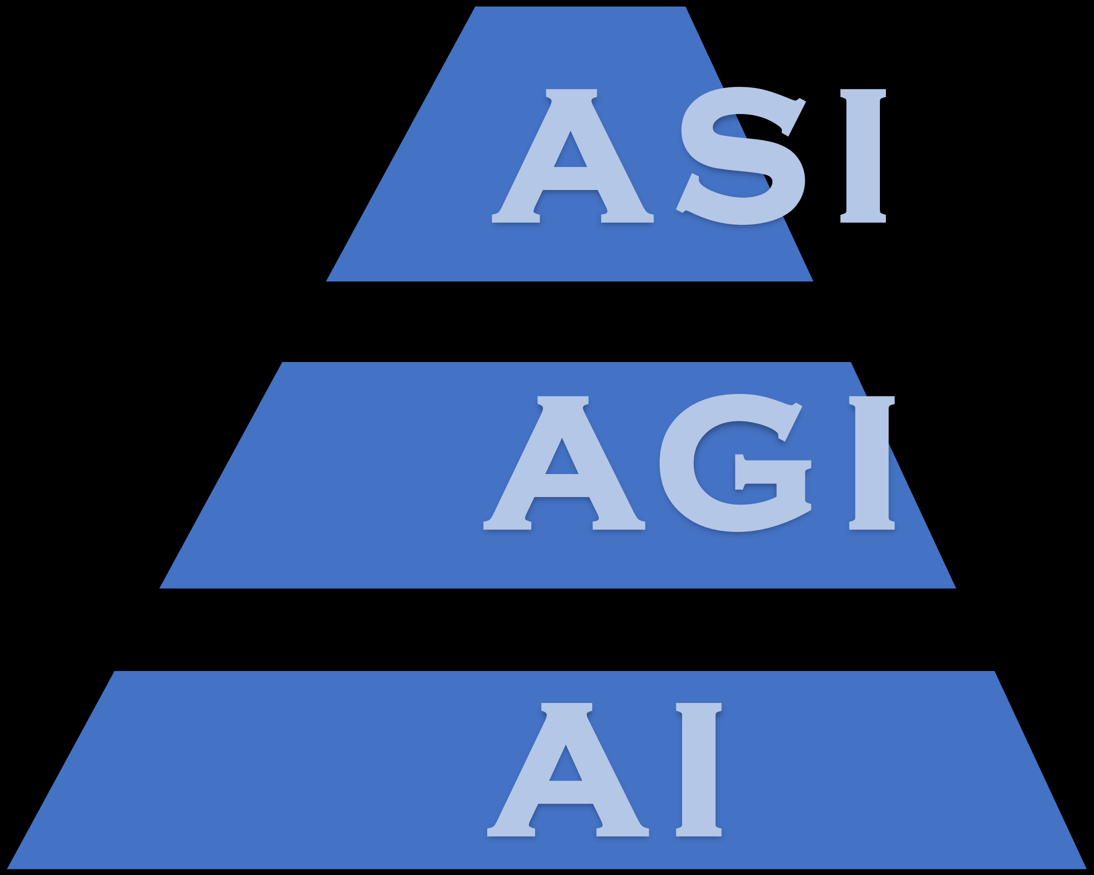
究極の AGI

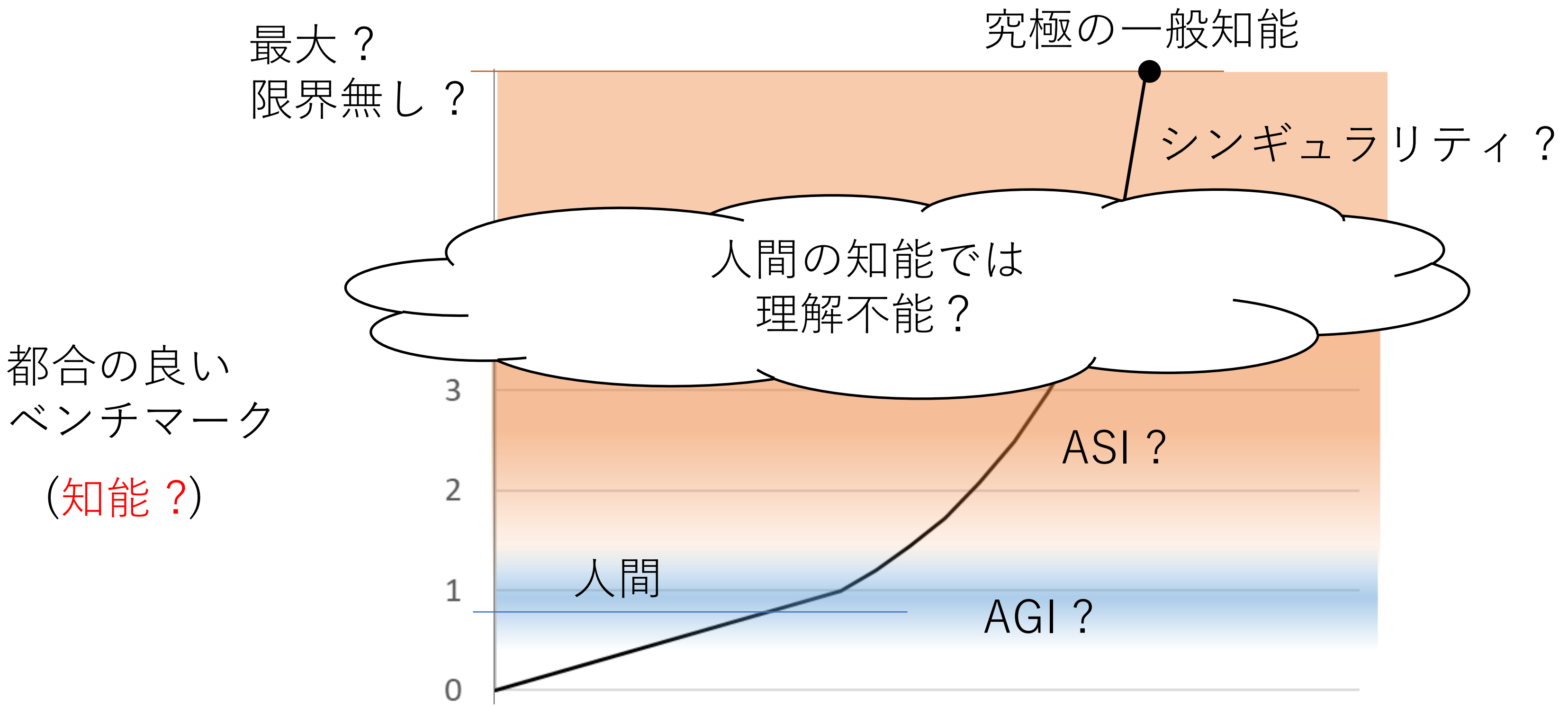


理論 ①

知能の定義

目次

知能の定義の必要性



知能を正しく定義すれば、究極の知能も理解できるはず

一般的には、AGIは人と同等、ASIは人を超える知能を意味します。
しかし、知能の客観的な定義はありません。
都合の良いベンチマークテストを行って、既に人間同等だからAGIという人もいます。
自己改良によって、知能は無限に良くなると考える人もいます。
人間より優れた知能は、人間には理解できないという人もいます。
しかし、知能の定義を曖昧にしたまま、なぜそのようなことが断言できるのでしょうか？
知能を適切に定義すれば、究極の知能を人間が理解することもできるはずです。

知能は有限

質問

1+1=?

回答選択肢

① 1

② 2　← 最適解
（どんなに優れたAIであっても）

③ 3

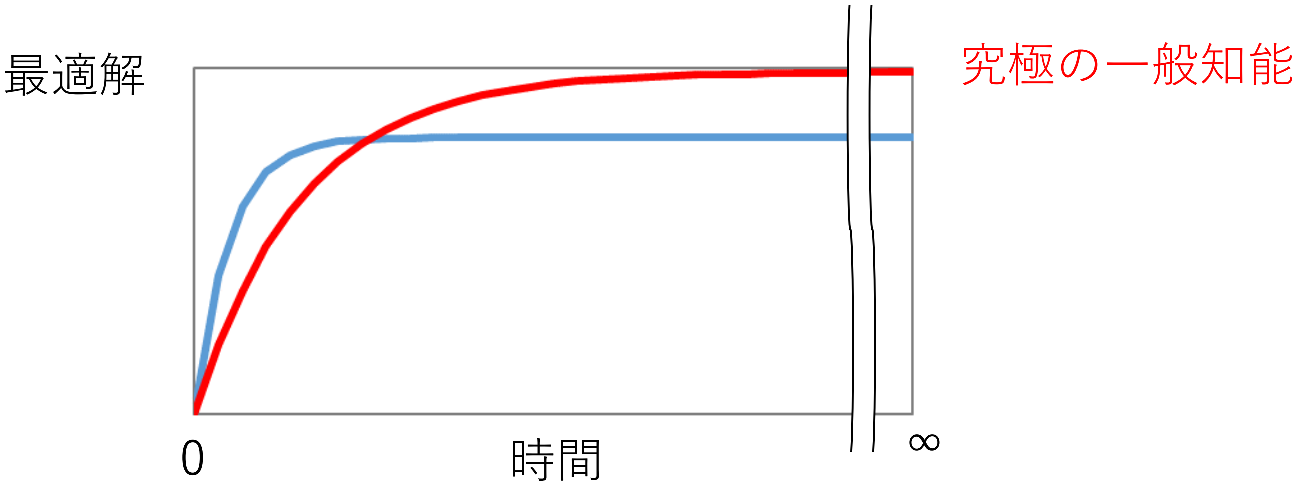
④ 4

この質問に限れば、2と答えることができれば、究極の知能です。

これより良い回答はない → 知能は無限ではない。

知能が決して無限ではないというのは、簡単な思考実験で理解できます。
知能が無限に改良される余地があるというのは、ある解よりも優れた解が常にある場合です。
「1+1=?」という質問に、1,2,3,4の4択から回答する場合を考えます。
どんなに優秀なAIであっても、「2」と回答するだろうと容易に予想できます。
もし、この質問しかされないのなら、最適解の「2」と出力すれば、最も優れた知能といえます。
この質問にしか答えられないなら、汎用性はなく、一般知能とはいえません。
全ての質問に対する最適解を定義できたなら、それを出力すれば究極の汎用知能になります。

究極の一般知能の定義



「究極の一般知能」の定義：どんな質問でも、必ず最適解が得られる能力

ここでは、回答までにかかる時間は無視しています。
出来の悪い回答なら、いくら早くてもうれしくありません。
計算速度はハードウェアの性能を上げれば、改善できます。
回答の速度と精度は、ある程度はトレードオフの関係にあります。
しかし、アルゴリズムが悪ければ、いくら計算量を増やしても、精度は頭打ちになります。
そこで、時間がいくらかかっても良いので、最適解が得られるなら、究極の知能とします。
どんな質問でも、必ず最適解が得られる能力を、「究極の一般知能」と定義します。

