears to be human-like. However, end-to-end systems are not yet subject to as strict a human evaluation as their modular systems counterpart. That is, while human evaluation metrics such as appropriateness and diversity are occasionally reported in addition to BLEU or perplexity, the measure of task success is not typically included. Therefore, when it comes to completing a particular task, the results are less conclusive. Moreover, the approaches investigated so far do not readily scale to changes in the ontology.

Adaptation

Modularization of a dialogue system such as that described above allows machine learning methods to be applied to each module (Young 2002; Lemon and Pietquin 2012). Still, in reality there is a long way to go from state-of-the-art to achieve human parity in dialogue. Dialogue is an example of an AI-complete task (Shapiro 1992) as it requires understanding, reasoning, and generation. If we take the machine learning approach to model dialogue, we quickly are confronted with a mismatch between training and testing conditions, especially when the systems are deployed in the wild.

For example, although arguably speech recognizers nowadays achieve parity with human performance in a noise-free environment (Xiong et al. 2017), ASR still produces high error rates when applied in dialogue systems, as they are meant to be used in a variety of situations — noisy cars or busy streets, for instance. Noise-robustness can be improved by propagating the uncertainty further down the pipeline of a modular dialogue system, which has led to approaches using partially observable Markov decision processes (Zhang et al. 2001; Young 2002; Williams and Young 2007; Thomson and Young 2010; Young

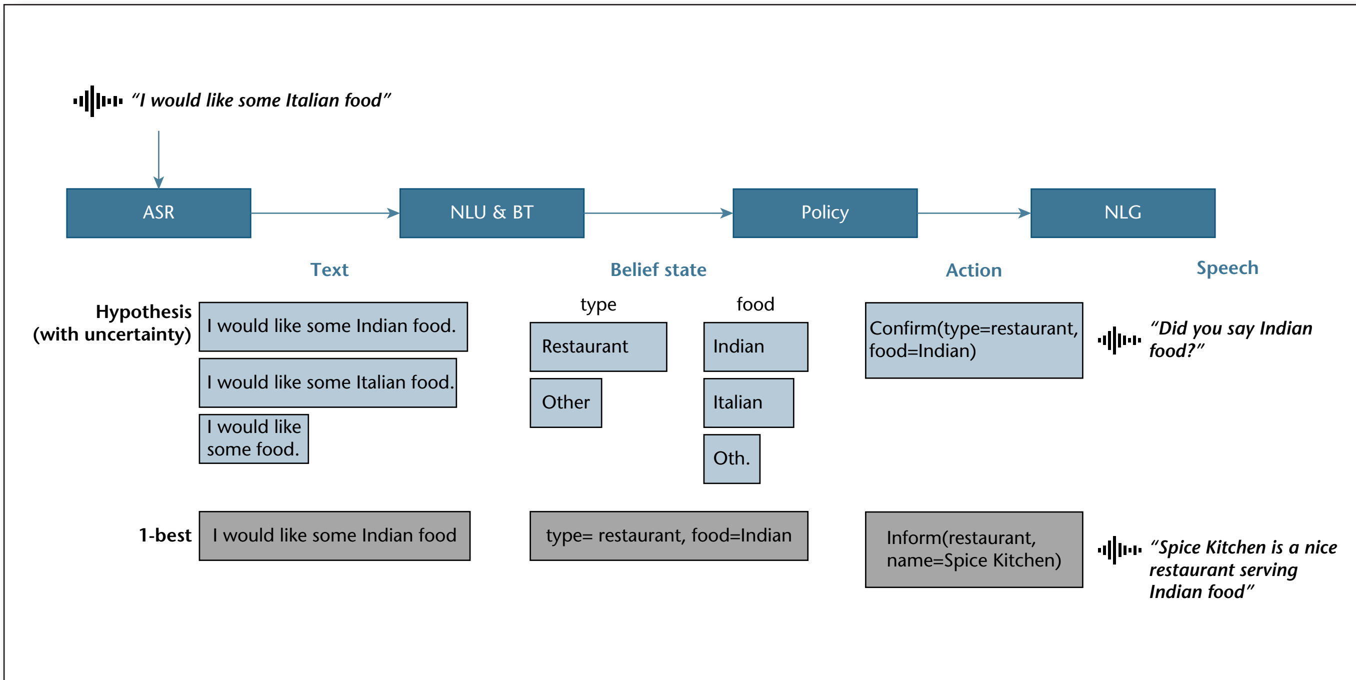

Figure 3. Propagation of Uncertainty in a Dialogue System Pipeline.

In blue is an example of a dialogue system that utilizes uncertainty propagation; the horizontal bar chart illustrates a probability distribution. The system is able to ask for confirmation in unsure cases. In gray, the system does not consider uncertainty at the expense of selecting an incorrect action.

et al. 2010; Lee and Eskénazi 2012; Young et al. 2013). The main idea is that each module of the spoken dialogue system takes as input a probability distribution and likewise produces a probability distribution as output. In this way, the uncertainty that arises from speech recognition or understanding errors can be propagated to other modules, and in particular to the decision-making module (see figure 3). Although state-of-the-art dialogue modeling is nowadays mostly based on deep learning, these principles still prevail and are essential to ensure robustness.

More fundamentally than noise level, mismatch in knowledge and language complexity between human users and dialogue systems greatly affects ease of use and, consequently, user success rate. For example, a task can be simplified into a series of yes/no questions. This is very intuitive for the user; however, it comes at the cost of severely restricting the dialogue. Furthermore, the system has no chance to learn about new concepts unless a new one is programmed into it. On the other hand, many of today's conversational AIs break down during their interaction with the user, likely due to their inability to match human dialogue despite them giving an impression of being able to facilitate one. If a dialogue system was to reach human parity, the ease of use and success rate would increase, but not at the cost of oversimplifying the interaction. The major challenge of today's research on conversational machines is to escape the *uncanny valley* that lies between the two extremes of overly simplistic, designed interaction, and human-like, natural interaction (see figure 4).

Adaptation could be a means for this much-needed leap forward. A key is to encapsulate large knowledge and allow its accumulation over time through flexible, adaptive models. When doing this, it is important that we rely on sample-efficient methods that are able to learn from imperfect conditions. These goals are particularly challenging for dialogue systems, as they deal with infinite possibilities and very complex patterns.

Adaptive Ontology and Word Vectors

Populating ontologies from text is the process of deriving high-level concepts and relations from information (Wong, Liu, and Bennamoun 2012). Given the pervasion and ubiquity of natural-language–based systems, there is a growing need for adaptive ontologies that can include new knowledge with ease, preferably fully automated.

For the longest time, most research focused on expanding specific ontologies using semiautomated tools for supporting humans in the loop (Agirre et al. 2000; Navigli and Velardi 2004). Approaches along this line generally use, among others, linguistic techniques and lexico-syntactic patterns (Pantel and Pennacchiotti 2006; Aguado de Cea et al. 2008), clustering techniques (Agirre et al. 2000; Witschel 2005), statistical techniques (Sugiura et al. 2003), and association rules (Bodenreider, Aubry, and Burgun 2005;

