第6回 人工知能学会 汎用人工知能研究会 (SIG-AGI)

強化学習と汎用エージェント

相澤　彰子
国立情報学研究所
aizawa@nii.ac.jp

「汎用」 ＝ ユニバーサル（universal）

背後にある確率モデルに関する
事前の知識を必要としない

「汎用」 ＝ ユニバーサル（universal）

- ユニバーサル符号化
  - データ系列に関する事前知識を必要とせず、データ長が長くなるにつれて漸近的に最適の圧縮を行う符号化方式
- コルモゴロフ複雑性
  - ある万能チューリングマシンを使って、与えられた文字列を出力するための最小のプログラムの長さ
- ソロモノフの万能推論
  - コルモゴロフ複雑性に基づく予測の最適化

話題①

Compression / Information distance

AIXI / universal algorithmic intelligence

話題②

情報距離・圧縮距離

汎用的な「距離」の尺度

- 圧縮による距離の計算
  - データ系列の複雑さを、生成プログラムのサイズで測る
  - 理想の圧縮プログラムを想定する

- 相対エントロピーによる距離の計算
  - 相対エントロピーの値を、圧縮を使って近似的に求める。

圧縮距離：歴史（1）

- 情報距離（Information Distance） 1998

  C.H.Bennet, P.Gacs, M.Li, P.M.B Vitani, and W.H. Zurek: “Information Distance,” IEEE trans. on Information Theory, 44(4), 1407-1423.

  - ２つのデータの間の距離を圧縮の度合いによって測る
  - 両者の間で重なりが大きければ、両者をあわせたデータはより効率的に圧縮できる
  - 「理想の圧縮プログラム」を想定した理論的な距離

↓

データの種類によらない
汎用的な距離尺度

コルモゴロフ複雑性

文字列の複雑さはそれを出力する最小のプログラムの長さで表される。

- アルゴリズム情報理論：記号列の複雑さを定義する
- x, y：２進数列$\{0,1\}^*$
- xのコルモゴロフ複雑性（アルゴリズム的エントロピー） $K(x)$
  - 万能チューリングマシンUが与えられたとき、xを出力するために必要な最小のプログラムx^*のサイズ

注意：$K(x)$の真の値は計算不可能

コルモゴロフ複雑性と情報距離

- yに対するxの条件付きコルモゴロフ複雑性$K(x|y)$
  - 補助データyを用いて復元することを想定したときにxを復元するために必要な最小のプログラムのサイズ
  - $K(x) - K(x|y)$はy中に含まれるxの情報量とみることができる。
- 情報距離$E(x,y) = \max\{K(y|x), K(x|y)\}$
  - 以下の距離空間（metric space）の条件を満足する
    - $D(x,y) = 0 \ iff \ x = y$ (the identity axiom)
    - $D(x,y) + D(y,z) \geq D(x,z)$ (the triangle inequality)
    - $D(x,y) = D(y,x)$ (the symmetry axiom)

圧縮距離：歴史（2）

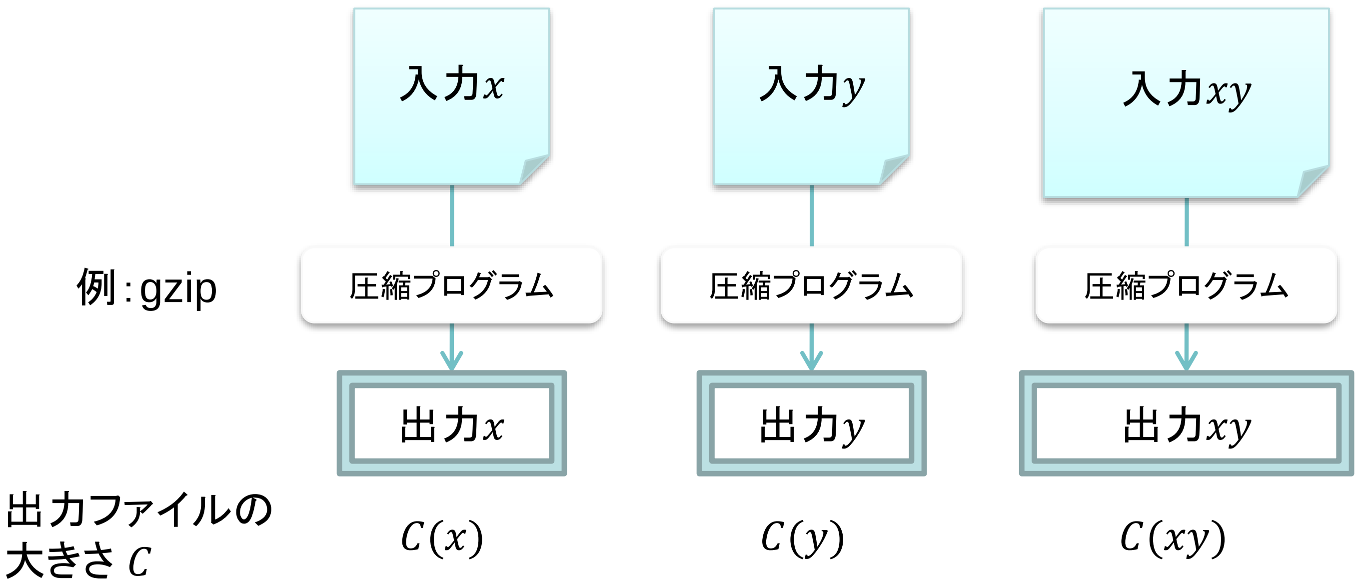
- 正規化圧縮距離（Normalized Compression Distance) 2004

