これからの自然言語処理

これからのコンピュータ技術は，超並列，高信頼の方向に向かって大きく進化するだろう．21 世紀の到来とともに，われわれは従来とは比較にならないほど大量のコンピュータパワーを手に入れることができる．また，従来とは比較にならないくらい大量の情報を安価に蓄積することが可能になり，大規模な知識ベースの構築が各所で進む．そして超高速情報ネットワークが構築され，いわゆる情報インフラが整備された新しい情報化社会に突入するだろう．高速な情報の転送を可能にする光ファイバーが各家庭に入り込み，超高速情報ネットワークが国内外に張り巡らされ，情報通信，情報処理に関する空間的，時間的な制約の多くは取り払われることになるだろう．

昨今のマスコミを巻き込むマルチメディアフィーバーは，主として次の三つ，コンピュータパワー，大規模な知識ベース，超高速な情報ネットワークを統合したところに巨大なビジネスチャンスが生まれ，新しい情報産業としてのマルチメディア産業が急激に立ち上がるということを多くの人が予感していることの反映でもある．ここで，マルチメディア産業が巨大な産業に成長するためには，どのような条件が満足されなければならないかを考えてみたい．

大学間，病院間，企業間，オフィス間などといった特定の拠点間が超高速情報ネットワークで接続されることにより，教育形態，社会構造，経済構造が部分的に変わりはする．それらがマルチメディア産業の成長に寄与するということも事実である．しかしそれだけでは，21 世紀の初頭に，巨大なマルチメディア産業が勃興するとは到底いえない．おそらく，各家庭にまで張り巡らされた超高速情報ネットワークを，一般個人が自在に使いこなせる時代がくるかどうかが，巨大なマルチメディア産業勃興の鍵になるのではないか．そのために，われわれがなすべきことは何だろうか?

専門家ではなく家庭にいる一個人が，ネットワークを介した高度な機能をもつシステムを自在に使いこなす時代がくるかどうかは，システムと人間との間のインタフェース技術の確立に依存していると思う．自然言語による対話，音声による対話は，もっとも自然で友好的なインタフェースを提供する．その意味で，音声や自然言語に関する優れたソフトウェアの開発が特に重要になる．これらはハードウェアというよりソフトウェアの問題であることに注意したい．マルチメディア技術に関連して，ソフトウェア技術の重要性がわが国ではしばしば見過ごされてる．ことの本質はソフトウェアにあるということをここで強調しておきたい．

いずれにせよ，音声や自然言語による優れたインタフェース技術の確立は，巨大なマルチメディア産業の勃興と密接に関連している．自然言語処理の研究者，技術者の果たす役割は，今後ますます重要になる．われわれはそうした自覚を持ってこれからの自然言語処理に取り組む必要がある．

これまで述べてきた，大量のコンピュータパワー，大規模な知識ベースや超高速情報ネットワークの構築は，自然言語処理技術そのものにも新しい視点を提供する．以下では自然言語処

理への超並列計算技術の応用という側面に焦点をあてて論じてみたい．

まずはじめに超並列計算技術は，計算速度の向上に関して万能薬のように考える傾向があるが，超並列計算は万能薬ではないということをあえて指摘しておきたい．実際，超並列計算が有効に働かない問題が多数存在する．たとえば並列計算同士の通信コストのために，並列計算の実がほとんど上がらないといった問題がある．自然言語処理に関して，超並列計算が有効な問題は何で，何が有効でないかを今後明らかにしておく必要がある．このことは，自然言語処理に大量のコンピュータパワーを導入し，力づくで問題を解こうとするアプローチの限界をも示唆することになると思われる．

力づくのアプローチをとろうとする場合，意味処理などの困難な問題に深入りせず，比較的浅いレベルの処理にとどめることがよく行われる．意味処理・文脈処理技術が現在未熟であるだけに，現段階では，処理を浅いレベルにとどめることが，一見成功したように見えることがある．しかし，長期の観点に立つなら，困難ではあっても，自然言語処理の本質は，深い意味処理，文脈処理のための理論と技術の確立にあることを忘れてはならない．そうした方向の新しい理論，新しい技術の確立がこれからの自然言語処理技術に求められている．大量のコンピュータパワーと大規模な知識ベース (たとえば辞書やコーパスなど) は，こうした困難な技術の確立にも有用であるという視点を忘れてはならないと思う．

自然言語には一定の秩序がある．秩序を一般的に説明する理論もなしに，これからの自然言語処理に明るい未来はないだろう．豊かなマルチメディア社会の実現もないと思う．大量のコンピュータパワーを背景に，意味処理，文脈処理といった困難な問題の解決に努力を傾注する必要がある．

田中穂積
東京工業大学大学院
情報理工学研究科