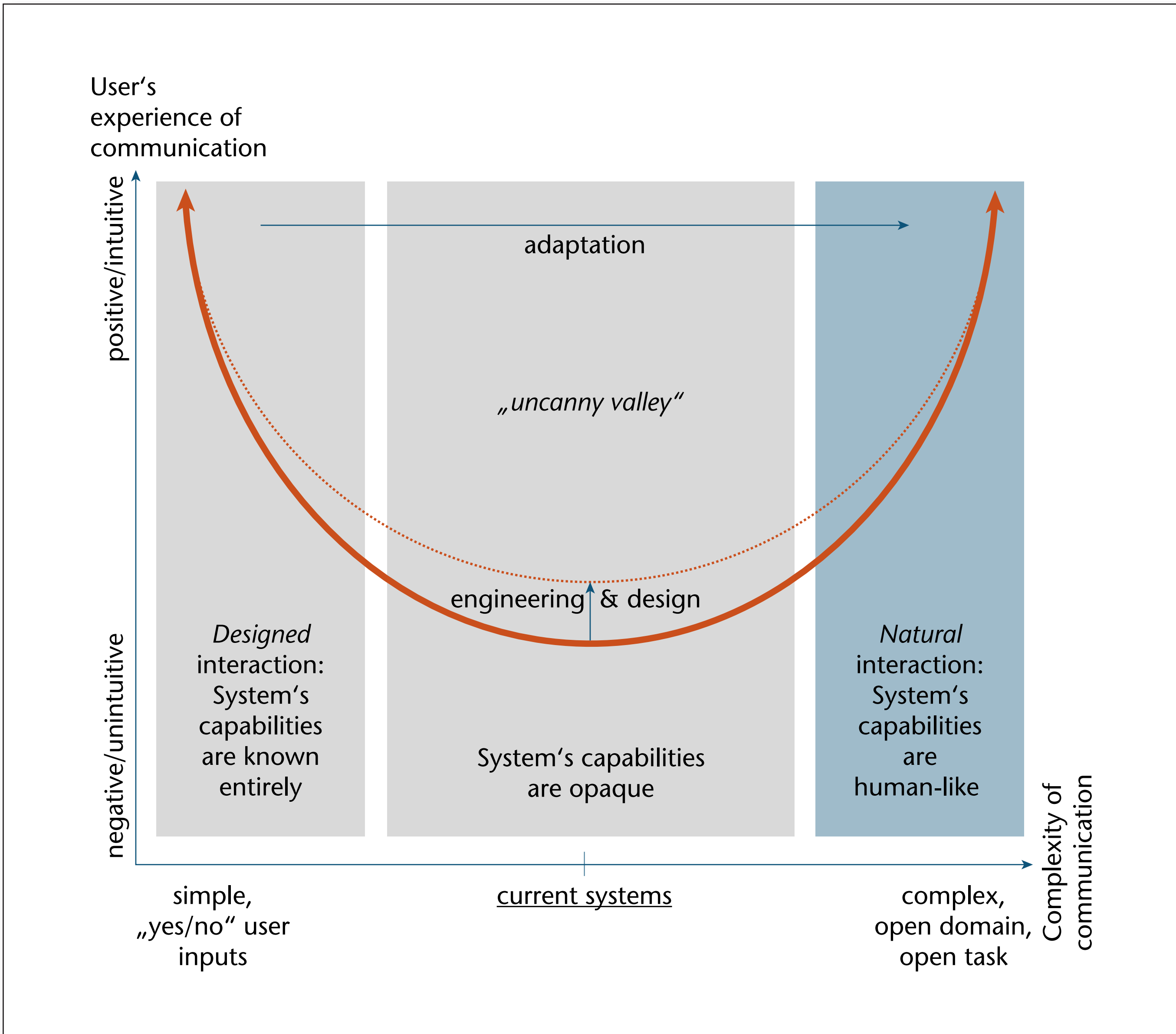

Figure 4. Relationship between User Experience and System Complexity.

One important indicator of human-like dialogue systems is positive user experience, which is highly influenced by system complexity. In this regard, engineering and design may alleviate shortcomings to a certain extent. However, adaptation is one of the key technologies to overcome the *uncanny valley* of dialogue systems.

Gulla, Brasethvik, and Kvarv 2009). This angle demands manual work, rules, and heuristics, which naturally results in severe limitations.

For building truly adaptive and flexible ontologies, we need to start from the beginning and look at what is the nature of the information that flows into these models. For instance, it is intuitive to represent words as discrete, atomic units, or sequences thereof, such as characters. The problem with atomic units is that no meaningful comparison is possible except equality testing (figure 5). Even with character-level comparisons between words, only lexical similarity can be measured. The similarity in meaning remains unquantifiable. To match potential slot values in a user's utterance with ontology entries, heuristics and rules have long been indispensable.

Distributional representations are continuous multidimensional representations of tokens. Modeling words as real-valued vectors open up an entirely different approach to modeling concepts and their relations. One great advantage of such approaches is their ability to represent concepts in a more compact manner, which helps to fight the curse of dimensionality (Bengio et, al. 2003). Popularly known as *word*

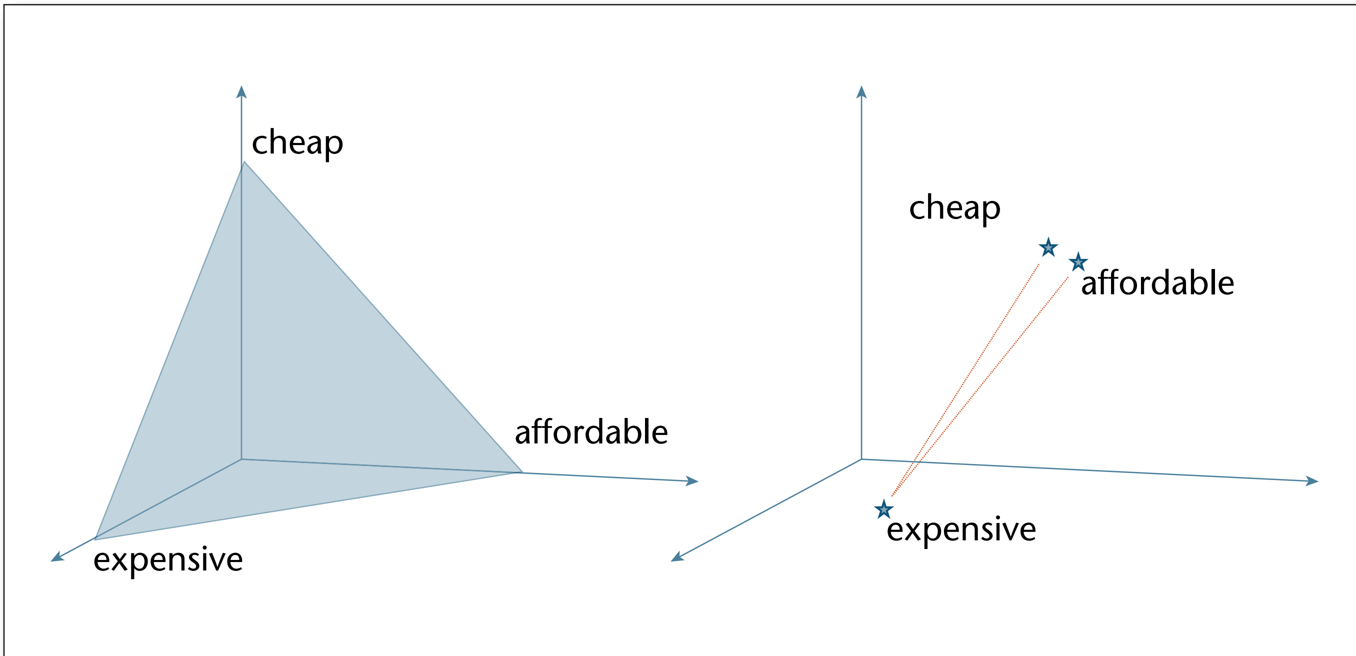

Figure 5. Words Represented as Atomic Units or Vectors.

(*Left*) Words that are represented as atomic units live at the corners of a simplex. The distance between any two units is identical to the distance of any other pair, that is, the only possible similarity test is the test for equality. (*Right*) Words that are represented as vectors in a continuous semantic embedding space become comparable by applying suitable distance measures. Similar words inhabit the same neighborhoods. Conversely, dissimilar words are located apart from each other.

vectors, distributional representations have been around for some time in the sphere of natural language processing (Collobert et al. 2011). The distributional hypothesis, which states that "the meaning of words lie in their use" (Wittgenstein 1953), provides a basis for distributional semantics, that is, a data-driven statistical study of word meanings. In the 1990s, latent semantic analysis emerged as an early method to compute vectors such that words that appear in similar contexts possess similar representations. This approach has been massively popularized by neural methods emerging in the early 2010s, such as Word-to-Vector (Mikolov et al. 2013a, 2013b) and Global Vectors for Word Representation (Pennington, Socher, and Manning 2014), which learn word vectors by being trained to reconstruct the context of words. These methods all have in common that they embed semantic similarities between words in a continuous vector space (see figure 5). Improvements and extensions soon followed, with fastText (Bojanowski et al. 2017) providing a way to meaningfully embed previously unseen words, and contextualized word vectors (McCann, Bradbury, Xiong, and Socher 2017), enabling the inference of contextualized vector representations. The field has since seen a surge of sophisticated language modeling techniques used to learn latent contextual word and sentence representations. Most prominent among these are Embeddings from Language Models (ELMo) and Bidirectional Encoder Representations from Transformers (BERT), along with others (Peters et al. 2018; Devlin et al. 2019; Sun et al. 2019), most of which are based on the popular transformer neural network architecture (Vaswani et al. 2017). These new and massive models led to a wave of drastic improvements on traditionally very hard natural language processing problems such as sentiment classification, entailment detection, question answering, and co-reference resolution (Wang et al. 2019a, 2019b). For dialogue tasks specifically, the potency of these models has recently been demonstrated impressionably by the chatbot Meena (Adiwardana et al. 2020). Distributed word representations have been shown to help in building adaptive models that represent knowledge in one form or another (Mitchell et al. 2018). Recently proposed models for dialogue modeling fuse semantic parsing and belief tracking by relying on distributed representations of concepts as well (Mrkšić et al. 2016b; Mrkšić and Vulić 2018; Ramadan, Budzianowski, and Gašić 2018). With this approach, learned semantic similarity is used to probe dialogue utterances for the presence of ontology terms. The ontology can be substantially expanded without the need to hand-code specialized semantic dictionaries, simply through the use of comparable vector representations. In addition to that, ambiguity is inherently captured by these representations, which naturally alleviates input restrictions. Recent research on ontology adaptation and ontology growing follows these lines as well. Initial work expands the set of values for any given slot by applying similarity measures to known values and words in the input that have not been seen before (Jayawardana et al. 2017). Other work exploits the presence of common slots in multiple domains to facilitate cross-domain adaptation (Wu et al. 2019). The main underlying challenge here is to learn distributed representations that capture meaningful properties of concepts and semantic similarity, which is a nontrivial task and therefore a very active field of research.

Dialogue modeling differs from other natural language processing tasks in that the distributional hypothesis is problematic in task-oriented systems, where words like *cheap* and *expensive* share near-identical contexts but clearly differ in their impact on a dialogue. Specialized word vectors help alleviate such issues (Mrkšić et al. 2016b). A method called *retrofitting* shifts vectors in the embedding space so that semantically similar words, based on external relational information, have similar vector representations (Faruqui et al. 2015). Counter-fitting injects antonymy and synonymy constraints into vector space representations to improve their capability to judge semantic similarity (Mrkšić et al. 2016a). Semantic quality of word vectors can also be refined by introducing constraints considering information from multilingual corpora (Mrkšić et al. 2017). Specialization of this kind so far has only been achieved for static word embeddings, which on their own underperform in comparison with contextual representations. Targeted specialization of contextual word embeddings is nontrivial because they do not live in a fixed space, and it therefore remains an open problem.