  M. Li, X. Chen, X. Li, B. Ma, and P. Vitani: “The Similarity Metric,” IEEE Trans. on Information Theory, 50(12), 3250-3264

  – 「理想の圧縮プログラム」ではなく、gzipなどの普通の圧縮プログラムを用いて近似的に距離を求める

正規化圧縮距離

- 基本的な考え方
  - 情報の似ている度合いを圧縮プログラムで測る



正規化
圧縮距離

$$NCD(x,y) = \frac{C(xy) - \min\{C(x), C(y)\}}{\max\{C(x), C(y)\}}$$

圧縮距離：歴史（3）

- 圧縮によるクラスタリング（2005）

  Rudi Cilibrasi and Paul M. B. Vitanyi : “Clustering by Compression,” IEEE Trans on Information Theory, 51(4), 1523-1545

  – NCDによる距離を実際のクラスタリング問題に適用してみせた
    - 遺伝子の間の距離に基づく系統学的分析
    - 言語間の距離（52言語）
    - 文学作品の著者推定（ロシア文学）
    - 音楽（ロック、ジャズ、など）
    - OCR文字認識
    - 天文観測（X線）データ

正規化圧縮距離(NCD)の歴史

理想の圧縮プログラム

C.H.Bennet, P.Gacs, M.Li, P.M.B Vitani, and W.H. Zurek: “Information Distance,” IEEE trans. on Information Theory, 44(4), 1407-1423 (1998)

◆情報距離
◆理想の圧縮プログラム

M. Li, X. Chen, X. Li, B. Ma, and P. Vitani: “The Similarity Metric,” IEEE Trans. on Information Theory, 50(12), 3250-3264 (2004)

◆正規化圧縮距離
◆既存の圧縮プログラム

Rudi Cilibrasi and Paul M. B. Vitanyi : “Clustering by Compression,” IEEE Trans on Information Theory, 51(4), 1523-1545 (2005)

◆圧縮によるクラスタリング
◆多様なアプリケーション

身近な圧縮プログラム

[M.Cebrian et.al. 2005]

◆Calgary Corpus（種類が異なる文書群）
◆bgip2, gzip, PPMZ / 正規化方法
◆ファイルの大きさの影響、ウィンドウサイズの影響

[D.Sculley et.al. 2006]

◆Unix User Data
◆LZ77, LZW, PPM
◆PPMが最良

著者同定問題における文字列カーネル、SVMなどとの比較

[D.Pavelec et.al 2009]

◆著者同定問題
◆PPM, サポートベクタマシン
◆互角の性能

1998　2004　2005　2006　2009

汎用的な「距離」の尺度

- 圧縮による距離の計算
  - データ系列の複雑さを、生成プログラムのサイズで測る
  - 理想の圧縮プログラムを想定する

- 相対エントロピーによる距離の計算
  - 相対エントロピーの値を、圧縮を使って近似的に求める。

LZ符号化（辞書式圧縮方式）

- ユニバーサル符号化方式の1つ（情報源の確率分布が未知）
- LZ78(compress,Gif)、LZ77(gzip)
- ポイント：辞書を使って、一続きの文字列を符号に対応させる。