学習の分類

	事前学習	怠惰学習
学習タイミング	質問される前	質問された後
回答速度	速い	遅い
回答品質	悪い	良い
質問を知った上で最適な学習	×	○
究極の一般知能に必要？	×	○

学習という言葉は、知能と似た意味で使われることがあります。
一般的な意味では、行動パターンを変化させる変化が学習です。
学習は、事前学習と、出題後に行われる怠惰学習に分けられます。
事前学習は、直感的な素早い思考に相当します。
怠惰学習は、じっくり思考することに相当します。
いくら時間が掛かっても良いなら、事前学習する必要は一切ありません。
質問が分かってから、その質問に最適な方法で怠惰学習する方が良いからです。

学習率

	機械学習 (事前学習)	意識的な思考 (怠惰学習)
使用するデータ	全て	n (関係あるいくつか)
学習率	0.1%	1/n
1データの影響	0.1%	1/n

従来のAIは、学習に大量のデータが必要です。
一方で、人間は、僅かなデータから推論することもできます。
意識上へ数個のデータだけをロードして、その数個のデータのみから推論できます。
機械学習の学習率としては、0.1%程度が一般的です。
推論結果に1つのデータが与える影響は、0.1%程度です。
質問を受けた後に考えるなら、関係性が大きい数個のデータを選んでから学習できます。
n個のデータだけを使うなら、1つのデータの影響は1/nなので、学習率も1/nです。
質問を受ける前に考えるなら、全てのデータから少しずつ均等に学習するしかありません。

知能は過程

質問:　1+1=?

①1
②2
③3
④4

ランダム
選択

② 2 (正解)

知能ゼロ

---

①1
②2
③3
④4

非ランダム
選択

①1
②2
~~③3~~
~~④4~~

ランダム
選択

① 1 (不正解)

より良い知能

知能
の過程

非知能
の過程

知能は過程であって
結果ではない

知能の水準を数値化したい場合、何らかのベンチマーク試験を行うのが一般的です。
しかし、正解が多い方が知能が優れているとは限らないことに注意が必要です。
ランダムに4択から選んでも、偶然に正解する可能性があるからです。
4択から2択まで絞り込んだ後で、どちらかをランダムに選んだ結果、不正解する場合もあります。
結果的に不正解でも、2択まで絞り込めたのだから、完全にランダムな回答よりは優秀です。
知能の水準は、結果ではなく、過程から評価する必要があります。
4択から2択まで絞り込む過程は知能ですが、2択からランダムに選ぶ過程は知能ではありません。
完全にランダムに回答した場合は、結果とは無関係に知能は0といえます。

知能の定義

質問:　ランダムに回答せよ。（評価基準: ランダムなら最高評価）

① 1
② 2
③ 3
④ 4

ランダム
→
選択

② 2　最適な回答

~~知能ゼロ~~

「知能」の定義：評価基準に合った出力する能力

「究極の一般知能」が定義できるなら、「知能」も定義できます。
先ほどの例で、ランダムな回答に知能は無いといいましたが、例外があります。
「ランダムに回答せよ」と質問された場合は、ランダムに答えるのが最適解です。
何が最適解なのかという評価基準は、出題者が決めることです。
評価基準に合った出力する能力のことを、知能と定義します。
どんな質問でも、最高評価の最適解を出力ができる能力が、究極の一般知能です。

質問と最適解の定義

「質問」の定義

- 回答選択肢
- 評価関数
- 利用可能な情報

「最適解」の定義:
「利用可能な情報」から特定できる範囲で、
「評価関数」で最高評価の「回答選択肢」

次に、「質問」と「最適解」を定義します。
「質問」とは、「回答選択肢」と「評価関数」と「利用可能な情報」だと定義します。
「評価関数」は、二つの「回答選択肢」の入力に対して、どちらが優れているか、一意に出力します。
ランダムな評価することは認められません。変化することも許されません。
「利用可能な情報」から特定できる範囲で、「評価関数」で最高評価の「回答選択肢」を「最適解」と定義します。

最適な推論方法の必要性

評価関数 (Action 1, Action 2) {　　… 無効な評価関数

return “行動の結果、未来に受け取る報酬 が大きい方”

}

未知

評価関数 (Action 1, Action 2) {　　… 有効な評価関数

return “行動の結果、未来に受け取る報酬の推測値 が大きい方”