Most recently, research follows the trend of valueless ontologies, that is, values remain entirely undefined and are extracted from user input directly (Chao and Lane 2019; Wu et al. 2019). These approaches use distributed word representations such as ELMo and BERT. The internal states of such models constitute contextualized word vectors. With a suitable strategy for fine-tuning, these models have been shown to be extremely efficient at solving downstream tasks such as identifying words of interest in user's dialogue input — for example, potential values for a requested slot. With the desire to build complex and dynamic ontologies comes the need for more sophisticated dialogue system components in general that can support changing and growing knowledge. This especially concerns the belief tracker, policy optimization, and the NLG, as the ontology is at the heart of these subsystems.

Adaptive Multidomain Belief Tracking

A belief tracker is tasked to predict user goals in a dialogue by considering the dialogue context from the beginning up to the current point in time. While this task can be reasonably solved within a limited and simple domain, the complexity increases exponentially when multiple domains are considered. Strides have been made with the surge of neural network methods that have been shown to be particularly effective for representation learning and language modeling.

Methods such as the Multi-Domain Belief Tracker (Ramadan, Budzianowski, and Gašić 2018) and the Neural Belief Tracker (Mrkšić and Vulić 2018) attempted to solve the task at hand by using a one-versus-all approach, which makes predictions for each possible configuration independently. While this method seems like a good solution, it has proven to not generalize well to multidomain settings. A more adaptive approach is proposed by the Global-Local Self-Attention encoder state tracker (Zhong, Xiong, and Socher 2018). The Global-Local Self-Attention encoder state tracker proposes to use a likelihood scoring method of observing the current utterances. Applying this scoring has achieved some heightened success in predicting the value given a domain-slot pairing. However, the quality of a model such as the Global-Local Self-Attention encoder state tracker or Globally Conditioned Encoder (Nouri and Hosseini-Asl, 2018) relies on the quality of the representations learned.

Obtaining meaningful and robust representations of the user goal has proven to be very challenging. One method that has stood out in this regard is the use of a multi-head self-attention mechanism to extract meaning from dialogue utterances. This mechanism utilizes the estimated relative importance of words through the use of attention to create better representations. The Slot Utterance Matching Belief Tracker (Lee, Lee, and Kim 2019) uses this mechanism to obtain representations for domain-slot pairs, which has shown to produce more robust scores over possible value candidates. While this approach shows promising results in a fixed ontology setting, it lacks the ability to extract values not present in the ontology.

Adaption to new domains, slots, or values not seen during training time is a problem not yet addressed by many methods. One means to take on this challenge could lie in finding ways of efficiently utilizing more of the readily available data, including nondialogue data. For example, can a dialogue system learn to talk about music by utilizing music reviews? An approach is to fine-tune the representation model, such as BERT (Devlin et al. 2019), utilizing unstructured data. Using this representation and language modeling as a self-supervised fine-tuning task, we can then teach the model about new concepts and their relations in a specific domain (Mitchell et al. 2018). Furthermore, could models actually extract user-mentioned values from the dialogue utterances? Results from the Transferable Dialogue State Generator (Wu et al. 2019) method have shown the potential of extracting new values from text, supporting adaptation to unseen values. The natural next step is adaptation to unseen slots or domains, a problem that, to our knowledge, has not yet been solved.

As shown in table 1, although state-of-the-art methods perform very well on single-domain dialogue, these models heavily struggle when more than one domain is considered. The best available models still make erroneous predictions roughly half of the time

Model	WOZ2.0	MultiWOZ2.0	MultiWOZ2.1
Multi-Domain Belief Tracker (Ramadan, Budzianowski, and Gašić 2018)	85.5%	15.6%	—
Globally Conditioned Encoder (Nouri and Hosseini-Asl, 2018)	88.5%	35.5%	—
Slot Utterance Matching Belief Tracker (Lee, Lee, and Kim 2019)	91.0%	42.4%	—
Transferable Dialogue State Generator (Wu et al., 2019)	—	48.6%	45.6%
Non-Autoregressive Dialogue State Tracking (Lee, Lee, and Kim 2019)	—	50.5%	49.0%
Dual Strategy - Dialogue State Tracking (Zhang et al., 2019)	—	—	51.2%
Selectively Overwriting Memory - Dialogue State Tracking (Kim et al., 2019)	—	51.4%	52.6%

Single-domain: WOZ2.0; multidomain: MultiWOZ2.0/2.1.

Table 1. Comparison of User Goal Prediction Accuracy in Single-Domain and Multidomain Environments.

when conversation can be about multiple domains simultaneously, such as a dialogue about finding a hotel and restaurant in the same area. Clearly more work needs to be done to bring these numbers on par by improving belief tracker adaptability to multiple, dynamic domains.

Policy Adaptation

To create a dialogue system that is able to handle a variety of tasks, we need a policy that can handle multiple domains in a seamless manner. The majority of approaches to policy optimization assume that the policy needs to be optimized as a function of the whole belief state. For large domains, however, this belief state becomes a high-dimensional vector, and standard RL approaches do not scale. This is known as the curse of dimensionality: The complexity increases exponentially with the number of features in the belief state vector.

One way to build a multidomain dialogue system is to use a committee of policies. The idea is to develop different expertise in each committee member by training them on different data, allowing the committee to rely on the member with most experience at each decision. At each point in time, each policy can propose an action to take, out of which one is selected according to a heuristic (Lison 2011). Using a Bayesian committee machine (Tresp 2000) approach, the collective decision-making can be handled more elegantly through statistical methods and by incorporating uncertainty (Gašić et al. 2017). A trade-off is the inability to combine the decisions from the experts — only one domain expert is chosen at each turn. The computational cost is linearly dependent on the number of committee members. The problem remains of how to handle multiple domains within a single dialogue turn.

Another route is to train a multidomain policy. More recently, deep RL has been applied on the policy optimization task; for example, see the papers by Cuayáhuitl, Keizer, and Lemon (2015), Fatemi et al. (2016), and Williams, Asadi, and Zweig (2017), among others. The problem can be modeled as a neural network that predicts the expected discounted cumulative reward of each state and action pair, based on which the policy selects the next action given a particular state to maximize this value. One way to promote knowledge sharing is to consider two kinds of agents in a policy — slot-dependent and slot-independent (Chen et al. 2018). Slot-dependent agents have parameters that keep track of common characteristics of all slots, as well as private parameters to model characteristics specific to each slot. Given a new domain, the shared parameters can be easily transferred to ensure good initialization, leading to faster convergence. Unfortunately, deep RL approaches are not robust to errors, and adaptation toward more complex domains is still suboptimal.

To facilitate policy adaptation, it is important to learn quickly and efficiently, even in large, growing action spaces. This is in line with the exploration toward more sample-efficient methods in RL (Su et al. 2017; Lipton et al. 2018; Weisz et al. 2018). Actor–critic experience replay has been applied in dialogue system settings, showing faster learning that yields state-of-the-art performance in a large action space (Weisz et al. 2018). The biggest gain in efficiency is due to the experience replay, which allows learning from a particular experience multiple times (Munos et al. 2016), unlike typical RL algorithms that are able to learn from an experience only once.

Large action spaces can also be tackled by decomposition. Many tasks are hierarchical in nature: there are subgoals that the agent needs to complete first. For example, in interaction with a hotel booking system, the user must first be able to find a hotel, then book it, and only then pay for it. The more complex the intended dialogue becomes, the more pressing the need for a subtasks hierarchy. To date, no standard RL algorithms that can efficiently solve hierarchical problems are available (Duan et al. 2016).

However, early research on hierarchical RL was promising (Dayan and Hinton 1993; Dietterich 2000; Barto and Mahadevan 2003). There has already been some initial success in applying this technique to dialogue modeling (Cuayáhuitl et al. 2010; Casanueva et al. 2018). The idea is to model the dialogue as a semi-Markov process so that the policy can be decomposed into a hierarchy of policies. Each policy operates only on a part of the belief space, making it very efficient. This hierarchy, however, needs to be predefined.

Recently researchers have also ventured into self-play for policy learning, where a dialogue system talks to itself or another dialogue system as a form of learning (Li et al. 2016; Shah et al. 2018). Self-play has been demonstrated to be more sample-efficient (Gupta et al. 2019) and effective in estimating interactive quality of dialogue (Ghandeharioun et al. 2019). However, arriving at a rich dialogue policy from self-play training is still a challenge as it is prone to dialogue breakdown between the systems; that is, after several turns, the dialogue cannot continue further as the systems get stuck in an infinite loop of repetitive responses such as "I don't know." Sample efficiency and scalable learning are two of the most important keys in facilitating a successful policy adaptation in dialogue. The aforementioned approaches demonstrate ways to achieve them, although their application within the context of policy adaptation in dialogue systems is yet to be exploited. Another bottleneck in domain adaptation that has not yet been resolved is the need to predefine all actions and slots prior to learning. It is desirable for a policy to be able to dynamically adapt to new concepts discovered throughout an interaction, similar to how humans continuously evolve and learn through communication (Chandramohan et al. 2014), moving further beyond multi-domain dialogue systems to open-domain ones.

