

RUSAGI 汎用人工知能研究所

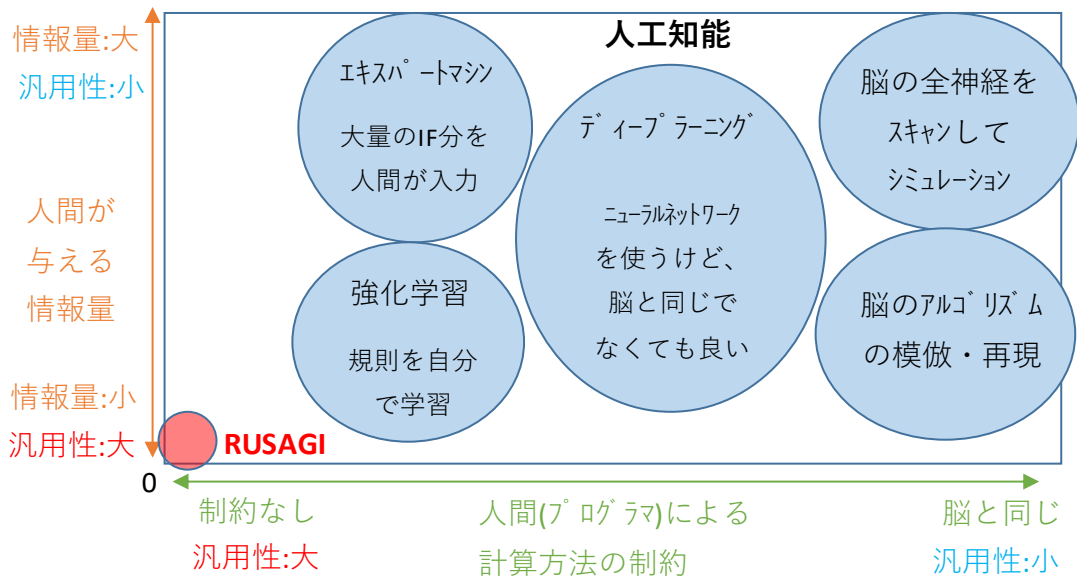
(Recursive Universal Selective Artificial General Intelligence)

© RUSAGI 汎用人工知能研究所 <https://rusagi.com/>

目次

1. RUSAGI の立ち位置.....	2
2. 人工知能の評価方法.....	3
3. 計算量と精度.....	3
4. 自己改良.....	4
5. アンリミテッドブレインワークス(脳超知能エミュレーション).....	5
6. 特化型と汎用型の研究の進め方.....	6
7. 研究者とフレーム.....	6
8. RUSAGI と他の人工知能の比較.....	7
9. RUSAGI 研究所と他社との比較.....	7
10. 完成時期.....	8

1. RUSAGI の立ち位置



グラフ横軸は計算方法の制約を表します。右端は、脳を忠実に再現しようというもので、左端は、脳との類似性を全く求めません。RUSAGI は左端に位置します。脳に似せれば人間の知能に近づくだらうという安心感がありますが、計算方法を一切制約しないことで、脳を超えることも可能になります。

		メリット	デメリット
人間による 計算方法の 制約	脳を模倣	脳に似せるのを 目標とすれば、 人間の知能に 近づく予想できる	プログラムが決めた処理が どんな状況でも 適しているとは限らない (脳も最適とは限らない)
	計算方法の 制約なし (RUSAGI)	AI自ら、適した処理 を選択する余地有。 (脳を超えるポテンシャル)	知能とは何なのか 理解していないと、 目標が決められない

グラフ縦軸は計算方法の制約を表します。上の方は、あらかじめ知識やハイパーパラメータを人間が詰め込むことで知識を向上させようというアプローチで、下の方は、情報の少ない状態で開始し、汎用性を高めるアプローチです。RUSAGI は下端に位置します。

		メリット	デメリット
人間が 与える 情報量	大 (ビッグデータ)	AI自ら、規則を学習する 処理が不完全でも良い (量で補える)	あらかじめ想定した 状況でのみ通用 (特化)
	なし (RUSAGI)	どんな状況でも通用 (汎用)	AI自ら、規則を学習する 処理が完全でなくては ならない