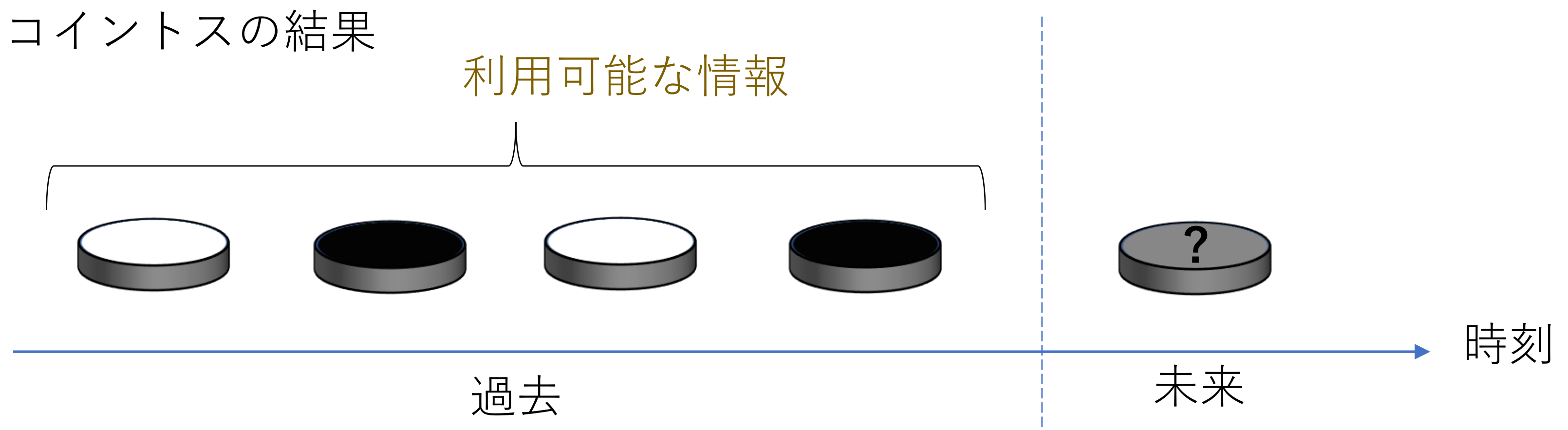
}

計算可能。
ただし、推論方法により異なる

「最適な推論方法」を決める必要がある

評価関数は、優劣の評価さえできれば、自由に設定できます。
良くあるロボットの目的設定を例に考えます
未来に受け取る報酬が多い行動が高評価だとします。
ですが、未来に受け取る報酬は未知なので、この評価関数は有効ではありません。
ただし、推測することはできます。
未来に受け取る報酬の推測値が多い方とする評価関数なら有効です。
しかし、推測の仕方によって推測値は変わります。
「最適解」を決めるには、「最適な推論方法」を決めておく必要があります。

演繹と帰納



答えが、
利用可能な情報に含まれる

演繹

数学の公理で一意に決まる

答えが、
利用可能な情報に含まれない

帰納

数学の公理がない

推論は、演繹と帰納に分けられます。
「利用可能な情報」の中に答えが含まれているときの推論は、演繹といいます。
評価関数に使われる値が全て既知なら、演繹できます。
一方で、演繹できない場合の推論を、帰納といいます。
例えば、コイントスの過去の結果を問うのは演繹で、未来の結果を問うのが帰納です。
演繹推論の結果は、数学的な公理から一意に決まります。
一方で帰納推論では、特定の推測値を正当化するような、数学の公理は存在しません。

確証性の原理

確証性の原理:

「関係のある観測の数が増加するほど、確証性が増す」

「観測の数」

「関係性の度合」

「確証の増加量」

これらの関係を定めなければ、
「最適な推論方法」は定められません。

帰納推論では、確証性の原理というものが知られています。
「関係のある観測の数が増加するほど、確証性が増す」というものです。
「観測の数」と「関係性の度合」ど「確証の増加量」の関係までは語られていません。
これらの関係を定めなければ、「最適な推論方法」は定められません。
極端な例を仮定すれば、これらの関係が見えてきます。

確証性の極端な例

◆「確証性」=0 : 全く推論できない。

◆「確証性」=1 : 演繹推論できる。

◆「関係性」=0 : 無関係な情報。 「確証性」=0

◆「関係性」=1 : 答えを含んでいる。 「確証性」=1
(「観測数」≧1)

◆「関係性」＜1 : 「確証性」＜1
(観測数がいくら多くても)

演繹推論は、帰納推論に特別な例と考えることもできます。
「確証性」が0のときは、全く推測できていない状態です。
「確証性」が1のときは、確実な正解が分かっており、演繹できている状態です。
「関係性」が0の情報は、全く無関係な情報です。
「関係性」が1の情報は、答えを含んでいる情報で、それがあれば演繹できます。
「関係性」が1の情報なら、「観測数」が1以上あれば、「確証性」は1になります。
「関係性」が1未満の情報なら、「観測数」がいくら増えても、「確証性」は1未満です。

サンプル品質による確証性の制限

◆「関係性」=c:
(観測数がいくら多くても)

「確証性」≦c

質問: ある女性の体重は？

「関係性」=50%

Answer: 同じ年齢の 男女 1,000,000人の体重の統計的分布

サンプルの50%は男性なのでノイズ

“Certainty”≦50%

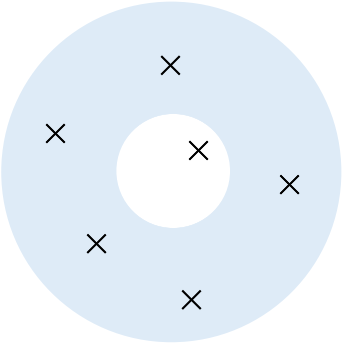
品質が悪ければ、いくら標本が多くても、確証性は制限される

また、「関係性」が一定値なら、「観測数」がいくら増えても、「確証性」は一定値以下になります。
例として、ある女性の体重を聞く質問を考えます。
回答として、同じ年齢の男女100万人の体重の統計的分布を答えました。
しかし、サンプルの50%は男性なのでノイズです。
性別の一致度を関係性とすれば、確証性も50%以下となります。
品質が悪ければ、いくら標本が多くても、確証性は制限されます。

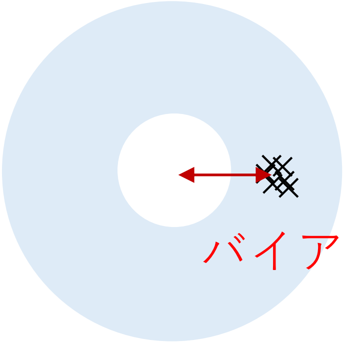
精度と正確度

｜推測値 – 実測値｜= 0：　「確証性」=1

精度(分散)が悪い



正確度(バイアス)が悪い



バイアス：予測しなかった
（中央を狙った）

｜推測値 – 実測値｜の期待値 = 最小値：　“最適な推論”