今日はいろいろなことがあっていろいろ大変だった。

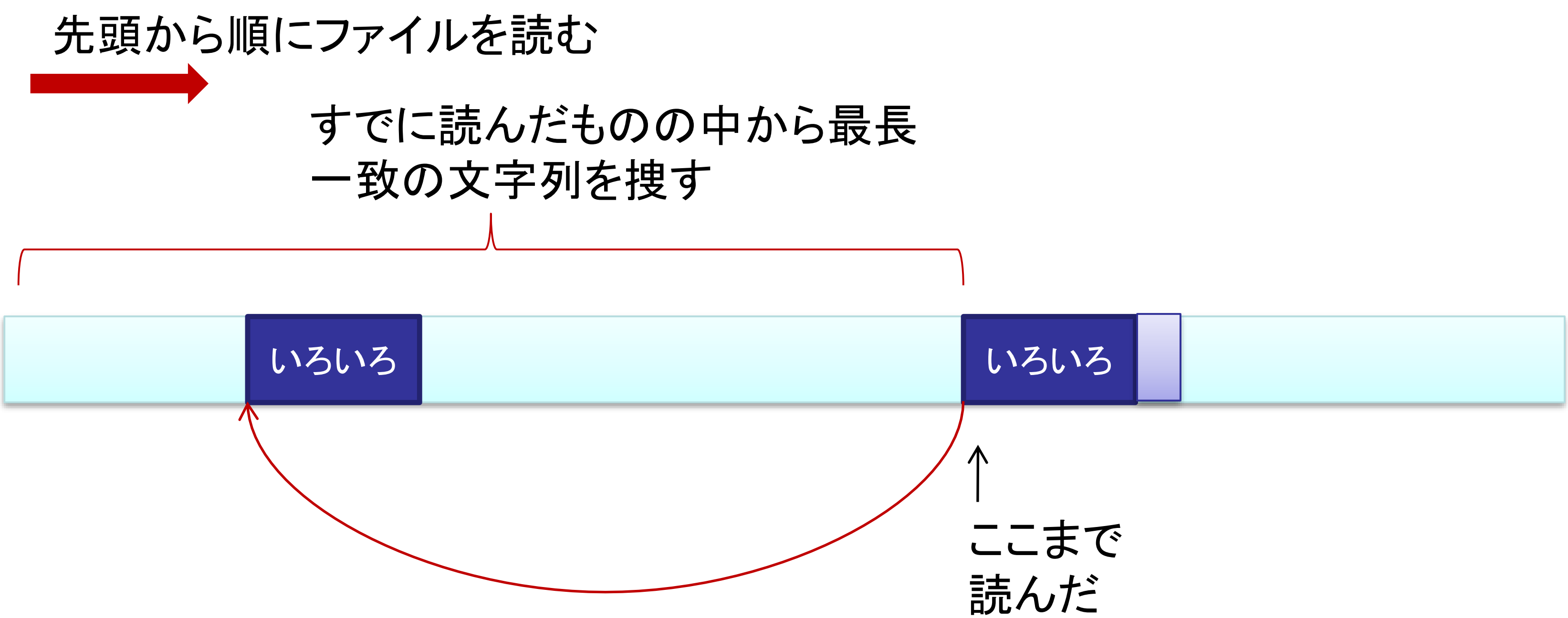
辞書

いろいろ ⟶ 0100

LZ符号化と情報距離

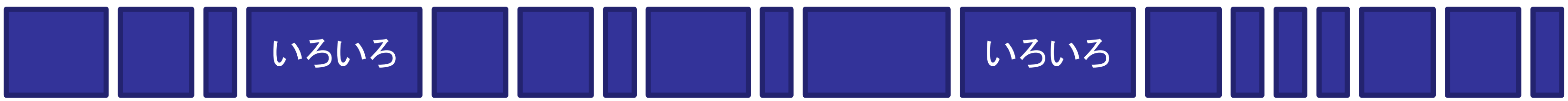
一定幅のウィンドウをスライドさせ、その中で最長一致文字列を捜すのがLZ77

トライなど検索に適した構造で辞書を逐次的に(別に)構築するのがLZ78



LZ符号化

- 十分に長い系列に対して、平均符号長がエントロピーに収束することが証明されている。（簡単なクラス＝k次マルコフ情報源について）



経験的なエントロピーの計算方法を与える

- 系列を空でない相異なる部分列に分解する。
- 与えられた系列xのLZ複雑量とは、そのような部分列の最大個数

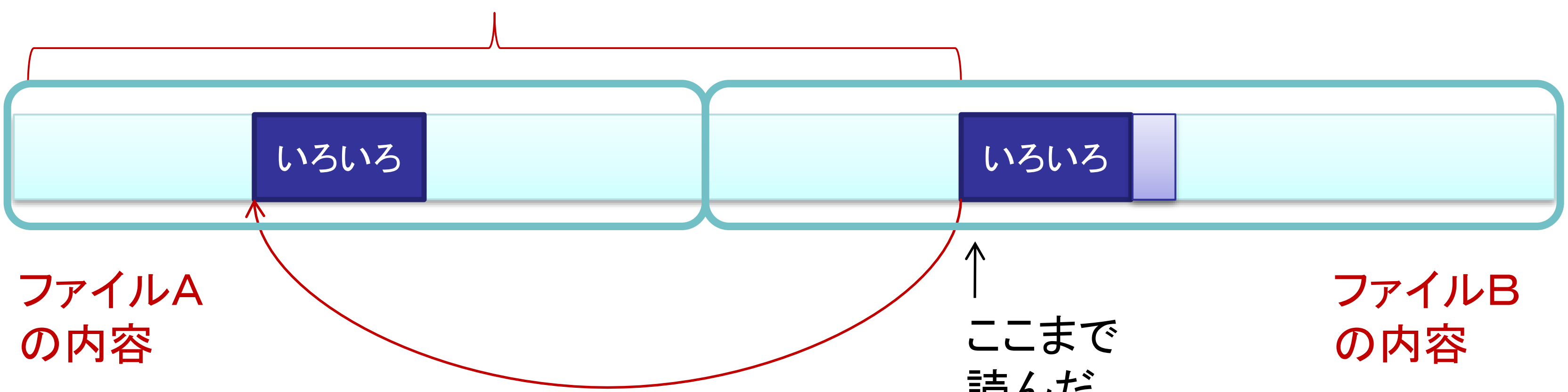
LZ符号化と情報距離

ファイルAとファイルBの近さを求めたい

圧縮距離の計算では、ファイルBを圧縮するときに、ファイルAから作った辞書もあわせて参照している。

先頭から順にファイルを読む

すでに読んだものの中から最長一致の文字列を捜す



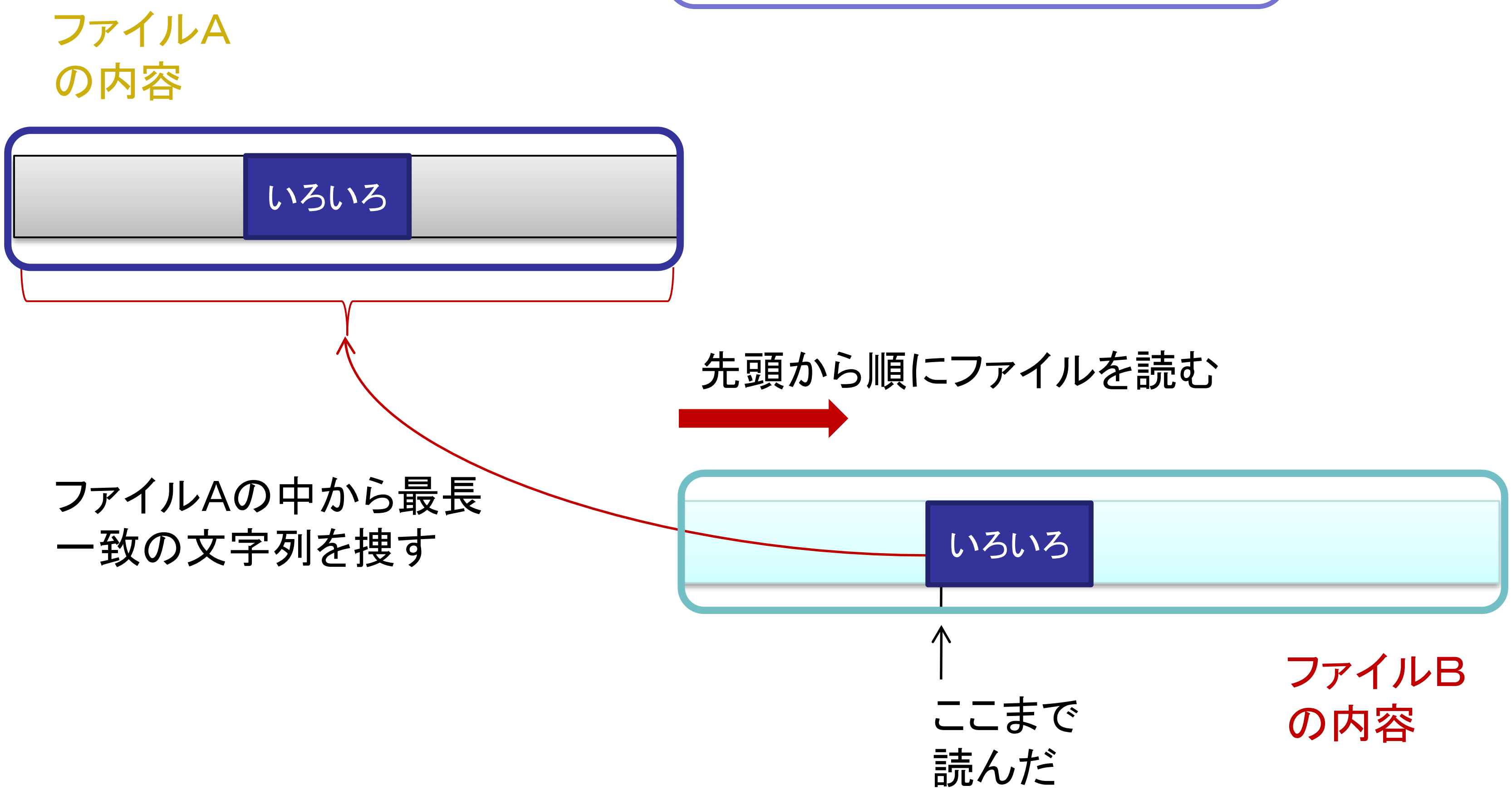
ファイルA
の内容

ここまで
読んだ

ファイルB
の内容

Ziv-Merhav Cross-parsing (1993)

ファイルBを圧縮するとき、辞書は更新しない。



Ziv-Merhav Cross-parsing (1993)

- ２つの文字列x, zが与えられたとき、両者の間のクロスエントロピー（KL情報量、ダイバージェンスとも）の経験値を求める
- まず、zに増分分解法（incremental parsing、LZ78の方法）を適用する。このときの部分列の個数を$c(z)$とする。
- 次に、xから得られた辞書を使ってzを分解する。このときの部分列の個数を$c(z|x)$とする。