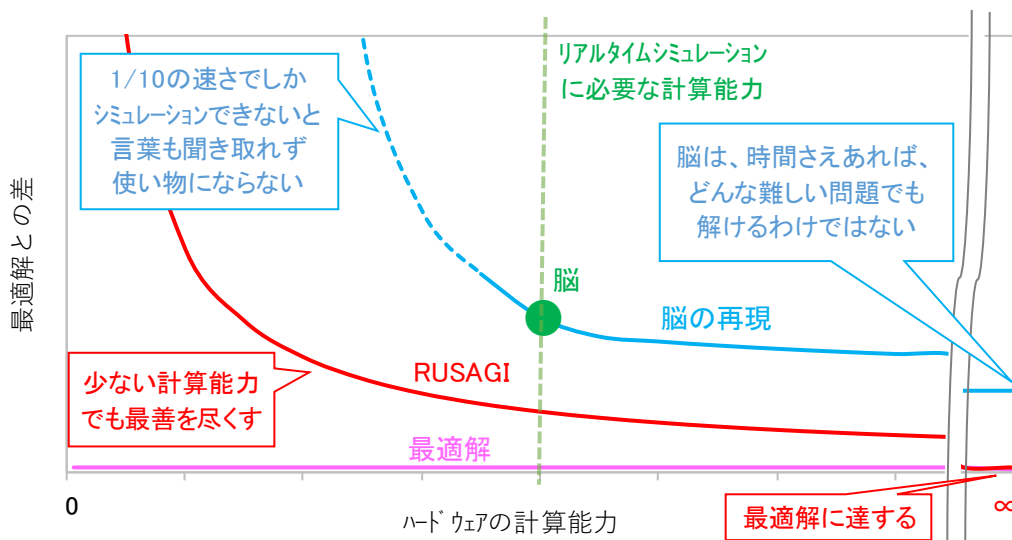
2. 人工知能の評価方法

人工知能の評価方法

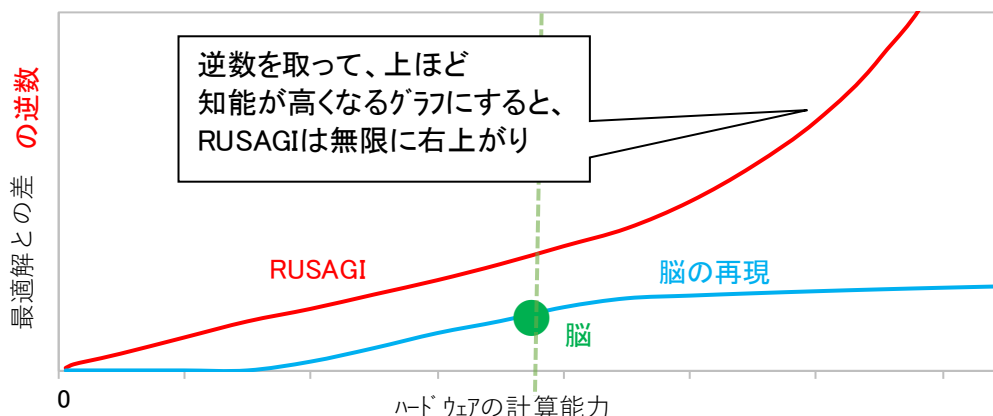
基準	実験して評価			実験せず評価
	処理が脳と同じか？	振る舞いが脳と同じか？	試験成績が良いか？	定義通りか？
実験例	ある映像を見たときの脳の活動とシミュレーション結果が一致するか	チューリングテスト(会話)	問題の正答率 ゲームの勝率	バグがなく 知能の定義(仕様) 通りのアルゴリズムと なっていれば、 実験する必要ない (RUSAGI)
汎用性の評価	5000兆種類の映像で完全に一致→汎用？	5000兆人が人間と区別できなかった→汎用？	5000兆種類のゲームで人間に圧勝→汎用？	
	映像認識、会話、ゲームをする知能はあるが、実験したこと以外も出来るかはわからない。			

何らかの実験を行って、試験結果が良ければ、知能が優れていると評価するのが一般的です。しかし、いくら様々な実験を行ったとしても、それらの実験を行った特定の条件下でのみ知能が優れるとしか言えません。RUSAGIは、あらゆる状況で成り立つ真の汎用性（一般性）を示すため、まず知能を定義して、アルゴリズムがその定義に従っていると確認することで、実験するまでもなく汎用的な知能であることを示します。

3. 計算量と精度



脳を忠実にシミュレーションする方法では、計算能力が足りないと、リアルタイムでは使い物になりません。また、計算能力を増やしても、脳の物理的な制約まで再現しているの、どんな問題でも解けるわけではありません。RUSAGIは、計算能力をどの処理に宛がうのか最適な選択をしようとするため、計算能力が足りなくてもベストを尽くします。また、計算能力が上がれば、いくらでも精度があがり、計算能力が ∞ なら、どんな問題でも最適解に達します。