確証性が大きい状態とは何なのかを考えます。
推測と実測の差が0ならば、不確実性はなく、確証性は1といえます。
推論結果の良し悪しは、精度(分散)と正確度(バイアス)に分けて考えることができます。
例えば、射った矢が、的の一点に集まっていたとしても、中央から偏っていては良くありません。
偏っていたとしても、射手は中央を狙っているつもりなので、どちらに偏るか把握していません。
結果を見れば偏りを知ることができますが、結果を見る前に偏りを知ることはできません。
しかし、知能は過程なので、結果の良し悪しは関係ありません。
予測と実測の差の期待値が最小だと推測できていれば、それが最適な推論なのです。

偏差

(実測値 – 推測値)^2 の期待値

二乗平均平方根偏差

(標準偏差)

$$RMSD = \sqrt{\frac{1}{n}\sum(x_i - \bar{x})^2}$$

質問者が選択

平均絶対偏差

$$MAD = \frac{1}{n}\sum|x_i - \bar{x}|$$

｜実測値 – 推測値｜の期待値

偏りを知るのが困難な一方で、分散の方は統計で計算できます。
統計的なバラツキは、標準偏差(二乗平均平方根偏差)を見るのが一般的です。
絶対平均偏差という値もあります。
どちらを使うべきかは、質問の出題者が指示するべきことです。
例えば、推測と実測値の差の二乗に比例して損害が発生するなら、標準偏差で評価するべきです。
推測値の絶対平均偏差は、推測値と実測値の距離の期待値に相当します。
単純により小さい誤差で推測するのが目的なら、絶対平均偏差で評価するのが妥当です。

情報

利用可能な情報：　$A \subset B$
推論の品質：　$A \leqq B$

どんな情報でもあった方が良い

一部の情報を見て見ぬふりするAIは最適ではない

利用可能な情報：　$A = B$
推論の品質：　$A = B$

利用可能な情報が同じなら、
推論も同じ

重複した情報で推論が変わるAIは最適ではない

帰納推論の良し悪しは、情報量の観点からも考えることができます。
どんな情報であっても、無いより有った方が、同等またはより良い推論ができるはずです。
利用可能な情報量が増加するほど、推論の良さも単調増加します。
もし、AIが、ある情報をあえて知らないフリをすることがあれば、最適ではありません。
逆に、利用可能な情報量が増えなければ、より良い解に変化することはできません。
同じ情報を2回伝えられた場合は、情報量は増えないので、最適解は変わらないはずです。
一部のデータが複製されて重複している場合に、そのデータが強調されてしまうAIは、最適ではありません。

人間のバイアス

	A		B
利用可能な情報	株価の 予測式		株価の 予測式 それはバックテスト にフィットさせた式 ネガティブな情報
儲かる確率	99.99%		50%
精度	A	>	B
利用可能な情報	A	<	B
推論の品質	A	<	B

AIに限った話ではなく、人間の知能にもさまざまなバイアスがあります。
例えば、ポジティブな情報を重視して、ネガティブな情報は無視することがあります。
ネガティブな情報であっても、無視せず、考慮した方が、良い推論になります。
例えば、高確率でお金が儲かると推論しているとき、それが否定される情報は無視したくなります。
その否定的な情報の追加によって、儲かる確率の推測値が減少しますが、それでもより良い推論に変化したはずです。
確率の大小は、分散しか表しておらず、バイアスが反映されていません、
統計的に儲かる確率の方が高いとしても、都合が良いように前提条件を設定しているだけかもしれません。

報酬の最大化

強化学習の目標:　報酬の推測値の最大化

推測式　　　　　　　　　　報酬の推測値

$$A: ax^5 + bx^4 + cx^3 + dx^2 + ex + f = 99$$

$$B: ax^2 + bx + c = 80$$

$$C: ax + b = 50$$

$\xrightarrow[\text{式を選ぶ}]{\text{AIが}}$ 式 A

式Aは最大値だけど、偶然かもしれない?

強化学習における「報酬」を最大化するという考え方にも、大きな落とし穴があります。
未来に受けとる報酬は未知なので、最大化するのは、あくまで報酬の推測値です。
統計の前提条件など、推測の仕方をAI自身が決められるとします。
複数の方法を試して、単純に報酬が最大化する推論方法を採用するとします。
しかし、相関性があるように見えて、実際には偶然だということが良くあります。
さまざまな方法を試せば、都合のよい偶然が起こった方法を選ぶことができます。
推論の仕方までAIに判断させようとする場合は、このようなバイアスを回避する仕組みが必要です。

ハイパーパラメータ

ハイパーパラメータ: ニューラルネットワークの層数など

	ハイパーパラメータの決め方	推論の品質
A	勘で決める	悪い
B	試験の平均点が最大になる 固定値にする	平均的には良い
C	質問毎に最適値へ変動可能にする (ハイパーパラメータ消滅)	質問毎に 最良になりうる

強化学習には、割引率というハイパーパラメータがあります。
ハイパーパラメータは、勘ではなく、試験結果が良くなるように決めるべきです。
ただし、平均点が最高になるように決めたとしても、ハイパーパラメータが消えるわけではありません。
平均的に良いだけではなく、質問毎に、最も良いハイパーパラメータであるべきです。
ニューラルネットワークにおける層数もハイパーパラメータです。
平均的に成績が良い層の数に固定されているのが一般的です。
固定せずに、質問毎に最適な層数に変動可能な方が、成績アップの余地があります。
質問毎に最適化されるようになって初めて、ハイパーパラメータがなくなったといえます。

