

修 士 論 文 の 和 文 要 旨

研究科・専攻	大学院 情報システム学研究科 社会知能情報学専攻 博士前期課程		
氏 名	NGUYEN MINH THE	学籍番号	0751009
論 文 題 目	実世界行動マイニングによる意味ネットワークの構築 ~ Web を対象とした条件付確率場と 自己教師あり学習を用いた行動属性の自動抽出と評価 ~		
要 旨	<p> 計算機がユーザの行動意図を把握し、それに応じたサービスを提供することは、ユビキタスコンピューティングとソーシャルコンピューティングの双方において重要な課題とされている。得られた情報リソースから行動意図を認識するためには、行動の構成要素（行動属性）の把握と相関ルール（行動間の関係）が必要である。これを予め定義しておくことは手間がかかるだけでなく、未知の意図にも対応できず問題がある。一方、Web やセンサイベントから行動データを収集し、行動属性の抽出と行動間の相関ルールの発見を行う先行研究がある。しかし、抽出のための準備コストが大きいため、抽出できる行動属性が少ない、適用可能な文の種類が少ない、行動属性間の係り受け関係が十分に考慮されていないなどの問題がある。 </p> <p> そこで本研究では、通常の人間の行動は { 行動主, 動作, 対象, 時刻, 場面, 時間, 場所, 利用オブジェクト, 時制, 感情, 評価, 事実性 } の 12 種類の行動属性で表現できると仮定し、Web ページに記述された文に含まれるこれらの行動属性を全て抽出する (タスク 1)。次に、抽出した行動属性の正規化と意味的な分類を実施し、人間の行動を表す意味ネットワークを構築する (タスク 2)。この意味ネットワークは人間の行動の属性と行動間の関係を表す集合知である。そして、この意味ネットワークを参照し、得られた情報リソースから行動意図を推測する (タスク 3)。最後に、個々のユーザの行動意図に基づき、提供しようとする商品・情報のランキング手法を作り、そのユーザに最適な商品・情報 (行動パターンなど) を提供する (タスク 4)。 </p> <p> 本論文では、タスク 1 に焦点を当て、日本語の Web ページを対象とし、12 種類の行動属性のうち { 動作, 対象, 行動主, 時刻, 場面, 場所 } とこれらの行動属性間関係を自動的に抽出する手法を提案する。提案手法では、Web コーパスから相互関連抽出の最先端技術である O-CRF をベースにして、条件付確率場と自己教師あり学習を利用する。また、実験結果から、提案手法はシングルパスで 2 つから 6 つまでの要素 { 動作, 対象, 行動主, 時刻, 場面, 場所 } から構成される関係を自動的に抽出でき、高い抽出精度 (88.75%) を得た。また、{ 動作, 対象, 行動主, 時刻, 場面, 場所 } の抽出精度はそれぞれ { 99.38% ,98.10% ,94.37% ,95.65% ,92.94% ,92.31% } であった。 </p>		