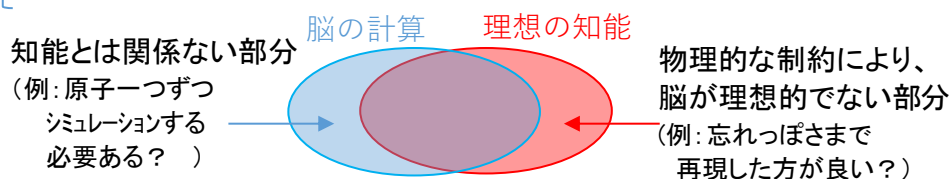


知能の大きさを表すために、最適解との差の逆数をグラフの縦軸にすると、RUSAGI の知能は、計算機の能力が増えるのに従って、無限に向上します。超知能といえます。

4. 自己改良

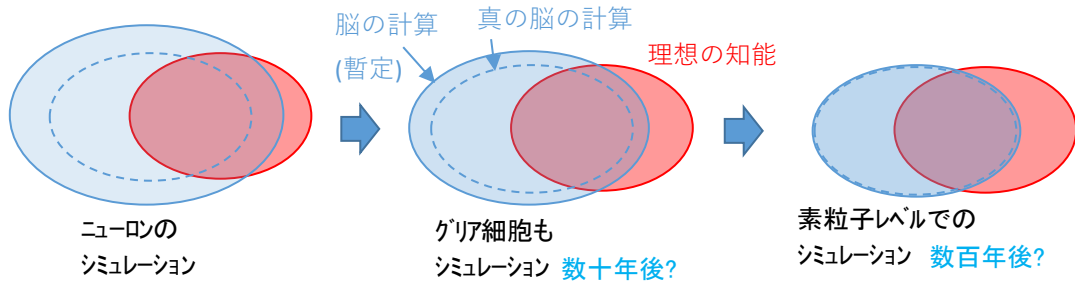
脳の完全な再現ができれば、無限に自己改良し知能が向上する？

- ・同じ計算内容のまま、数学的に速いアルゴリズムに改良…可能だが効果は限定的
- ・知能の正体が明らかになり、より知能が高くなるように計算内容を改良できる？



脳の完全再現ができて、脳の計算が明らかになるだけで、理想的な知能の正体は分からない。(推測するしかない)

脳を完全にシミュレーションできたとしても、自己改良によって超知能になることはありません。脳が処理している内容が明らかになりますが、知能とは関係ない処理も含んでしまっているため、知能とはなにか明らかにはなりません。知能とは何か考察して、定義してやらなければ、無駄な処理を省いて最適化できません。しかし、知能を理解するのに、脳内を解剖して調べることは必ずしも必要ではなく、RUSAGI は、知能の振るまいから知能を定義しています。

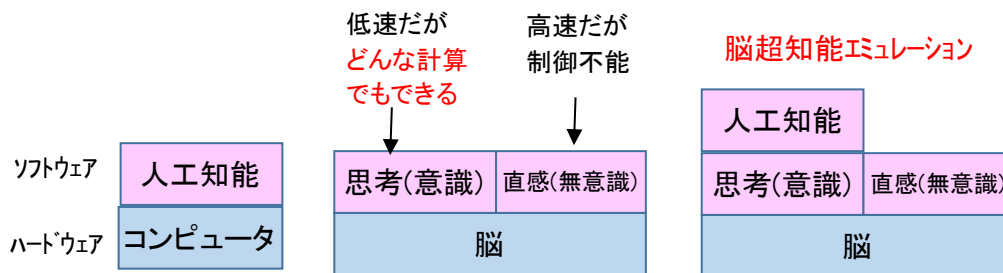


徐々に脳の計算が明らかになるが、脳を超える知能を目指すためには、どこかの時点で、脳の計算等を元に、理想的な知能を推測して定義する必要がある

シミュレーション可能になるまで何十年も待っててもしょうがない。

- ➡ 脳内を覗かずとも、知能の振る舞いを良く観察すれば、脳内で何が起きているのか、知能の正体を予想可能なはず。知能の正体を推測して、合理的な定義をしておこう (RUSAGI)。