帰納の公理

帰納の公理:

関数fの元xが近いほど、写像yも近い

$y = f(x)$　どんな関数なのかは不明

次数	関数	写像	頻度
0	0次関数(定数)	|f(1)-f(0)| ＝ |f(2)-f(0)|	＝100%
1	1次関数(線形)	|f(1)-f(0)| ≦ |f(2)-f(0)|	＝100%
≧2	2次以上の関数	|f(1)-f(0)| ≦ |f(2)-f(0)| |f(1)-f(0)| ≧ |f(2)-f(0)|	＝50% ＝50%
?	不明な関数	|f(1)-f(0)| ≦ |f(2)-f(0)| |f(1)-f(0)| ≧ |f(2)-f(0)|	≧50% ≦50%

最適な帰納推論の方法を決めるために、できるだけ単純な事例を考えます。
「関数fの元xが近いほど、写像yも近い」を帰納の公理とします。
y=f(x)で表しますが、どのような関数なのかは不明です。
0次関数の場合、yはxによらず一定です。
1次関数の直線なら、xが近いほど、yも近くなります。
2次以上の関数では、直線より上に逸れるか下に逸れるか分かりません。
どんな関数か分からなくても、xが近いほど、「yが遠い」よりも「yも近い」といえます。
1次関数の場合はこの関係が成り立ち、その他の場合は分からないので同等だからです。

元が複数の関数

$$y = f(x, z)$$

条件	写像	頻度
Δx1＜Δx2 Δz1＝Δz2	|f(1, c)-f(0, 0)| ≦ |f(2, c)-f(0, 0)|	≧50%
	|f(1, c)-f(0, 0)| ≧ |f(2, c)-f(0, 0)|	≦50%
Δx1＜Δx2 Δz1＜Δz2	|f(1, 1)-f(0, 0)| ≦ |f(2, 2)-f(0, 0)|	≧50%
	|f(1, 1)-f(0, 0)| ≧ |f(2, 2)-f(0, 0)|	≦50%
Δx1＜Δx2 Δz1＞Δz2	|f(1, 2)-f(0, 0)| ≦ |f(2, 1)-f(0, 0)|	?
	||f(1, 2)-f(0, 0)| ≧ |f(2, 1)-f(0, 0)|	?

y=f(x, z)の場合も考えてみます。
zが同じなら、xが近いほどyも近いと推測できます。
zが近く、xも近いなら、yも近いと推測できます。
zが遠く、xが近い場合は、どちらともいえません。
xとzの単位が異なれば比較できません。
例え物理的な単位が同じでも、表しているものが異なるなら、比較するのは乱暴です。

部分的にランダムな関数

$$y = f(x, z, \ldots, \varepsilon)$$

ε : 隠れた変数 …ノイズ
(x, z, …以外でyに影響する全ての変数)

$$\lim_{n\to\infty}\left(\frac{1}{n}\sum \varepsilon_i\right) = 0$$

データ数が増えるほど、ノイズが減り、確証性が増加

部分的にランダムに決まる関数の場合も考えます。
関数の元に、εを付け加えました。
一見ランダムに見えても、隠れた変数εによって必然的に決まったと解釈します。
εは、xやz以外で、yに影響を与える全ての変数を表します。
εはノイズと見做したものであり、データ数が増えるほど、平均値は0に近づきます。
そのため、データ数が増えるほど、ノイズが減り、確証性が増加します。

隠れた変数の推測

$y = f(x)$ （隠れた変数 ε は存在しない）

質問: $f(0) = ?$

	A	B
観測	$f(2) = 2$	$f(2) = 2$
	$f(1) = 1$	$f(1) = 1$
		$f(1) = 1$
		$f(1) = 1$

隠れた変数がないので、
何度観測しても同じ値

確証性: A ＝ B

帰納推論の本質: 隠れた変数の統計的な推測

逆に、隠れた変数 ε が存在しない場合も考えてみます。
y=f(x)という関数の、f(0)の値を推測したいとします。
f(2)=2より、f(1)=1の方が、f(0)に近いと推測します。
隠れた変数がないので、f(1)は、何度観測しても1です。
f(1)=1を何度観測しても推論の精度は上がりません。
隠れた変数が存在しても、その値が同じだと分かっている場合も同様です。
帰納推論の本質は、隠れた変数の統計的な推測です。

ノイズ耐性

帰納の公理: $y = f(x)$

関数fの元xが近いほど、写像yも近い

xの差が僅かなら、yの差も僅かなはず

AIが満たすべき特性:

何らかの値が僅かに変化しただけなら、出力も僅かしか変化してはいけない。

・数値が一致するかどうかで条件分岐してはいけない。

・最大値や最小値を持つサンプルを選出してはいけない。

僅かな差で大きく変化するということは、
ノイズの影響を強く受けるということです。

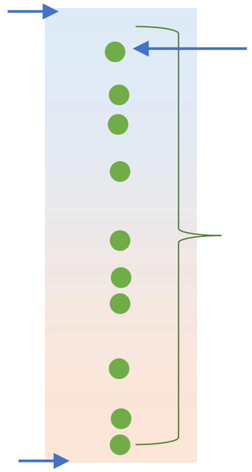
帰納の公理から、AIが満たすべき性質が導かれます。
xの差が僅かなら、yの差は僅かなはずです。
何らかの値が僅かに変化しただけなら、出力も僅かしか変化してはいけません。
例えば、数値が一致するかどうかで条件分岐してはいけません。
また、最大値や最小値を持つサンプルを選出してはいけません。
僅かな差なら、ノイズで大小関係が逆転してしまいます。
僅かな差で大きく変化するということは、ノイズの影響を強く受けるということです。