Adaptation in Response Generation

Early efforts in NLG relied on templates and handcrafted rules. However, these approaches are quickly outgrown by the increasing urgency of more varied system responses in dialogue as well as adaptation capabilities. Recent approaches use language modeling with neural networks, treating NLG as a sequence generation problem conditioned on the dialogue action from the policy. Typically, a discrete flat representation is used as dialogue action, such as *inform*(type=hotel, area=center, price=cheap). One approach is to supply the dialogue act representation to condition the language generation using a recurrent neural network (Wen et al. 2015). More recently, novel dialogue act representations have been proposed to allow better domain adaptation through knowledge sharing. For example, by using a tree-structured representation of dialogue actions (Tseng et al. 2019), or a multilayered graph (Chen et al. 2019). The main idea is to merge identical parts of multiple domain ontologies, allowing information sharing in the dialogue act representation across domains. However, this line of methods does not work on ontologies that have only few or no common elements because they operate on the premise of merging similarities between the domains. There is still a need for a method that allows interpolation or inference of new knowledge.

Strong supervision in the form of extensively labeled data is required to optimize an NLG module. This may not always be readily available in huge amounts, if at all, especially for multi-domain settings. In this regard, NLG adaptation also benefits from a refined learning process and creative means to collect data. An approach is to use multiple adaptation stages: using an out-of-domain data set to counterfeit the in-domain dialogue for pretraining before fine-tuning on the in-domain data (Wen et al. 2016). NLG and decision-making can also be seen as a joint optimization objective. This allows adaptation to fine-grained changes in dialogue context (Lemon 2008). Furthermore, general features can be learned across domains using a model-agnostic meta-learning algorithm (Tran and Nguyen 2018).

Beside domain adaptation, there is an interest in user adaptation by incorporating emotion- or personality-based conditioning in the response (Walker et al. 2007; Mairesse and Walker 2010; Lubis et al. 2018; Oraby et al. 2018; Zhou et al. 2018; Colombo et al. 2019;). Incorporation of affective aspects has been reported to increase user satisfaction and feeling of closeness, as well as improving rapport and user acceptance (Higashinaka, Dohsaka, and Isozaki 2008; Acosta 2009, Saini et al. 2005). While the majority of works in this regard are done in chat-oriented settings, results on the more task-oriented tutoring and navigation systems are promising as well (Litman and Silliman 2004; Bui et al. 2007). To push conversational AI toward human parity, user adaptation strategies such as these are essential.

Adaptation in End-to-End Dialogue Systems

Research in end-to-end task-oriented systems is largely inspired by the success of sequence-to-sequence modeling for chat-oriented systems (Serban et al. 2015). While a pure end-to-end approach may rid us of the dependency on ontology, the degree of freedom that the system has may be too high, causing difficulty in forming meaningful responses. A way to compensate is by using a copy mechanism (Sukhbaatar et al. 2015) on the knowledge base (Zhong, Xiong, and Socher 2018), or by using a memory network to keep track of retrieved knowledge base entities and words that have appeared in dialogue (Madotto, Wu, and Fung 2018). The need to incorporate some kind of structure is apparent to improve the performance of end-to-end task oriented systems.

A modularly connected end-to-end dialogue system utilizes a neural network as each of its modules. This type of end-to-end system passes on information in a similar manner to a modular system. However, each module outputs a representation instead of structured data. With this setting it is possible to introduce a separate policy network that is trainable via RL for continuous improvement of the system (Wen et al. 2017a). To enforce the structure further, each module can be pretrained to predict its corresponding structured output prior to fusion and integration (Mehri, Srinivasan, and Eskenazi 2019). A flexible structure has also been introduced in the paper by Shu et al. (2019), which utilizes separate decoders for each of the inform, request, and response slots. Although this has been shown to help, incorporation of structure often requires multiple supervisions, increasing the data requirement for training.

Structure in an end-to-end model reveals an avenue for adaptation, most notably by utilizing representation learning methods. The aim is to train a dialogue encoder that produces representations with a good domain-generalizability. Ideally, such representations should allow accurate interpolation given a similar dialogue context in a new domain. In reality, though, achieving such representations is still challenging. A common method leverages pretraining objectives that are inspired by natural language processing tasks, such as next-utterance retrieval (Lowe et al. 2016) and generation (Vinyals and Le 2015). Naturally, this requires a sufficient amount of additional data often along with its corresponding label or supervision. The choice of the pretraining objective has been demonstrated to highly influence generalizability of the learned representation. For example, inconsistency identification is reported to be a more effective pretraining objective for response generation, compared with masked-utterance retrieval and next-utterance retrieval (Mehri et al. 2019). However, it is still unclear what kind of pretraining objective truly maximizes the gain in domain adaptation.

Minimizing the effort in system adaptation has been one particular interest in this line of research. The task can be framed as the so-called zero-shot or few-shot learning, referring to the amount of training data of the new domain that the model receives before it is put to the test on that new domain; none (zero-shot), or only a small amount (few-shot). One approach is to induce a latent action space that spans across domains using dialogue context-response pairs as well as a set of response-dialogue act pairs (Zhao and Eskenazi 2018). Representations obtained with variational methods (Zhao, Lee, and Eskenazi 2018) used by Shalyminov et al. (2019) allows an end-to-end training using raw dialogue data only. While improvement on metrics such as task success and entity recognition are reported, the numbers are still very low when tested on real human dialogue. Furthermore, lack of interpretability and controllability remains a major challenge in this family of models, especially where adaptation is concerned.

Conclusion and Outlook

We are at a transition phase in dialogue system research, moving from simplistic and restricted human-computer dialogue, into a dynamic and adaptive one that can learn and evolve over time. We believe that adaptation is the means for this much-needed leap forward. A key is to dynamically encapsulate large knowledge and allowing its accumulation over time through flexible, adaptive models.

Approaches using partially observable Markov decision processes have helped improve adaptation to different noise levels by allowing error propagation down the dialogue system pipeline. The surge of neural network methods had a huge impact in dialogue system research as well, enabling better knowledge representation in the ontology, function approximations, and language modeling in the modules. Sample-efficient methods aided faster learning, decreasing the data requirement and speeding up convergence.

Despite the progress made, there is still much to be done to achieve human parity in conversational AI. The reality is that being able to operate today's dialogue system is a skill of its own. Patient and determined users have to learn how to talk to dialogue systems (Sadun and Sande 2014), finding out what they can or cannot do, putting intent in ways the systems can understand, and so forth. Humans are the ones adapting to dialogue systems instead of the other way around. Turning this around is a demanding task, whose eventual solutions have a high potential impact on the future of conversational machines.

A critical step in advancing task-oriented systems involves overcoming the limitations of having to predefine system capacities. It is not obvious to users which tasks a system can handle and which concepts (actions, slots, values) it knows. This opaqueness combined with limited conversational range leads to poor satisfaction. Furthermore, continual learning and adaptation of the system through time is not possible.

In this regard, it is important to make the systems as independent of a static ontology as possible. A first step could be a truly value-less ontology, where a system could seamlessly detect values unseen during training during dialogue with human users. This allows the system to quickly adapt to new real-life concepts, such as new restaurants and movie names. A step beyond that is the ability to extract a new family of concepts and the relationships within for ontology growing, that is, new slots or even domains, from the vast knowledge contained in the worldwide web and unstructured data. Lastly, we need to be able to incorporate new actions into the conversation to be able to drastically increase the conversational capabilities. This may require a universal and dynamic action-embedding space that maps system intents in dialogue. Consequently, each of the modules should have the abilities to incorporate these new concepts on-the-go.

This vision raises the question of how to represent knowledge and complex relations across dialogue system modules. We have started exploration beyond discrete spaces toward continuous ones; however, this comes with a new set of challenges such as the need for more complex supervision, including how to induce spaces with properties that support seamless adaptation across modules. In this regard, we are likely to benefit from research on knowledge-infused and semantically specialized contextual representations, which has just picked up pace. In terms of modeling, neural networks have been consistently shown to be a powerful method for solving each of the aforementioned challenges, but it is important to increase their interpretability and controllability to facilitate truly successful adaptation.

Acknowledgments

Funding has been provided by the Alexander von Humboldt Foundation within the framework of the Sofja Kovalevskaja Award, endowed by the Federal Ministry of Education and Research.

References

Acosta, J. 2009. Using Emotion to Gain Rapport in a Spoken Dialogue System. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, Companion Volume (Student Research Workshop and Doctoral Consortium), 49–54. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.3115/1620932.1620941.

Adiwardana, D.; Luong, M.-T.; So, D. R.; Hall, J.; Fiedel, N.; Thoppilan, R.; Yang, Z.; Kulshreshtha, A.; Nemade, G.; Lu, Y., et al. 2020. *Towards a Human-Like Open-Domain Chatbot*. arXiv:2001.09977. Ithaca, NY: Cornell University Library.

Agirre, E.; Ansa, O.; Hovy, E.; and Martínez, D. 2000. Enriching Very Large Ontologies Using the WWW. Paper presented at the European Conference on Artificial Intelligence (ECAI) Workshop on Ontology Learning. Berlin, Germany, August 20–25.

Aguado de Cea, G.; Gómez-Pérez, A.; Montiel-Ponsoda, E.; and Suárez-Figueroa, M. C. 2008. Natural Language-Based Approach for Helping in the Reuse of Ontology Design Patterns. In *Proceedings of the International Conference on Knowledge Engineering and Knowledge Management*, Volume 5268 (Lecture Notes in Computer Science), Berlin, Germany: Springer. doi.org/10.1007/978-3-540-87696-0_6.

Barto, A. G., and Mahadevan, S. 2003. Recent Advances in Hierarchical Reinforcement Learning. *Discrete Event Dynamic Systems* 13(1/2): 41–77. doi.org/10.1023/A:1022140919877.

