

# AGI への道程：2017 年版

中川裕志<sup>1,2</sup>

<sup>1</sup> 東京大学 情報基盤センター

<sup>2</sup> 理化学研究所 革新知能統合研究センター

## アブストラクト：

汎用人工知能(AGI)が実現する時期については諸説があり、例えばカーツワイルは 2045 年と想定している。その真偽はさておき、現在までに明らかになっている情報から AGI にまつわる技術状況を検討することは、AGI への道程を占うために役立つであろう。この発表では、このような目的に従って、以下の議論を展開する。

- (1) 目的特化型的人工知能を組み合わせることで AGI を作ることの困難さはすでにポストロムに指摘されているが、ここでは多種類の目的特化型人工知能の集合から解決すべき目標に適した目的特化型人工知能を探索することの困難さについて述べる。
- (2) Brooks のサブサンプリング・アーキテクチャを基礎として、AGI に関係の深い自意識を実現する可能性の候補として、最近、知られるようになった長距離ニューロンを組み合わせる方法について考察する。
- (3) 単一の AGI が超知能として他の AGI および人類を支配ないし駆逐するような人工知能脅威論は的外れであることを論ずる。

以上のように AGI ないし超知能が実現しそうもないにもかかわらず、現在の弱い人工知能ですら人間社会においては実質的な脅威になるケースが散見されるという現状の俯瞰をもって、まとめとする。