- 系列長が十分に長いとき、以下がクロスエントロピーの値に近づく。

何個に分解できるかという
個数がわかればよい

$$\Delta(z \| x) = \frac{1}{n}\left[c(z \mid x)\log_2 n - c(z)\log_2 c(z)\right]$$

経験的なクロスエントロピーの計算方法を与える

Ziv-Merhav法の歴史

ZM分解
→相対エントロピー

情報距離としての適用

ユニバーサル分類器

J.Ziv and N.Merhav: "A Measure of Relative Entropy between Individual Sequences with Application to Universal Classification," IEEE trans. on Information Theory, 39(4), 1270-1279 (1993)

◆Ziv-Merhav Crossparsing
◆相対エントロピーに近づく

[D.Benedetto et.al. 2002]

◆LZ増分分解法を使って相対エントロピーを計算

[D.Coutinho et.al 2005]

◆イタリア語文献の著者同定問題
◆LZ増分分解法とZM法を比較、後者が優位

[S.Helmer et.al. 2007]

◆XML文書のクラスタリング問題
◆NCDとZM法、従来のXML文書類似度尺度を比較
◆ZM法と従来尺度の相補的適用

[A.Martins et.al 2007]

◆ポルトガル語文献の著者同定問題
◆文字列カーネルとZM法を比較
◆互角の性能

Jacob Ziv: Classification With Finite Memory Revisited. IEEE Transactions on Information Theory 53(12): 4413-4421 (2007)