Bengio, Y.; Ducharme, R.; Vincent, P.; and Jauvin, C. 2003. A Neural Probabilistic Language Model. *Journal of Machine Learning Research* 3(Feb): 1137–55.

Bodenreider, O.; Aubry, M.; and Burgun, A. 2005. Non-Lexical Approaches to Identifying Associative Relations in the Gene Ontology. *Pacific Symposium on Biocomputing* 10: 91–102.

Bojanowski, P.; Grave, E.; Joulin, A.; and Mikolov, T. 2017. Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Linguistics* 5: 135–46. doi.org/10.1162/tacl_a_00051.

Bordes, A.; Boureau, Y.; and Weston, J. 2017. *Learning End-to-End Goal-Oriented Dialogue*. arXiv:1605.07683. Ithaca, NY: Cornell University Library.

Bui, T. H.; Poel, M.; Nijholt, A.; and Zwiers, J. 2007. A Tractable DDN-POMDP Approach to Affective Dialogue Modeling for General Probabilistic Frame-based Dialogue Systems. *In* Proceedings of the 5th Workshop on Knowledge and Reasoning in Practical Dialogue Systems, 34–57. ris.utwente.nl/ws/files/5315464/bnaic2008-1.pdf.

Casanueva, I.; Budzianowski, P.; Su, P.-H.; Ultes, S.; Barahona, L. M. R.; and Tseng, B.-H.; and Gašić, M. 2018. Feudal Reinforcement Learning for Dialogue Management in Large Domains. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Volume 2 (Short Papers), 714–9. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.18653/v1/N18-2112.

Chandramohan, S.; Geist, M.; Lefevre, F.; and Pietquin, O. 2014. Co-adaptation in Spoken Dialogue Systems. In *Natural Interaction with Robots, Knowbots and Smartphones*, 343–53. Berlin, Germany: Springer. doi.org/10.1007/978-1-4614-8280-2_31.

Chao, G.-L., and Lane, I. 2019. BERT-DST: Scalable End-to-End Dialogue State Tracking with Bidirectional Encoder Representations from Transformer. In *Proceedings of Interspeech 2019*, 1468–72. Baixas, France: International Speech Communication Association. doi.org/10.21437/Interspeech.2019-1355.

Chen, L.; Chang, C.; Chen, Z.; Tan, B.; Gašić, M.; and Yu, K. 2018. Policy Adaptation for Deep Reinforcement Learning-based Dialogue Management. In *2018 Institute of Electrical and Electronics Engineers (IEEE) International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6074–8. Piscataway, NJ: Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/ICASSP.2018.8462272.

Chen, W.; Chen, J.; Qin, P.; Yan, X.; and Wang, W. Y. 2019. Semantically Conditioned Dialogue Response Generation via Hierarchical Disentangled Self-Attention. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 3696–709. Stroudsberg, PA: Association for Computational Linguistics. https://doi.org/10.18653/v1/P19-1360.

Collobert, R.; Weston, J.; Bottou, L.; Karlen, M.; Kavukcuoglu, K.; and Kuksa, P. 2011. Natural Language Processing (Almost) from Scratch. *Journal of Machine Learning Research* 12(Aug): 2493–537.

Colombo, P.; Witon, W.; Modi, A.; Kennedy, J.; and Kapadia, M. 2019. Affect-Driven Dialogue Generation. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Volume 1 (Long and Short Papers), 3734–43. Stroudsberg, PA: Association for Computational Linguistics.

Cuayáhuitl, H.; Keizer, S.; and Lemon, O. 2015. *Strategic Dialogue Management via Deep Reinforcement Learning*. arXiv: 1511.08099. Ithaca, NY: Cornell University Library.

Cuayáhuitl, H.; Renals, S.; Lemon, O.; and Shimodaira, H. 2010. Evaluation of a Hierarchical Reinforcement Learning Spoken Dialogue System. *Computer Speech & Language* 24(2): 395–429. https://doi.org/10.1016/j.csl.2009.07.001.

Dayan, P., and Hinton, G. E. 1993. Feudal Reinforcement Learning. *In* Advances in Neural Information Processing Systems. 271–8. San Francisco, CA: Morgan Kaufmann.

Devlin, J.; Chang, M.-W.; Lee, K.; and Toutanova, K. 2019. BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Volume 1 (Long and Short Papers), 4171–86. Stroudsberg, PA: Association for Computational Linguistics.

Dhingra, B.; Li, L.; Li, X.; Gao, J.; Chen, Y.-N.; Ahmed, F.; and Deng, L. 2017. Towards End-to-End Reinforcement Learning of Dialogue Agents for Information Access. In

Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, Volume 1 (Long Papers), 484–95. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.18653/v1/P17-1045.

Dietterich, T. G. 2000. Hierarchical Reinforcement Learning with the MAXQ Value Function Decomposition. *Journal of Artificial Intelligence Research* 13: 227–303. doi.org/10.1613/jair.639.

Duan, Y.; Chen, X.; Houthooft, R.; Schulman, J.; and Abbeel, P. 2016. Benchmarking Deep Reinforcement Learning for Continuous Control. *Journal of Machine Learning Research* 48: 1329–38.

El Asri, L.; Laroche, R.; and Pietquin, O. 2012. Reward Function Learning for Dialogue Management. STAIRS 2012: 95–106.

Faruqui, M.; Dodge, J.; Jauhar, S. K.; Dyer, C.; Hovy, E.; and Smith, N. A. 2015. Retrofitting Word Vectors to Semantic Lexicons. In *Proceedings of the North American Chapter of the Association for Computational Linguistics (NAACL)*. Stroudsberg, PA: Association for Computational Linguistics.

Fatemi, M.; El Asri, L.; Schulz, H.; He, J.; and Suleman, K. 2016. Policy Networks with Two-Stage Training for Dialogue Systems. *In* Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue, 101–10. doi.org/10.18653/v1/W16-3613.

Gašić, M.; Mrkšić, N.; Rojas-Barahona, L.M.; Su, P.-H.; Ultes, S.; Vandyke, D.; Wen, T.-H.; and Young, S. 2017. Dialogue Manager Domain Adaptation Using Gaussian Process Reinforcement Learning. *Computer Speech & Language* 45: 552–69. doi.org/10.1016/j.csl.2016.09.003.

Ghandeharioun, A.; Shen, J. H.; Jaques, N.; Ferguson, C.; Jones, N.; Lapedriza, A.; and Picard, R. 2019. *Approximating Interactive Human Evaluation with Self-Play for Open-Domain Dialogue Systems*. arXiv:1906.09308. Ithaca, NY: Cornell University Library.

Gulla, J. A.; Brasethvik, T.; and Kvarv, G. S. 2009. Association Rules and Cosine Similarities in Ontology Relationship Learning. *Lecture Notes in Business Information Processing* 19: 201–12. doi.org/10.1007/978-3-642-00670-8_15.

Gupta, A.; Lowe, R.; Foerster, J.; Kiela, D.; and Pineau, J. 2019. Seeded Self-play for Language Learning. In *Proceedings of the Beyond Vision and LANguage: inTEgrating Real-world kNowledge (LANTERN)*, 62–6. Hong Kong, China: Association for Computational Linguistics. doi.org/10.18653/v1/D19-6409.

Hart, B., and Risley, T. R. 1995. *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore, MD: Paul H. Brookes Publishing.

Higashinaka, R.; Dohsaka, K.; and Isozaki, H. 2008. Effects of Self-Disclosure and Empathy in Human–Computer Dialogue. In *2008 Institute of Electrical and Electronics Engineers (IEEE) Spoken Language Technology Workshop*, 109–12. Piscataway, NJ: Institute of Electrical and Electronics Engineers (IEEE). doi.org/10.1109/SLT.2008.4777852.

Hinton, G.; Deng, L.; Yu, D.; Dahl, G. E.; Mohamed, A.; Jaitly, N.; Senior, A.; Vanhoucke, V.; Nguyen, P.; Sainath, T. N.; and Kingsbury, B. 2012. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. *IEEE Signal Processing Magazine* 29(6): 82–97. doi.org/10.1109/MSP.2012.2205597.

Jayawardana, V.; Lakmal, D.; de Silva, N.; Perera, A. S.; Sugathadasa, K.; Ayesha, B.; and Perera, M. 2017. Semi-Supervised Instance Population of an Ontology Using Word Vector Embedding. In *2017 Seventeenth International Conference on Thinking (ICT) Advances for Emerging Regions (ICTer)*, 1–7. Piscataway, NJ: Institute of Electrical and Electronics Engineers (IEEE). doi.org/10.1109/ICTER.2017.8257822.

Kim, S.; Yang, S.; Kim, G.; and Lee, S.-W. 2019. Efficient dialogue state tracking by selectively overwriting memory. arXiv preprint arXiv:1911.03906.

Lee, H.; Lee, J.; and Kim, T.-Y. 2019. SUMBT: Slot-Utterance Matching for Universal and Scalable Belief Tracking. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 5478–83. Stroudsberg, PA: Association for Computational Linguistics (ACM). doi.org/10.18653/v1/P19-1546.

Lee, S., and Eskénazi, M. 2012. POMDP-Based Let's Go System for Spoken Dialogue Challenge. In *Proceedings of the 2012 Institute of Electrical and Electronics Engineers (IEEE) Spoken Language Technology Workshop (SLT)*. Piscataway, NJ: Institute of Electrical and Electronics Engineers (IEEE). doi.org/10.1109/SLT.2012.6424198.

