

◇ 木曜日
◇ 7月8日
◇ 午後5時00分開始
◇ 情報科学研究科棟 2F 大講義室

話題：言語処理から知に迫る

--- ウェブ情報編集の未来像と自然言語処理の新展開

話題提供者： 乾 健太郎 教授

(システム情報科学専攻 知能情報科学講座 情報伝達学分野)

情報化社会の深化とウェブの爆発的な普及によって、ネットには日々莫大な量の情報や知識が流通し、その蓄積が進んでいる。こうしたネット上に散在する情報や知識を発掘し集約し再構成することによって、新しい価値を持った「知」を創出し、それを必要とするユーザに届ける。そうした情報や知識の高度な編集を自動化するソフトウェア技術の実現を我々は目標にしている。

さて、編集の対象となる情報や知識はその多くが日本語や英語のような自然言語（ことば）で表現されているのであるから、上の意味の「編集」を実現するためには、言語で伝達される情報、すなわち言語の意味を解せるように計算機自身がかしこくなる必要がある。これまでの計算機は言語の理解に必要な言語知識や世界知識が決定的に不足していたため、これはすぐには手の届きそうにない遠い目標であった。しかし、数億から数十億文書規模の超大規模言語データがウェブから入手可能になった今日、それらのデータから大量の言語/世界知識を計算機自らが獲得し、これまでよりも一段深い頑健な言語理解を実現するという、新しい可能性が見え始めている。

本講演では、(a) ブログに書かれた人々の意見や経験の情報を広く集めて共有の知とする「経験マイニング」、(b) ウェブ上の言明間に潜在する同意、対立、根拠等の隠れた論理的関係を解析し、言論空間を再構成する「言論マップ」という2つの応用例に、近未来のウェブ情報編集を目指した取り組みを紹介するとともに、それを支える自然言語の意味解析に関する最近の研究成果を報告する。情報の海からの自動知識獲得が機械に知をもたらし、賢くなった機械が情報の海から新しい知を創出する

--- そんな可能性を参加者の皆さんとともに考えたい。---

第38回 情報科学談話会