◆有限長のファイルに対するユニバーサルな分類器

Jacob Ziv: On Finite Memory Universal Data Compression and Classification of Individual Sequences. IEEE Transactions on Information Theory 54(4): 1626-1636 (2008)

1993 2002 2005 2007 2008

「汎用」 ＝ ユニバーサル（universal）

話題①

- ユニバーサル符号化
  - データ系列に関する事前知識を必要とせず、データ長が長くなるにつれて漸近的に最適の圧縮を行う符号化方式
- コルモゴロフ複雑性
  - ある万能チューリングマシンを使って、与えられた文字列を出力するための最小のプログラムの長さ
- ソロモノフの万能推論
  - コルモゴロフ複雑性に基づく予測の最適化

Compression / Information distance

AIXI / universal algorithmic intelligence

話題②

AIXI： UNIVERSAL ALGORITHMIC INTELLIGENCE

References

[1] Marcus Hutter: Universal Algorithmic Intelligence: A Mathematical Top → Down Approach, *Artificial General Intelligence*, Goertzel, Ben and Pennachin, Cassio (eds.), Springer Berlin Heidelberg, pp.227-290 (2007).

[2]小林亮太, 相澤彰子: 汎用エージェントの理論的枠組み ー Marcus Hutter が提唱するAIXI の紹介ー, *人工知能*, Vol. 29, No. 3, pp. 234-238 (2014).