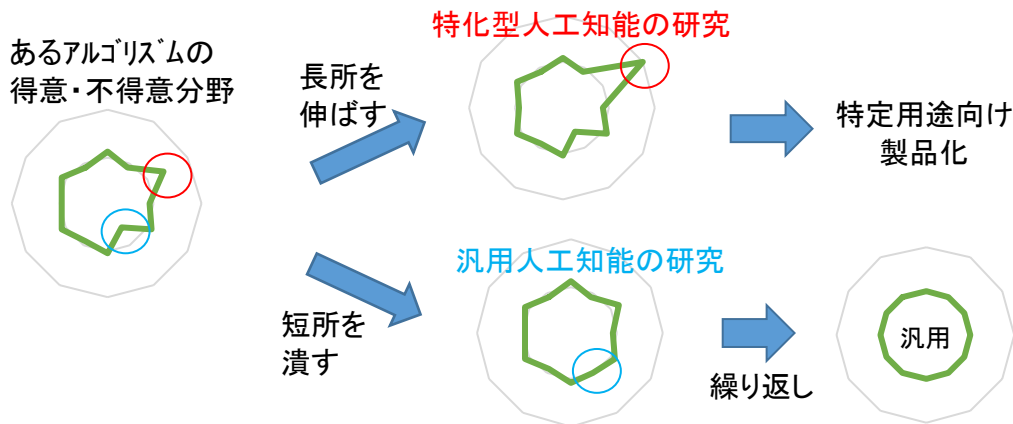
5. アンリミテッドブレインワークス(脳超知能エミュレーション)



- ・知能のアルゴリズムが分かれば、意識上で知能をエミュレートできる。
- ・低速だが、フレームを制御し無制限な発想や、思考のデバッグが可能になる。
- ・知能の研究が進むほどエミュレーションの精度も上がり、再帰的に研究は加速する。

RUSAGIはソフトウェアです。脳をハードウェアとしてRUSAGIを動かすこともできます。ネイティブよりも処理速度は落ちますが、理想的思考により、フレームや常識にとらわれることがないため、いつかは必ず最適解に到達します。RUSAGIの研究が進めば、エミュレーションの性能もあがるため、再帰的な改良で研究が加速します。

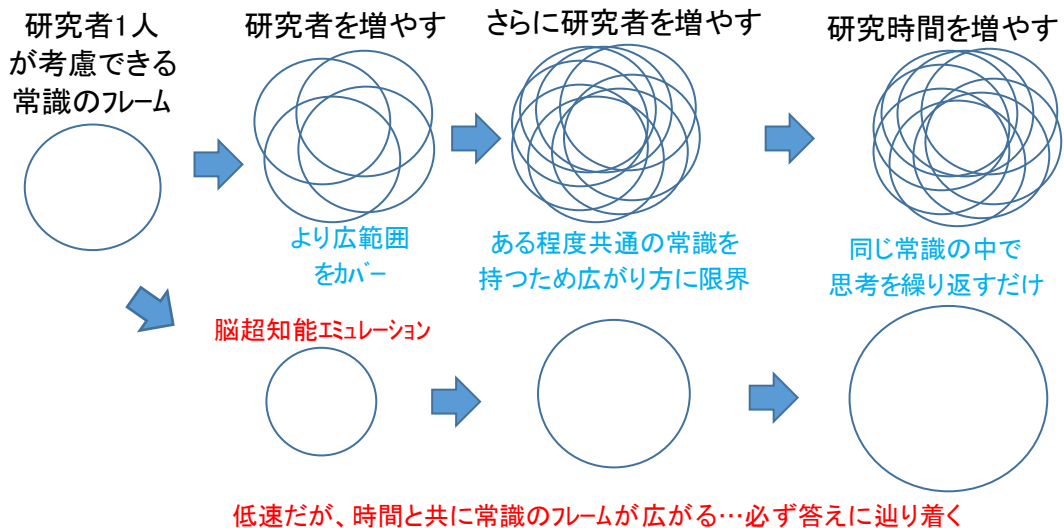
6. 特化型と汎用型の研究の進め方



- ・完成に時間が掛かるが、途中での成果のアピールが困難・・・研究費が得られない
- ・複数人での研究では、お互いに欠点を指摘し合うばかり・・・モチベーション維持困難

特化型 AI は、あるアルゴリズムが、ある用途で良い試験結果を得られれば、その長所を伸ばして、製品化を目指します。一方、RUSAGI は、ある状況でうまくいかないという短所を探して、延々と改善します。良い試験結果でアピールということができないため、研究費を得るには不利です。

7. 研究者とフレーム



研究者を増やせば増やすほど、研究が加速するかといえば、そうではありません。同じような常識に囚われて、その常識内でいくら考えても、同じ結果しか得られません。一方、RUSAGI 汎用人工知能研究所では、脳超知能エミュレーションにより常識の範囲を徐々に広げられます。人海戦術は、1を100に増やす仕事には有効ですが、0を1にする仕事には限界があります。