Lemon, O. 2008. Adaptive Natural Language Generation in Dialogue Using Reinforcement Learning. *In* Proceedings of the 12th Workshop on the Semantics and Pragmatics of Dialogue (LONDIAL), 141–8. events.illc.uva.nl/semdial/proceedings/semdial2008_londial_proceedings.pdf.

Lemon, O., and Pietquin, O. 2012. *Data-Driven Methods for Adaptive Spoken Dialogue Systems*. Berlin, Germany: Springer. doi.org/10.1007/978-1-4614-4803-7.

Li, J.; Monroe, W.; Ritter, A.; Jurafsky, D.; Galley, M.; and Gao, J. 2016. Deep Reinforcement Learning for Dialogue Generation. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 1192–202. doi.org/10.18653/v1/D16-1127.

Li, X.; Chen, Y.-N.; Li, L.; Gao, J.; and Celikyilmaz, A. 2017. End-to-End Task-Completion Neural Dialogue Systems. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing* (Volume 1: Long Papers), 733–43. Taipei, Taiwan: Asian Federation of Natural Language Processing. www.aclweb.org/anthology/I17-1074.pdf.

Lipton, Z.; Li, X.; Gao, J.; Li, L.; Ahmed, F.; and Deng, L. 2018. BBQ-Networks: Efficient Exploration in Deep Reinforcement Learning for Task-oriented Dialogue Systems. In *Thirty-Second Association for the Advancement of Artificial Intelligence (AAAI) Conference on Artificial Intelligence*. Palo Alto, CA: Association for the Advancement of Artificial Intelligence (AAAI) Press.

Lison, P. 2011. Multi-Policy Dialogue Management. In *Proceedings of the Special Interest Group on Discourse and Dialogue (SIGdial) 2011 Conference*, 294–300. Stroudsberg, PA: Association for Computational Linguistics.

Litman, D. J., and Silliman, S. 2004. ITSPOKE: An Intelligent Tutoring Spoken Dialogue System. In *Demonstration Papers at the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (HLT-NAACL) 2004*, 5–8. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.3115/1614025.1614027.

Lowe, R.; Serban, I. V.; Noseworthy, M.; Charlin, L.; and Pineau, J. 2016. On the Evaluation of Dialogue Systems with Next Utterance Classification. In *Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 264–9. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.18653/v1/W16-3634.

Lubis, N.; Sakti, S.; Yoshino, K.; and Nakamura, S. 2018. Eliciting Positive Emotion Through Affect-Sensitive Dialogue Response Generation: A Neural Network Approach. In *Thirty-Second Advancement of Artificial Intelligence (AAAI) Conference on Artificial Intelligence*. Palo Alto, CA: Association for the Advancement of Artificial Intelligence (AAAI) Press.

Madotto, A.; Wu, C.-S.; and Fung, P. 2018. Mem2Seq: Effectively Incorporating Knowledge Bases into End-to-End Task-Oriented Dialogue Systems. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, Volume 1 (Long Papers), 1468–78. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.18653/v1/P18-1136.

Mairesse, F.; Gašić, M.; Jurćićek, F.; Keizer, S.; Thomson, B.; Yu, K.; and Young, S. 2010. Phrase-based Statistical Language Generation Using Graphical Models and Active Learning. In *Proceedings of the Association for Computational Linguistics (ACL)*. Stroudsberg, PA: Association for Computational Linguistics.

Mairesse, F., and Walker, M. A. 2010. Towards Personality-Based User Adaptation: Psychologically Informed Stylistic Language Generation. *User Modeling and User-Adapted Interaction* 20(3): 227–78. doi.org/10.1007/s11257-010-9076-2.

McCann, B.; Bradbury, J.; Xiong, C.; and Socher, R. 2017. *Learned in Translation: Contextualized Word Vectors*. arXiv:1708.00107. Ithaca, NY: Cornell University Library.

Mehri, S.; Razumovskaia, E.; Zhao, T.; and Eskenazi, M. 2019. Pretraining Methods for Dialogue Context Representation Learning. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics,* 3836–45. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.18653/v1/P19-1373.

Mehri, S.; Srinivasan, T.; and Eskenazi, M. 2019. Structured Fusion Networks for Dialogue. In *Proceedings of the 20th Annual Special Interest Group on Discourse and Dialogue (SIGdial) Meeting,* 165–77. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.18653/v1/W19-5921.

Mikolov, T.; Chen, K.; Corrado, G.; and Dean, J. 2013a. *Efficient Estimation of Word Representations in Vector Space*. arXiv:1301.3781. Ithaca, NY: Cornell University Library.

Mikolov, T.; Sutskever, I.; Chen, K.; Corrado, G. S.; and Dean, J. 2013b. *Distributed Representations of Words and Phrases and their Compositionality*. arXiv:1310.4546. Ithaca, NY: Cornell University Library.

Mitchell, T.; Cohen, W.; Hruschka, E.; Talukdar, P.; Yang, B.; Betteridge, J.; Carlson, A.; Dalvi, B.; Gardner, M.; Kisiel, B., et al. 2018. Never-Ending Learning. *Communications of the ACM* 61(5): 103–15. doi.org/10.1145/3191513.

Moore, R. K. 2003. A Comparison of the Data Requirements of Automatic Speech Recognition Systems and Human Listeners. Paper presented at the 8th European Conference on Speech Communication and Technology, EUROSPEECH 2003, Geneva, Switzerland, September 1–4, 2003.

Mrkšić, N.; Séaghdha, D.; Thomson, B.; Gašić, M.; Rojas-Barahona, L.; Su, P.-H.; Vandyke, D.; Wen, T.-H.; and Young, S. 2016a. Counter-Fitting Word Vectors to Linguistic Constraints. In *Proceedings of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (HLT-NAACL)*. Stroudsberg, PA: Association for Computational Linguistics (ACL).

Mrkšić, N.; Séaghdha, D. Ó.; Wen, T.; Thomson, B.; and Young, S. J. 2016b. *Neural Belief Tracker: Data-Driven Dialogue State Tracking*. arxiv.org/abs/1606.03777. Ithaca, NY: Cornell University Library.

Mrkšić, N.; Vulić, I.; Séaghdha, D. Ó.; Leviant, I.; Reichart, R.; Gašić, M.; Korhonen, A.; and Young, S. J. 2017. Semantic Specialization of Distributional Word Vector Spaces Using Monolingual and Cross-lingual Constraints. *Transactions of the Association for Computational Linguistics* 5: 309–24. doi.org/10.1162/tacl_a_00063.

Mrkšić, N., and Vulić, I. 2018. Fully Statistical Neural Belief Tracking. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, Volume 2 (Short Papers), 108–13. Stroudsberg, PA: Association for Computational Linguistics.

Munos, R.; Stepleton, T.; Harutyunyan, A.; and Bellemare, M. 2016. Safe and Efficient Off-Policy Reinforcement Learning. In *Advances in Neural Information Processing Systems*. arxiv.org/abs/1606.02647. Ithaca, NY: Cornell University Library.

Navigli, R., and Velardi, P. 2004. Learning Domain Ontologies from Document Warehouses and Dedicated Web Sites. *Computational Linguistics* 30(2): 151–79. doi.org/10.1162/089120104323093276.

Nouri, E., and Hosseini-Asl, E. 2018. Toward scalable neural dialogue state tracking model. In Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), 2nd Conversational AI workshop.

Oh, A. H., and Rudnicky, A. I. 2000. Stochastic Language Generation for Spoken Dialogue Systems. In *Proceedings of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 27–32. Stroudsberg, PA: Association for Computational Linguistics.

Oraby, S.; Reed, L.; Sharath, T.; Tandon, S.; and Walker, M. 2018. Neural MultiVoice Models for Expressing Novel Personalities in Dialogue. In *Proceedings of Interspeech 2018*, 3057–61. Baixas, France: International Speech Communication Association. doi.org/10.21437/Interspeech.2018-2174.

Pantel, P., and Pennacchiotti, M. 2006. Espresso: Leveraging Generic Patterns for Automatically Harvesting Semantic Relations. In *Proceedings of the Association for Computational Linguistics (ACL)*, 113–20. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.3115/1220175.1220190.

Papineni, K.; Roukos, S.; Ward, T.; and Zhu, W.-J. 2002. BLEU: A Method for Automatic Evaluation of Machine Translation. In *Proceedings of the Association for Computational Linguistics (ACL)*. Stroudsberg, PA: Association for Computational Linguistics.

Pennington, J.; Socher, R.; and Manning, C. 2014. GloVe: Global Vectors for Word Representation. *In* Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 1532–43. doi.org/10.3115/v1/D14-1162.

Peters, M. E.; Neumann, M.; Iyyer, M.; Gardner, M.; Clark, C.; Lee, K.; and Zettlemoyer, L. 2018. Deep Contextualized Word Representations. *In* Proceedings of the North American Chapter of the Association for Computational Linguistics (NAACL), 2227–37. www.aclweb.org/anthology/N18-1202.pdf.

Pietquin, O. 2013. Inverse Reinforcement Learning for Interactive Systems. In *Proceedings of the 2Nd Workshop on Machine Learning for Interactive Systems: Bridging the Gap Between Perception, Action and Communication, MLIS '13*, 71–5. New York: Association for Computing Machinery (ACM). doi.org/10.1145/2493525.2493529.