- ユニバーサルなエージェント
  - いかなる環境のもとでも強化学習により最適な戦略をとることができるエージェント
  - 「汎用的な知能」

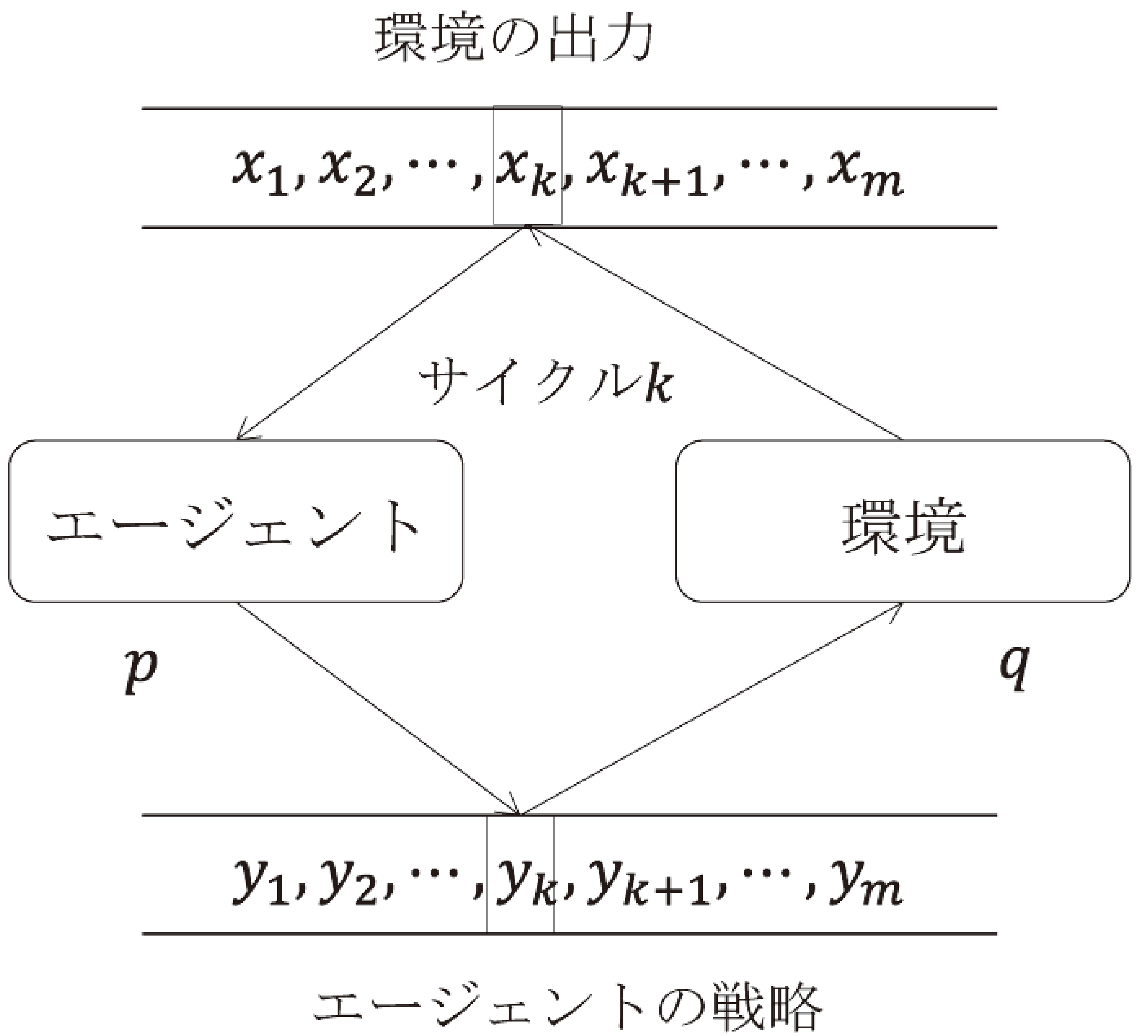
既知の環境の下でのエージェント

- エージェントと強化学習

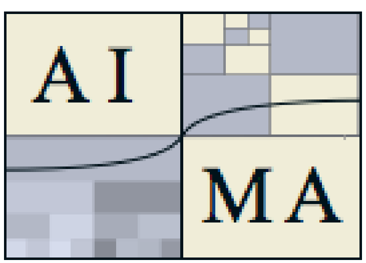
エージェントは環境とやりとりをするシステム。

各サイクルkで、戦略pにしたがってy_kを出力し、環境からx_kを受け取る。

エージェントと環境のやりとり



Artificial Intelligence: A Modern Approach



http://aima.cs.berkeley.edu/

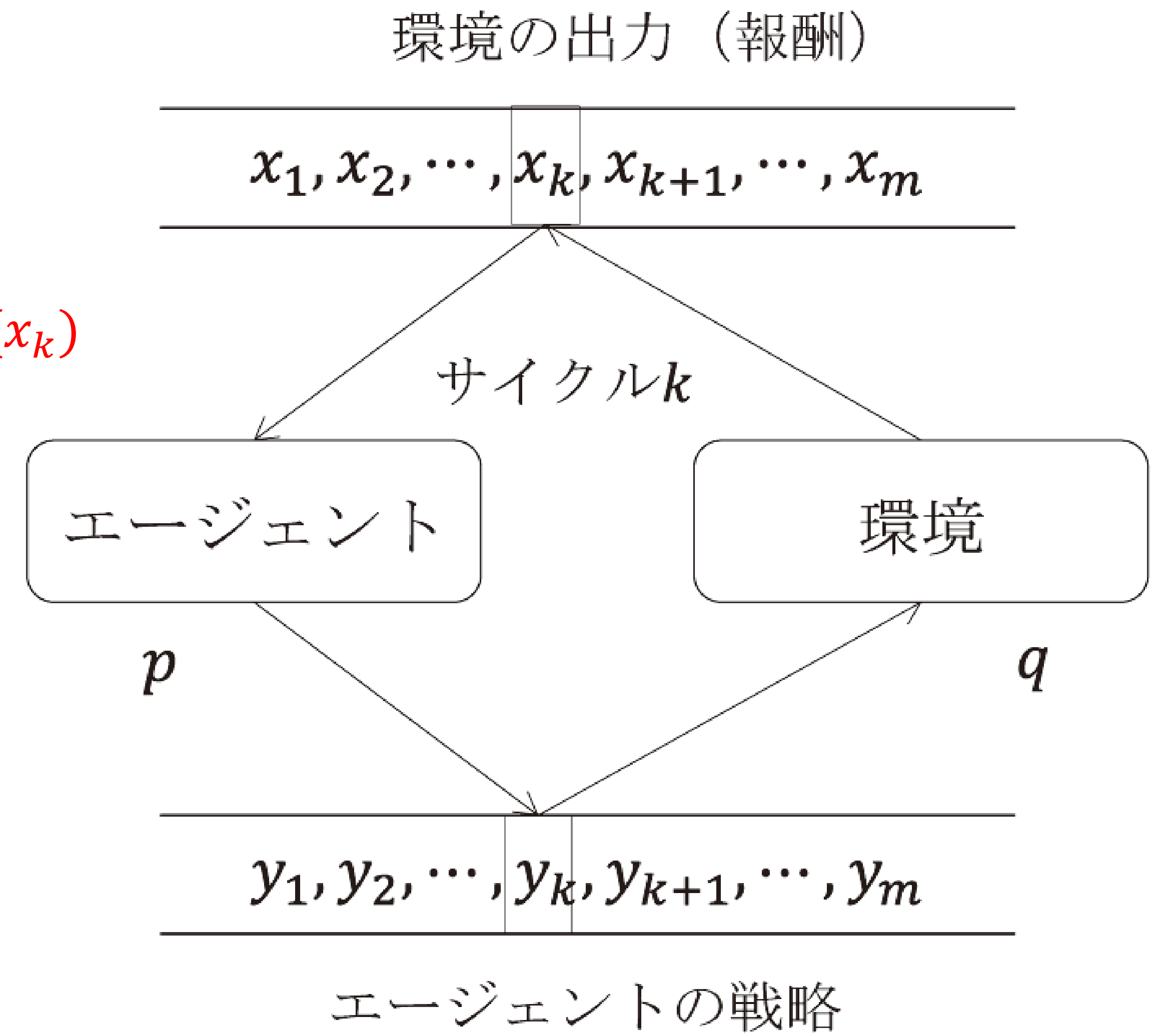
エージェントのふるまい

- これまでの説明では、「環境」と「エージェント」は対称的にふるまう
- どこで「知能」がはいってくるか？

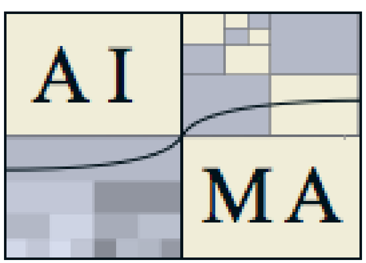
- 「報酬」を考えるところ
- エージェントの目標は「報酬」の最大化。これを実現するのが「強化学習」

エージェントと環境のやりとり

報酬　$r(x_k)$



Artificial Intelligence: A Modern Approach



http://aima.cs.berkeley.edu/

報酬

- エージェントの活動期間をサイクル1からmとする
- その期間での報酬

$$\text{報酬} = r(x_1) + \cdots + r(r_m)$$

- 環境のふるまいを表す確率分布μが既知のとき、この値を最大にする「最適な」エージェントp^*を考える

報酬の最大化(環境が既知)