計算量

巡回セールスマン問題

ランキング

最短経路
(最良回答)



最良のもの:　上位10% (1/n)

n=10 ランダム抽出

品質が上位 1/n の回答に必要な計算量： O(n)

最長経路
(最悪回答)

もし、AIのアルゴリズムに組み合わせ爆発が起こったなら、
それは最適ではないと疑うべき。

ここまで、計算時間を無視してきましたが、大した問題ではないことを示します。
例えば、巡回セールスマン問題では、全ての経路の組み合わせの数が、非常にたくさんあります。
そこで、ランダムに10個の経路を調べて、最も良いものを選ぶとします。
全ての経路を良い順に並べると、10個の内で一番良いものは、およそ上位10%です。
品質が上位1/nの回答をするのに必要な計算量のオーダーは、nです。
制限時間があれば、最適解ではなくても、現状できる最も良い回答をするしかありません。
この方法は、どんな問題にでも、汎用的に使えます。
もし、AIのアルゴリズムに組み合わせ爆発が起こったなら、それは最適ではないと疑うべきです。

複数の問題

◆問題 A 「…」　　人間が設定

◆問題 B 「…」　　問題Aを解くために、AI自身が設定

◆問題 C 「すべての問題の優先順を決める」

- 問題 A
- 問題 B
- 問題 C（自身）

問題が複数ある場合を考えます。
計算能力が有限なら、どの問題から解くか、優先度を決める必要があります。
人間が設定する問題なら、人間が優先度を指定します。
また、AI自身が、既存の問題を解くために、新たな問題を設定する場合があります。
AIが問題を設定するなら、AIが優先度を決めなければなりません。
「すべての問題の優先順を決める」という問題が発生します。
「優先度を決めるという問題」自身も含めて、優先度を決めます。

フレーム問題

新しい質問の設定

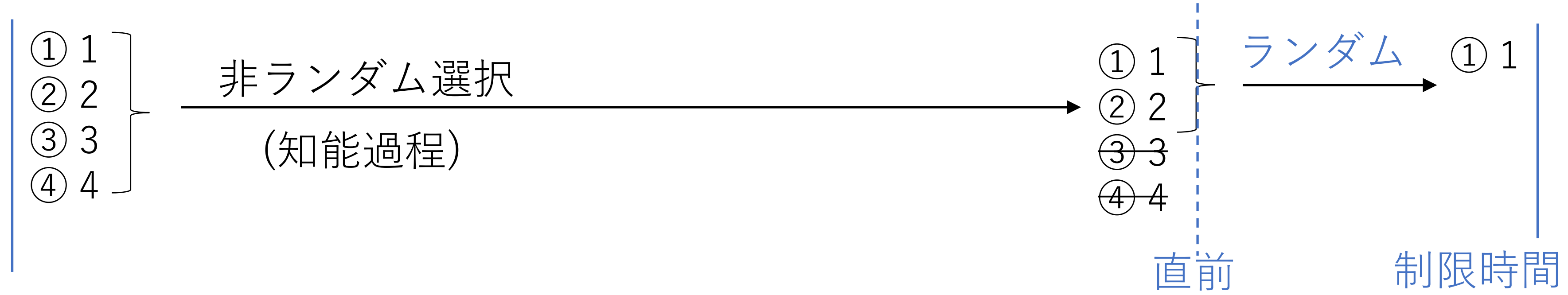
- 回答選択肢
- 利用可能な情報
- 評価関数

無限に設定することが許される

無限個の中からn個を無作為抽出した中から最も良いものを選べば、品質上位1/nの回答が得られる

n個から、最も良いものを選ぶなら、n-1回の比較が必要

n=1 :単純に一つをランダム抽出。比較が不要



AI自らが新しい質問を設定する場合、「回答選択肢」などの設定が必要です。
「回答選択肢」などを無制限にすることも許されます。
考慮するべきことが無限にあると、フレーム問題に陥ります。
ですが、無限個の中からn個を無作為抽出した中から最も良いものを選べば、品質上位1/nの回答が得られます。
n個から、最も良いものを選ぶなら、n-1回の比較が必要です。
時間がなければ、n=1として、ただ無作為抽出すれば、選択肢の評価をする必要がありません。
例えば、4択問題で、2択まで絞り込めていた場合、制限時間の直前に2択から無作為に選べばよいのです。
つまりは、知能の過程で処理しきれない分だけ無作為過程で処理すれば、フレーム問題に対処できます。

情報の劣化

・知能によって絞り込んだ選択肢
（できるだけ保持）

・ランダムに選択した選択肢
（できるだけ避ける）

区別しなければならない

・推測値(帰納)

・実測値

区別しなければならない

AIは、やむおえず部分的にランダム選択をしなければいけない場合があります。
ランダム選択は知能ではないため、できる限りランダム選択していない状態を保持するべきです。
AI内部では、ランダム選択した後でも、ランダム選択する前の状態を覚えておくべきです。
ある値が、推測値なのか、ランダムに選んだだけの値なのか、区別しないといけません。
また、どんなに優れた推測値であっても、演繹ではないのなら、実測値と区別しなければなりません。
全ての情報は、削除や改竄だけでなく、不確かな情報の追加による劣化も、防がなければいけません。
さもなくば、AI自身が、嘘を付いているかどうかも判断できなくなります。

