Research Statement – Shafiq Joty (NTU-NLP)

The Internet is a great source of human knowledge, but most of the information is hidden in unstructured texts. As a researcher in Natural Language Processing (NLP), my goal is to add structures to this text to uncover relevant information, and to use it in developing useful applications. To this end, my research interests span two areas of NLP – (*a*) developing language analysis tools to understand human language, and (*b*) to build robust NLP applications to support end users. For (*a*) , I am interested in parsing texts with syntactic, semantic and discourse structures. For (*b*) , my interests lie in NLP applications that involve not only language understanding but also generation. These include machine translation, multilingual and multimodal processing, text summarization and dialogue systems. Recently, I am also looking into robustness of NLP models. One methodology emphasized throughout my research is to first identify the inherent semantic structures in a given problem, and then to develop structured machine learning models to exploit such structures effectively. My work has relied on deep learning for better representation of the input text and on probabilistic graphical models and reinforcement learning for capturing dependencies in the output.

1 Parsing tools

Natural language is ambiguous. As humans, we can easily disambiguate the meaning of linguistic units (words, phrases, sentences) as we read or listen. However, for machines it is difficult to understand without explicit representations of syntax, semantic and discourse. In my group, we build NLP tools to parse natural language in terms of its syntax (constituency and dependency parsing), semantic structures (named entities, semantic roles) and discourse structures (coreference, coherence). Previously, researchers used statistical machine learning methods to build such NLP tools. However, like other fields in AI, NLP has gone through a neural tsunami, and researchers develop mostly deep learning models for NLP. Our current approach also uses deep learning (e.g., pointer networks) extensively.

2 NLP applications

One of the main reasons to build parsing tools is that they can support building downstream NLP applications, e.g., named entities are used to build knowledge graphs, constituency and dependency trees are used to build sentiment analysis tools, and so on. Alternatively, one can also build applications in an end-to-end way with neural methods, which is the current trend in NLP. In my lab, we use deep learning to build end-to-end models for machine translation (MT), text summarization, question answering, and dialog systems. In MT, my lab is focusing on (*i*) new deep learning architecture (e.g., Tree Transformer), (*ii*) data augmentation through diversification, (*iii*) unsupervised and semi-supervised neural MT to support low-resource languages like Malay, Nepali, Hindi, Sinhala and Tamil, (*iv*) word translation (cross-lingual embeddings), and (*v*) discourse-based MT. In summarization and dialogue systems, we are focusing more on new architectures (submodular coverage) and few-shot learning methods for low-resource scenarios.

3 Multilingual and multimodal processing

A significant part of my current research focuses on *multilingual and multimodal processing*. MT is a multilingual application. We are interested in multilingual parsing tools (e.g., multilingual NER, parsing) as well as applications (e.g., multilingual question answering, summarization, natural language inference). Our methods focus on zero- and low-resource transfer learning scenarios, where our goal is to transfer knowledge from English. I also collaborate with the vision group at NTU to build multimodal applications including image and vidoe captioning models.

4 Adversarial robustness

With the aim to make our NLP tools and applications robust and ethical (minority or L2 speaker bias, gender bias), we also work on adversarial attacks (Morpheus, BITE) and defense systems to combat those attacks.