

令和元年度 情報工学コース卒業研究報告要旨

武田 (浩) 研究室	氏 名	小 野 瞭
卒業研究題目	牌譜情報を用いた麻雀の実況解説文の分類	
<p>ネット麻雀等の普及により、プロをはじめとした上級者の牌譜データを Web 上から取得することが可能となっている。しかし、初中級者が、上級者の牌譜データだけを見て、その意図を理解するのは難しいと考えられる。一方、動画共有サービスやインターネットテレビの普及に伴い、Web 上に公開されている実況解説付き麻雀動画も増加しており、その一部については牌譜データも公開されている。そこで、これらの対局動画、および牌譜データを学習データとして利用することで、麻雀の牌譜に対する実況解説文を自動生成することを考える。このとき問題となると考えられる点が2つ存在する。まず第一に、Web から取得できる実況解説の多くは牌譜と関係のない雑談を含んでいるが、これらは牌譜からの実況解説生成においてはノイズになると考えられる。第二に、麻雀の実況には盤面全体の情報に基づいて行われるものと、対局者のみが知ることのできる手牌情報に基づいて行われるものが存在するが、実況解説モデルの学習の際にはこれらを区別して扱う必要があると考えられる。そこで本研究では、麻雀の牌譜データからの実況解説生成の第一ステップとして、実況解説文のタイプ分類に取り組む。</p> <p>本研究ではまず、牌譜データと実況解説、および、実況解説文のタイプ情報を含むデータの作成に取り組んだ。このうち、牌譜データは Web 上で公開されているものを使用し、実況解説文は実況解説が行われた対局動画を文字起こしすることで取得し、これらを動画の時刻を基準に紐付けた。さらに、実況解説文を人手により、1) 対局者全員が知っている状況の実況、2) 注目している対局者のみが知っている状況の実況、3) 対局者が知らない状況の実況、4) 対局者が既に行った行為に関する解説、5) 対局者がこれから取ると予想される行為に対する解説、6) 具体的な打牌に関係しない戦略に関する解説、7) 局面と関係しない雑談、の7タイプに分類し、最終的に5半荘分、約2500文からなる文タイプ情報付き実況解説文で構成されるデータセットを作成した。</p> <p>続いて、作成したデータセットに基づく分類モデルの構築に取り組んだ。具体的には、本研究では2つのモデルを構築した。一つ目は、教師あり学習の代表的な分類器の一つであるサポートベクターマシン (SVM) に基づくモデルであり、分類対象の文を入力として、文タイプの分類を行う。二つ目は、ニューラルネットに基づくモデルであり、文ベクトルに対し、Bidirectional LSTM を用いて双方向の文脈情報を加味し、条件付き確率場 (CRF) により系列ラベリングを行う Bi-LSTM-CRF に基づくモデルである。これら二つのモデルに対し、入力がテキストのみである場合と、牌譜データも用いた場合の2つの条件で分類を行い、それらの精度を報告する。</p>		