- 環境のふるまいが**既知の確率分布μ**で与えられる
- ステップkにおいて、ステップ$k,\cdots,m$までの報酬の期待値:

$$\sum_{x_k}\cdots\sum_{x_m}(r(x_1)+\cdots+r(r_m))\times\mu(x_k^m|y_1^{k-1},x_1^{k-1},y_k^m)$$

ステップkにおいてy_1^{k-1},x_1^{k-1}が観察されたときにエージェントの戦略y_k^mに対して環境がx_k^mを出力する確率

この値を最大にするy_kがエージェントのサイクルkにおける最適の戦略

AIμモデル

報酬の最大化（環境が未知）

- 未知の環境について議論する。

- 未知の環境のもとで、過去のデータから将来を予測する

“universal sequence prediction”

コルモゴロフの複雑性、ソロモノフの universal prior

コルモゴロフ複雑性（再掲）

文字列の複雑さはそれを出力する最小のプログラムの長さで表される。

- アルゴリズム情報理論：記号列の複雑さを定義する
- x, y：２進数列$\{0,1\}^*$
- xのコルモゴロフ複雑性（アルゴリズム的エントロピー）　$K(x)$
  - 万能チューリングマシンUが与えられたとき、xを出力するために必要な最小のプログラムpのサイズ

$$K(x) \coloneqq \min_{p}\{l(p): U(p) = x\}$$

ソロモノフの universal prior

プログラムpの長さを、pに関する「信念」に結び付けたもの

プログラムpは語頭符号を使ってバイナリ形式で表現されているものとする

$$\sum_{p} 2^{-l(p)} \leq 1$$

$2^{-l(p)}$をpに関する事前の「信念」として、系列xに対するもっともらしさ$\xi(x)$を定義

$$\xi(x) := \sum_{p:U(p)=x^*} 2^{-l(p)} \geq 2^{-K(x)}$$

x^*はxではじまる文字列を出力するすべてのプログラム

$\xi(x)$を使ったμに関する事後測度の更新

$$\xi(x_k | x_1^{k-1}) := \frac{\xi(x_1^k)}{\xi(x_1^{k-1})}$$

この予測分布は、時間の経過とともに、（かなりはやく）真の分布に近づくことが知られている

未知の環境に対応するエージェント

環境の分布としてソロモノフの universal prior を用いる（AIXI）

AIμモデル
$$\sum_{x_k} \cdots \sum_{x_m} (r(x_1) + \cdots + r(r_m)) \times \mu(x_k^m | y_1^{k-1}, x_1^{k-1}, y_k^m)$$

AIξモデル
$$\sum_{x_k} \cdots \sum_{x_m} (r(x_1) + \cdots + r(r_m)) \times \xi(x_k^m | y_1^{k-1}, x_1^{k-1}, y_k^m)$$

$y_1^{k-1}, x_1^{k-1}, y_k^m$に対して$x_k^m$を返すすべての可能な環境を考えて、以下で計算

$$\xi\left(x_k^m \middle| y_1^{k-1}, x_1^{k-1}, y_k^m\right) = \sum_{q:q\left(x_k^m\right)=y_1^{k-1}, x_1^{k-1}, y_k^m} 2^{-l(q)}$$

汎用エージェント AIXIの戦略

$$y_k^{\xi} = \operatorname*{argmax}_{y_k} \sum_{x_k} \max_{y_{k+1}} \sum_{x_{k+1}} \cdots \max_{y_{m_k}} \sum_{x_{m_k}} \left(r(x_k) + \cdots + r(r_{m_k}) \right) \times \xi(x_k^m | y_1^{k-1}, x_1^{k-1}, y_k^m)$$

主張: AIXI は universally optimal である。

- パラメタフリー
- ξはμに収束する
- AIμは最適
- 他に、よりはやくAIμに収束するモデルが存在しない

以下の有限性は仮定している

- 環境からの入力の取り得る値は有限
- エージェントの出力の取り得る値は有限
- 報酬は非負値で0から上限値の範囲内に収まる
- 先読みの区間は有限

AIXI: 汎用的なエージェントのモデル

- Ray Solomonoff のアルゴリズム情報理論の考え方を踏襲
- ユニバーサルな事前分布を報酬の最適化戦略に取り入れたもの
- マルコフ性が成り立たない非定常的な環境にも適用できる強力な枠組み
- ただし，そもそも有限時間で計算可能でない.

「計算可能な」汎用人工知能モデル

- Hutterの方向感
  - プログラム長、計算時間に制限を加えたモデルは、有限時間で計算可能で、ある定義のもとではもっとも知能が高い
  - 先読み区間 m_k をどうやって決めるか？
  - 離散と連続の対応づけをどうするか？

ご清聴ありがとうございました