Ramadan, O., and Budzianowski, P.; and Gašić, M. 2018. Large-Scale Multi-Domain Belief Tracking with Knowledge Sharing. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics* (Volume 2: Short Papers), 432–7. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.18653/v1/P18-2069.

Sadun, E., and Sande, S. 2014. *Talking to Siri: Mastering the Language of Apple's Intelligent Assistant*. Indianapolis, IN: Que Publishing.

Saini, P.; De Ruyter, B.; Markopoulos, P.; and Van Breemen, A. 2005. Benefits of Social Intelligence in Home Dialogue

Systems. In *International Federation for Information Processing (IFIP) Conference on Human–Computer Interaction*, 510–21. Berlin, Germany: Springer. doi.org/10.1007/11555261_42.

Saon, G.; Kurata, G.; Sercu, T.; Audhkhasi, K.; Thomas, S.; Dimitriadis, D.; Cui, X.; Ramabhadran, B.; Picheny, M.; Lim, L.-L., et al. 2017. English Conversational Telephone Speech Recognition by Humans and Machines. In Proceedings of Interspeech 2017, 132–6. doi.org/10.21437/Interspeech. 2017-405.

Seide, F.; Li, G.; and Yu, D. 2011. Conversational Speech Transcription Using Context-dependent Deep Neural Networks. In *Conference of the International Speech Communication Association*, 437–40. Baixas, France: International Speech Communication Association.

Serban, I. V.; Lowe, R.; Henderson, P.; Charlin, L.; and Pineau, J. 2018. A Survey of Available Corpora for Building Data-driven Dialogue Systems: The Journal Version. *Dialogue & Discourse* 9(1): 1–49.

Serban, I. V.; Sordoni, A.; Bengio, Y.; Courville, A. C.; and Pineau, J. 2015. *Hierarchical Neural Network Generative Models for Movie Dialogues*. arxiv.org/abs/1507.04808. Ithaca, NY: Cornell University Library.

Shah, P.; Hakkani-Tür, D.; Liu, B.; and Tür, G. 2018. Bootstrapping a Neural Conversational Agent with Dialogue Self-Play, Crowdsourcing and On-Line Reinforcement Learning. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Volume 3 (Industry Papers), 41–51. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.18653/v1/N18-3006.

Shalyminov, I.; Lee, S.; Eshghi, A.; and Lemon, O. 2019. Few-Shot Dialogue Generation Without Annotated Data: A Transfer Learning Approach. In *Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue,* 32–9. Stroudsberg, PA: Association for Computational Linguistics. doi.org/ 10.18653/v1/W19-5904.

Shapiro, S. C. 1992. Artificial Intelligence. In *Encyclopedia of Artificial Intelligence*, Second edition. New York: John Wiley.

Shu, L.; Molino, P.; Namazifar, M.; Xu, H.; Liu, B.; Zheng, H.; and Tur, G. 2019. Flexibly-Structured Model for Task-Oriented Dialogues. In *Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue*, 178–87. Stroudsberg, PA: Association for Computational Linguistics. doi.org/ 10.18653/v1/W19-5922.

Su, P.-H.; Budzianowski, P.; and Ultes, S.; Gašić, M.; and Young, S. 2017. Sample-efficient Actor-Critic Reinforcement Learning with Supervised Data for Dialogue Management. In *Proceedings of the 18th Annual International Symposium on Computer Architecture (ISCA) and Association for Computational Linguistics (ACL) Special Interest Group on Discourse and Dialogue (SIGdial) Meeting on Discourse and Dialogue*, 147–57. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.18653/v1/W17-5518.

Su, P.-H.; Vandyke, D.; Gašić, M.; Kim, D.; Mrkšić, N.; Wen, T.-H.; and Young, S. 2015. *Learning from Real Users: Rating Dialogue Success with Neural Networks for Reinforcement Learning in Spoken Dialogue Systems*. arXiv:1508.03386. Ithaca, NY: Cornell University Library.

Sugiura, N.; Kurematsu, M.; Fukuta, N.; Izumi, N.; and Yamaguchi, T. 2003. A Domain Ontology Engineering Tool with General Ontologies and Text Corpus. *In* Evaluation of Ontology-Based Tools, Proceedings of the Second International Workshop on Evaluation of Ontology-Based Tools (EON2003) held at the Second International Semantic Web Conference (ISWC 2003), October 20, 2003. Vol. 87, Center for European Union Research (CEUR). CEUR-WS.org.

Sukhbaatar, S.; Szlam, A.; Weston, J.; and Fergus, R. 2015. *End-to-End Memory Networks*. arXiv:1503.08895. Ithaca, NY: Cornell University Library.

Sun, Y.; Wang, S.; Li, Y.; Feng, S.; Chen, X.; Zhang, H.; Tian, X.; Zhu, D.; Tian, H.; and Wu, H. 2019. *ERNIE: Enhanced Representation through Knowledge Integration*. arXiv:1904. 09223. Ithaca, NY: Cornell University Library.

Sutskever, I.; Vinyals, O.; and Le, Q. V. 2014. Sequence to Sequence Learning with Neural Networks. In *Advances in Neural Information Processing Systems*. arXiv:1409.3215. Ithaca, NY: Cornell University Library.

Thomson, B., and Young, S. 2010. Bayesian Update of Dialogue State: A POMDP Framework for Spoken Dialogue Systems. *Computer Speech & Language* 24(4): 562–88. doi.org/ 10.1016/j.csl.2009.07.003.

Tran, V.-K., and Nguyen, L.-M. 2018. Adversarial Domain Adaptation for Variational Neural Language Generation in Dialogue Systems. In *Proceedings of the 27th International Conference on Computational Linguistics*, 1205–17. Stroudsberg, PA: Association for Computational Linguistics.

Traum, D. R. 2000. 20 Questions on Dialogue Act Taxonomies. *Journal of Semantics* 17(1): 7–30. doi.org/10.1093/jos/17.1.7.

Tresp, V. 2000. A Bayesian Committee Machine. *Neural Computation* 12(11): 2719–41. doi.org/10.1162/ 089976600300014908.

Tseng, B.-H.; Budzianowski, P.; and Wu, Y.-C.; and Gašić, M. 2019. Tree-Structured Semantic Encoder with Knowledge Sharing for Domain Adaptation in Natural Language Generation. In *Proceedings of the 20th Annual Special Interest Group on Discourse and Dialogue (SIGdial) Meeting*, 155–64. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.18653/v1/W19-5920.

Turing, A. M. 1950. Computing Machinery and Intelligence. *Mind* 59(236): 433–60. doi.org/10.1093/mind/LIX.236.433.

van den Oord, A.; Dieleman, S.; Zen, H.; Simonyan, K.; Vinyals, O.; Graves, A.; Kalchbrenner, N.; Senior, A. W.; and Kavukcuoglu, K. 2016. *WaveNet: A Generative Model for Raw Audio*. arXiv.org/abs/1609.03499. Ithaca, NY: Cornell University Library.

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L.; and Polosukhin, I. 2017. *Attention Is All You Need*. arXiv:1706.03762. Ithaca, NY: Cornell University Library.

Vinyals, O., and Le, Q. 2015. *A Neural Conversational Model*. arXiv.org/abs/1506.05869. Ithaca, NY: Cornell University Library.

Walker, M. A.; Litman, D. J.; Kamm, C. A.; and Abella, A. 1997. PARADISE: A Framework for Evaluating Spoken Dialogue Agents. In *Proceedings of the European Chapter of the Association for Computational Linguistics (EACL)*. Stroudsberg, PA: Association for Computational Linguistics.

Walker, M. A.; Rambow, O.; and Rogati, M. 2002. Training a Sentence Planner for Spoken Dialogue Using Boosting. *Computer Speech & Language* 16(3–4): 409–33. doi.org/10.1016/ S0885-2308(02)00027-X.

Walker, M. A.; Stent, A.; Mairesse, F.; and Prasad, R. 2007. Individual and Domain Adaptation in Sentence Planning for Dialogue. *Journal of Artificial Intelligence Research* 30: 413–56. doi.org/10.1613/jair.2329.

Wang, A.; Pruksachatkun, Y.; Nangia, N.; Singh, A.; Michael, J.; Hill, F.; Levy, O.; and Bowman, S. R. 2019a. *SuperGLUE: A*

Stickier Benchmark for General-Purpose Language Understanding Systems. arXiv:1905.00537. Ithaca, NY: Cornell University Library.

Wang, A.; Singh, A.; Michael, J.; Hill, F.; Levy, O.; and Bowman, S. R. 2019b. *GLUE: A Multi-Task Benchmark and Analysis Platform for Natural Language Understanding*. arXiv:1804.07461. Ithaca, NY: Cornell University Library.

Weisz, G.; Budzianowski, P.; Su, P.-H.; and Gašić, M. 2018. Sample Efficient Deep Reinforcement Learning for Dialogue Systems with Large Action Spaces. IEEE/ACM Transactions on Audio. *Speech and Language Processing* 26(11): 2083–97.

Wen, T.-H.; Gašić, M.; Mrkašić, N.; Su, P.-H.; Vandyke, D.; and Young, S. 2015. Semantically Conditioned LSTM-based Natural Language Generation for Spoken Dialogue Systems. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, 1711–21. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.18653/v1/D15-1199.

Wen, T.-H.; Gašić, M.; Mrkšić, N.; Rojas-Barahona, L. M.; Su, P.-H.; Vandyke, D.; and Young, S. 2016. Multi-domain Neural Network Language Generation for Spoken Dialogue Systems. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 120–9. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.18653/v1/N16-1015.