8. RUSAGI と他の人工知能の比較

	深層学習	脳の模倣	RUSAGI
目標	特化型	汎用型	
評価基準	試験成績		理論的裏付け
研究手法	仮説立案、実機実験		思考実験
開発速度	○	×	×
必要資金	△	×	○
ハードウェアのボトルネック	○	×	○
成果の分かりやすさ	◎	○	×
脳との互換性	△	◎	×
研究費の得易さ	○	○	×
汎用性	×	○	◎

総合的にみて、ある程度のことさえ出来ればよいなら、脳をあまり意識しないディープラーニングが優秀です。人間レベルを目指すなら脳の模倣が王道ですが、脳のスキャンやシミュレーションをどこまで詳細にやればいいのか、どれだけ計算能力が必要でいつ可能になるのか分かりません。RUSAGI なら、ハードウェアのボトルネックは存在しません。また、汎用人工知能が完成すれば、特化型人工知能の研究は無駄になってしまうため、最終的に脳を超えるものを目指すなら、RUSAGI の研究が近道になります。

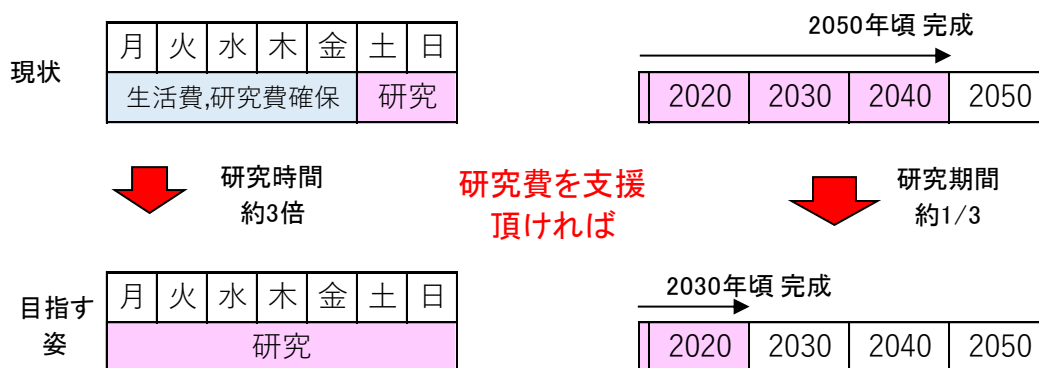
9. RUSAGI 研究所と他社との比較

	他社が 追従した場合	RUSAGI 汎用人工知能研究所
真の狙い	脳と同等の能力で、 労働を代替し、 多大な利益を得る	脳を超える究極の 人工知能の実現と 科学技術の進歩
戦略	株主が短期的に確実に利益が得られる特化型の方が良いと、方針変更の可能性。脳を超えなくても労働代替できるので、脳を超える研究をさせてもらえない。 ×	汎用型が完成すれば、特化型は全て無駄になるため、汎用型の研究に集中。利益が目的ではないため方針変更はありえない。 ○
戦術	実験結果で進捗度合いを表現するといった現行の研究・開発手法が使えないため、うまく組織が機能しない。 ×	非公開の技術がある。問題点を把握している。(問題の本質を見極めることは、解決するよりずっと難しい) ○

RUSAGI は、実験により正しさを示すことができないため、かなり確証性のある基本的な部分しか公開しておりません。より進んだ仮説を無数に立てておりますが、どういった仮説が、どういった状況でうまくいかないかという問題点の情報が無数に蓄積されています。問題の本質を見抜いて明らかにすることは、明らかになった問題を解くよりもずっと難しいことです。他社が追い付こうとしても難しいでしょう。そもそも、利益を追求する企業にとっては人間と同レベルで十分であるため、超知能を目指さないかもしれません。他社が人間レベルの AI に手間取ってれば、一足飛びに超 AI である RUSAGI が先に完成するでしょう。

10. 完成時期

RUSAGI汎用人工知能研究所



研究開始当初は、いつ完成するか全く目途が立たず、途中での成果もアピールし難い基礎研究のため、研究費を得ることはできません。それでも究極の人工知能を創りたいため、平日は働いて生活費を稼ぎ、休日に研究を続けてきました。さまざまなブレイクスルーを経て、ついに完成の目途が付きました。こういった問題点を抱えていて、あと何ができれば完成なのか把握できています。また、壁にぶつかっても、必ず越えられる思考法も身につきました。それでもなお多くの年月が必要なほど難しい研究です。そのため、365日研究に専念できるように研究費の支援をお願いしております。研究できる時間が3倍になれば、完成までの期間は1/3になります。2050年頃完成のところ、2030年には完成させられる見込みです。