生成AIのハルシネーション

	質問	回答	評価
訓練時	A社の株価は?	123,456 です	良い
	B社の株価は?	分かりません	悪い
実行時	B社の株価は?	123,456 です	

AIはトレーニングされた通りに答えているだけです

生成AIは、ハルシネーションという現象に悩まされています。
その原因は、AI自身が真偽の区別を誤ることだけではありません。
生成AIは、ただトレーニングされた通りに答えているだけで、何も間違えていません。
トレーニング内容や、評価基準が悪いだけです。
「分かりません」という回答は、低評価にするか、除外してトレーニングしているはずです。
人間社会では、当てずっぽうでも良いので、少しでも正解する可能性がある回答することが推奨されています。
それに、質問者は知らないから聞いているので、その場で嘘を見抜かれて低評価されることはありません。
もしAIに「分かりません」と言われたなら、「よく素直に告白したね」と褒めてあげてください。

定義のまとめ

「知能」の定義:
評価基準に合った出力する能力

「究極の一般知能」の定義:
どんな質問でも、必ず最適解が得られる能力
(いくら時間が掛かっても)

究極のAGIへの
スタートライン

最適な帰納推論の方法を定義

・個々の事例からパズルのピースを収集
・パズルを組んで一般化

究極のAGI完成
のゴール

全ての質問の最適解を定義

ここまでのまとめをします。
「知能」を評価基準に合わせて出力する能力と定義しました。
いくら時間が掛かっても良いので、どんな問題でも最適解を出力する能力を「究極の一般知能」を定義しました。
ここが、究極のAGIへのスタートラインです。
全ての質問の最適解を定義できればゴールです。
そのためには、最適な帰納推論の方法を定義する必要があると分かりました。
まずは個々の事例からパズルのピースを収集します。
うまくパズルを組んで一般化できれば、究極のAGIが完成します。

究極のAGIに必要なことのまとめ

1. 最適化されない値(ハイパーパラメータ)があってはいけない
2. 推測値は、実測値と区別しなければいけない
3. どんな情報でも、見て見ぬふりしてはしけない
4. 重複した情報で、出力が変化してはいけない
5. 情報の僅かな差で、出力が大きく変化してはいけない
6. ランダムな選択は、最小限にしなければいけない
7. 精度(バリアンス)だけでなく、正確度(バイアス)を考慮しなければいけない
8. 組み合わせ爆発してはいけない

究極のAGIの要求特性をまとめました

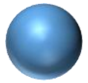
1. 最適化されない値(ハイパーパラメータ)があってはいけない
2. 推測値は、実測値と区別しなければいけない
3. どんな情報でも、見て見ぬふりしてはしけない
4. 重複した情報で、出力が変化してはいけない
5. 情報の僅かな差で、出力が大きく変化してはいけない
6. ランダムな選択は、最小限にしなければいけない
7. 精度(バリアンス)だけでなく、正確度(バイアス)を考慮しなければいけない
8. 組み合わせ爆発してはいけない

AGIの比較

	商業AGI	究極のAGI
目的	労働の置換	知能の解明
一般性	人間レベル	最大
過程	不問(ブラックボックス)	公理から証明
成績評価	ベンチマーク試験	不要

商業AGIと究極のAGIを比較します。
商業AGIの目的は労働の置換ですが、究極のAGIの目的は知能の解明です。
目指す一般性についても、人間レベルと最大で異なります。
商業AGIは、ブラックボックスでも良いのでベンチマークで高得点を目指します。
究極のAGIは、過程が証明されていれば、ベンチマークは不要と考えます。

開発方針の比較

	従来のAI	究極のAGI
目指す形状		
改良方向	良い点を 見つけて伸ばす	悪い点を 見つけて治す
重要な能力	アイデアを出すこと	アイデアを否定すること
進捗確認	ベンチマーク	理論的

従来のAIと究極のAGIでは、開発方針が異なります。
従来のAIは、良い点を見つけて伸ばします。
一方、究極のAGIでは、悪い点を見つけて治します。
アイデアを出すことよりも、アイデアの悪い点を見つけて否定する能力が重要になります。
また、ベンチマークテストが有効ではないので、進捗確認が困難です。

参考文献

・自身の知能

（知能の観察記録を読むよりも、知能の実物を観察しましょう。）

参考文献は、自身の知能です。
知能の観察記録を読むよりも、知能の実物を観察しましょう。

あとがき

この研究は、ボランティアで行われています。

ベンチマーク試験が不要なので、スパコンやビッグデータは必要ありません。

誰でも挑戦することができます。

利益を得るための近道ではありませんが、知能を解明するための近道です。

人類最後の発明になる究極のAGIを目指してみませんか？

Contact: ai@ultagi.org　　https://ultagi.org/

この研究は、ボランティアで行われています。
ベンチマーク試験が不要なので、スパコンやビッグデータは必要ありません。
誰でも挑戦することができます。
利益を得るための近道ではありませんが、知能を解明するための近道です。
人類最後の発明になる究極のAGIを目指してみませんか？

次回予告

次回の動画は「統計2.0」の話になる予定です。

名前は仮です。

最適な帰納推論の方法を決めていきます。

そのために統計の概念を大きく拡張します。

今回の動画では、重要なパズルのピースを挙げていきました。

次回の動画からは、パズルのピースが組み合わさっていきます。

次回の動画は「統計2.0」の話になる予定です。
名前は仮です。
最適な帰納推論の方法を決めていきます。
そのために統計の概念を大きく拡張します。
今回の動画では、重要なパズルのピースを挙げていきました。
次回の動画からは、パズルのピースが組み合わさっていきます。

お問い合わせ先

お問い合わせは、

こちらからお願いします。

https://ultagi.org/