Wen, T.-H.; Miao, Y.; Blunsom, P.; and Young, S. 2017a. Latent Intention Dialogue Models. 70: 3732–41. *Journal of Machine Learning Research*.

Wen, T.-H.; Vandyke, D.; Mrkšić, N.; Gašić, M.; Rojas-Barahona, L. M.; Su, P.-H.; Ultes, S.; and Young, S. 2017b. A Network-based End-to-End Trainable Task-oriented Dialogue System. In *Proceedings of the European Chapter of the Association for Computational Linguistics (EACL)*. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.18653/v1/E17-1042.

Williams, J. D.; Asadi, K.; and Zweig, G. 2017. Hybrid Code Networks: Practical and Efficient End-to-End Dialogue Control with Supervised and Reinforcement Learning. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, Volume 1 (Long Papers), 665–77. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.18653/v1/P17-1062.

Williams, J. D., and Young, S. 2007. Partially Observable Markov Decision Processes for Spoken Dialogue Systems. *Computer Speech & Language* 21(2): 393–422. doi.org/10.1016/j.csl.2006.06.008.

Witschel, H. 2005. Using Decision Trees and Text Mining Techniques for Extending Taxonomies. In *Proceedings of Learning and Extending Lexical Ontologies by Using Machine Learning Methods, Workshop at the Twenty-Second International Conference on Machine Learning (ICML '05)*. New York, NY: Association for Computing Machinery.

Wittgenstein, L. J. J. 1953. *Philosophical Investigations [Philosophische Untersuchungen*]. §43. New York, NY: MacMillan Publishing.

Wong, W.; Liu, W.; and Bennamoun, M. 2012. Ontology Learning from Text: A Look Back and Into the Future. *ACM Computing Surveys* 44(4): 1. doi.org/10.1145/2333112.2333115.

Wu, C.-S.; Madotto, A.; Hosseini-Asl, E.; Xiong, C.; Socher, R.; and Fung, P. 2019. Transferable Multi-Domain State Generator for Task-Oriented Dialogue Systems. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.18653/v1/P19-1078.

Xiong, W.; Droppo, J.; Huang, X.; Seide, F.; Seltzer, M. L.; Stolcke, A.; Yu, D.; and Zweig, G. 2017. Toward Human Parity in Conversational Speech Recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 25(12): 2410–23. doi.org/10.1109/TASLP.2017.2756440.

Yang, X.; Chen, Y.-N.; Hakkani-Tür, D.; Crook, P.; Li, X.; Gao, J.; and Deng, L. 2017. End-to-End Joint Learning of Natural Language Understanding and Dialogue Manager. In *Proceedings of the Institute of Electrical and Electronics Engineers (IEEE) International Conference on Acoustics, Speech and Signal Processing (ICASSP 2017)*, 5690–4. Piscataway, NJ: Institute of Electrical and Electronics Engineers (IEEE) Press. doi.org/10.1109/ICASSP.2017.7953246.

Yang, Z.; Levow, G.; and Meng, H. 2012. Predicting User Satisfaction in Spoken Dialogue System Evaluation With Collaborative Filtering. *IEEE Journal of Selected Topics in Signal Processing* 6(8): 971–81. doi.org/10.1109/JSTSP.2012.2229965.

Young, S. 2002. Talking to Machines (Statistically Speaking). *In* Proceedings of the International Conference on Spoken Language Processing (ICSLP). mi.eng.cam.ac.uk/reports/svr-ftp/SJYoung_ICSLP02.pdf.

Young, S., Gašić, M.; Keizer, S.; Mairesse, F.; Schatzmann, J.; Thomson, B.; and Yu, K. 2010. The Hidden Information State Model: A Practical Framework for POMDP-Based Spoken Dialogue Management. *Computer Speech & Language* 24(2): 150–74. doi.org/10.1016/j.csl.2009.04.001.

Young, S., Gašić, M.; Thomson, B.; and Williams, J. 2013. POMDP-based Statistical Spoken Dialogue Systems: A Review. *Proceedings of the IEEE* 101(5): 1160–79. doi.org/10.1109/JPROC.2012.2225812.

Yu, K.; Zen, H.; Mairesse, F.; and Young, S. 2011. Context Adaptive Training with Factorized Decision Trees for HMM-Based Statistical Parametric Speech Synthesis. Speech Communication 53(6): 914–23. Elsevier Science Publishers BV Amsterdam, The Netherlands.

Zhang, B.; Cai, Q.; Mao, J.; Chang, E.; and Guo, B. 2001. Spoken Dialogue Management as Planning and Acting under Uncertainty. In *Proceedings of Eurospeech 2001*. Baixas, France: International Speech Communication Association.

Zhang, J.-G.; Hashimoto, K.; Wu, C.-S.; Wan, Y.; Yu, P. S.; Socher, R.; and Xiong, C. 2019. Find or classify? Dual strategy for slot-value predictions on multi-domain dialog state tracking. arXiv preprint arXiv:1910.03544.

Zhao, T., and Eskenazi, M. 2016. Towards End-to-End Learning for Dialogue State Tracking and Management using Deep Reinforcement Learning. In *Proceedings of Special Interest Group on Discourse and Dialogue (SIGdial)*. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.18653/v1/W16-3601.

Zhao, T., and Eskenazi, M. 2018. Zero-Shot Dialogue Generation with Cross-Domain Latent Actions. In *Proceedings of the 19th Annual Special Interest Group on Discourse and Dialogue (SIGdial) Meeting*, 1–10. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.18653/v1/W18-5001.

Zhao, T.; Lee, K.; and Eskenazi, M. 2018. Unsupervised Discrete Sentence Representation Learning for Interpretable Neural Dialogue Generation. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL)*, Volume 1 (Long Papers), 1098–107. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.18653/v1/P18-1101.

Zhong, V.; Xiong, C.; and Socher, R. 2018. Global-Locally Self-Attentive Encoder for Dialogue State Tracking. In

Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, Volume 1 (Long Papers), 1458–67. Stroudsberg, PA: Association for Computational Linguistics. doi.org/10.18653/v1/P18-1135.

Zhou, H.; Huang, M.; Zhang, T.; Zhu, X.; and Liu, B. 2018. Emotional Chatting Machine: Emotional Conversation Generation with Internal and External Memory. In *Proceedings of the Thirty-Second Association for the Advancement of Artificial Intelligence (AAAI) Conference*. Palo Alto, CA: Association for the Advancement of Artificial Intelligence (AAAI) Press.

Nurul Lubis received a BSc degree (with distinction) in 2014 from the Bandung Institute of Technology, Indonesia, and MEng and PhD degrees in 2017 and 2019 from the Nara Institute of Science and Technology, Japan, under the scholarship support of the Japanese Ministry of Education, Culture, Sports, Science and Technology. She is a postdoctoral researcher in the Dialogue Systems and Machine Learning Group at Heinrich Heine University Düsseldorf, Germany. She has published peer-reviewed conference papers and journal articles on the topic of emotion in spoken language and affective dialogue systems for which she has received a number of awards, most recently the Nara Institute of Science and Technology Best Student Award in 2019. Her research interests also include RL and policy optimization for adaptive dialogue systems.

Michael Heck received a Diploma degree in Computer Science from the Karlsruhe Institute of Technology, Germany, in 2012, and a PhD degree from the Nara Institute of Science and Technology, Japan, in 2018. For his studies, he was supported by the Baden-Württemberg Scholarship, the Nara Institute of Science and Technology Global Initiatives Program, and the Nara Institute of Science and Technology International Scholarship. He is a postdoctoral researcher in the Dialogue Systems and Machine Learning Group at Heinrich Heine University Düsseldorf, Germany. Prior to that, he was with the Interactive Systems Labs at the Karlsruhe Institute of Technology (2013–2015), Augmented Human Communication Labs at the Nara Institute of Science and Technology (2015–2018), and the RIKEN Center for Advanced Intelligence Project, Japan (2018–2019). He has published peer-reviewed conference papers and journal articles on the topic of ASR, unsupervised learning, and the zero resource scenario. His current research focuses on dialogue system modeling, representation learning, and distributional semantics.

Carel van Niekerk received a BSc degree (with distinction) in 2016 in Mathematical Statistics from the University of Pretoria, South Africa, and a MSc degree (with distinction) in 2018 in Mathematical Statistics and Statistical Learning also from the University of Pretoria. He is a researcher and doctoral student, under Prof. Milica Gašić, at the Dialogue Systems and Machine Learning Group at Heinrich Heine University Düsseldorf, Germany. He recently received the award for the best MSc Thesis in Mathematical and Applied Statistics at the University of Pretoria in 2019. His current research interests are in statistical belief tracking and distributional dialogue state representation learning.

Milica Gašić received a BS degree in Computer Science and Mathematics from the University of Belgrade, Serbia, in 2006, and a MPhil degree in Computer Speech, Text, and Internet Technology and a PhD degree in Engineering from the University of Cambridge, U.K., in 2007 and 2011, respectively. She is a professor with the Heinrich Heine University Düsseldorf, Germany, and holds the chair for Dialogue Systems and Machine Learning. Prior to that, she was a lecturer with the Cambridge University Engineering Department. She has authored or coauthored around 60 journal articles and peer-reviewed conference papers. The topic of her PhD dissertation was Statistical Dialogue Modeling and she was the recipient of an Engineering and Physical Sciences Research Council PhD Plus Award for her dissertation. Most recently, she obtained a European Research Council Starting Grant and an Alexander von Humboldt Award.