A Theory of Universal AI

Literature
Marcus Hutter
Kircherr, Li, and Vitanyi

Presenter
Prashant J. Doshi

CS594: Optimal Decision Making

Roadmap

- Claim
- Background
- Key Concepts
- The AIμ Model in Functional Form
- The AIμ Model in Recursive Form
- The Universal AIξ Model
- Example Constants and Limits
- Application
  - Sequential Decision Theory
- Conclusions

Claim

Development of a universally optimal AI model

Universal $\Rightarrow$ parameterless, unbiased, model-free

Optimal $\Rightarrow$ No other program can learn or solve the task faster

Background

- Decision Theory
  - Solves the problem of rational agent behaviour in uncertain worlds given an *environment*
  - Known prior probability distribution μ over the environment

- Solomonoff's Universal Induction
  - Solves the problem of sequence prediction for an unknown prior distribution
  - Predict the continuation x_n of a given binary sequence $x_1 \ldots x_{n-1}$

$$\mu(x_n|x_{<n}) = \frac{\mu(x_1 \ldots x_n)}{\mu(x_{<n})}$$

Background

- Solomonoff's result:
  - μ expected Euclidean distance between ξ and μ is finite

$$\sum_{k=1}^{\infty}\sum_{x_{1:k}}\mu(x_{1:k})(\xi(x_k|x_{<k})-\mu(x_k|x_{<k}))^2 \stackrel{+}{<} \frac{1}{2}ln2\cdot K(\mu)$$

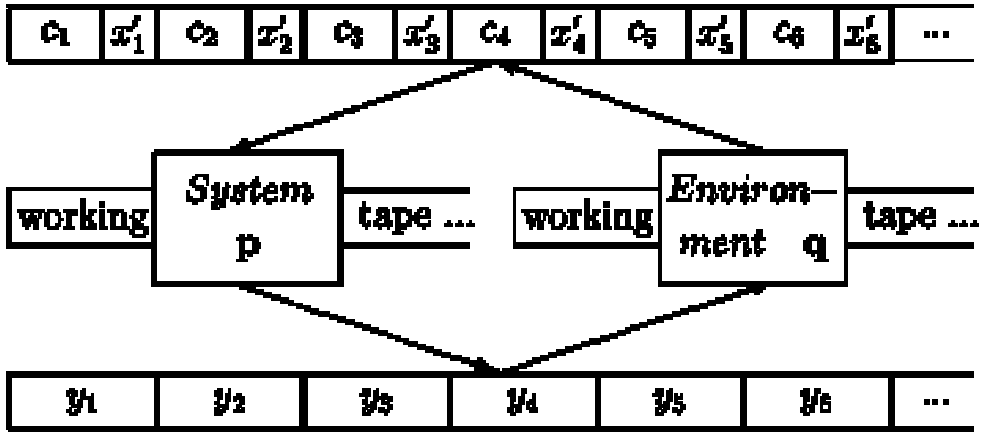
  - Convergence

$$\lim_{n\to\infty}\xi(x_n|x_{<n})\to\mu(x_n|x_{<n})$$

  - ξ is the *Universal Probability Distribution*
  - $\xi(x)\stackrel{\times}{\geq}2^{-K(\rho)}\cdot\rho(x)$
  - $\xi(x)=2^{-K(x)+c}$ $K(x)$ is the Kolmogorov Complexity of x

Key Concepts

The Cybernetic or Agent Model



- $p : X^* \to Y^* \;||\; y_{1:k} = p(x_{<k})$
  - p is partial function / chronological Turing machine
- $q : Y^* \to X^* \;||\; x_{1:k} = q(y_{1:k})$
  - q is partial function / chronological Turing machine

Key Concepts

- History :

$$y_{<k}x_{<k} := y_1x_1y_2x_2y_3x_3 \ldots y_{k-1}x_{k-1}$$

- Probability of input given the history

$$\rho(x_k|y_{<k}x_{<k}y_k) = \frac{\rho(y_{1:k}x_{1:k})}{\rho(y_{<k}x_{<k})}$$

$$\rho(y_{1:k}x_{1:k}) = \rho(x_k|y_{<k}x_{<k}y_k)\rho(x_{k-1}|y_{<k-1}x_{<k-1}y_{k-1}) \ldots$$
$$\ldots \rho(x_2|y_{<2}x_{<2}y_2)\rho(x_1|y_1)$$

The AIμ Model in Functional Form

- **Task**: Derive p^* which maximizes the total credit over a predefined lifetime(T)

  For a known deterministic environment q

  - $V_{km}(p,q) := \sum_{i=k}^{m} r(x_i)$
  - Optimal policy

$$\boxed{p^* := \underset{p}{argmax} V_{kT}(p,q) \Rightarrow V_{kT}(p^*,q) \geq V_{kT}(p,q)\ \forall p}$$

  For a prior distribution over environments $\mu(q)$

  - Let $\dot{Q}_k := \{q : q(\dot{y}_{<k}) = \dot{x}_{<k}\}$ be the set of all environments that produces the history $\dot{y}_{<k}\dot{x}_{<k} := \dot{y}_1\dot{x}_1\dot{y}_2\dot{x}_2 \ldots \dot{y}_{k-1}\dot{x}_{k-1}$
  - $V^{\mu}_{km}(p|\dot{y}_{<k}\dot{x}_{<k}) := \sum_{q\varepsilon\dot{Q}_k} \mu(q) V_{km}(p,q)$

The AIμ Model in Functional Form

- AIμ maximizes expected future reward over the next $h_k = m_k - k + 1$ (horizon) cycles
  - Optimal policy

$$p_k^* := \underset{p:p(\dot{x}_{<k}|\dot{y}_{<k})=\dot{y}_{<k}y_k}{argmax} V_{km}^{\mu}(p|\dot{y}_{<k}\dot{x}_{<k})$$

$$p^*(\dot{x}_{<k}|\dot{y}_{<k}) := p_k^*(\dot{x}_{<k}|p_{k-1}^*(\dot{x}_{<k-1}|\ldots p_1^*)$$

Best output $\dot{y}_k := \underset{y_k}{argmax}\ \underset{p}{max}\ V_{km}^{\mu}(p|\dot{y}_{<k}\dot{x}_{<k})$

- p^* is computable if X,Y and m_k are finite

The AIμ Model in Recursive Form

- **Task**: Derive expected reward sum in cycles k to m using expected reward sum in cycles $k+1$ to m
  - $V^{\mu}_{km}(\dot{y}_{<k}\dot{x}_{<k}y_k) := \sum_{x_k}[r(x_k) + V^{*\mu}_{k+1,m}(\dot{y}_{1:k}\dot{x}_{1:k})]\mu(x_k|\dot{y}_{<k}\dot{x}_{<k}y_k)$
  - Optimal expected reward
    $V^{*\mu}_{km}(\dot{y}_{<k}\dot{x}_{<k}) := \max_{y_k} V^{\mu}_{km}(\dot{y}_{<k}\dot{x}_{<k}y_k)$

$$\boxed{V^{*\mu}_{km}(\dot{y}_{<k}\dot{x}_{<k}) := \max_{y_k} \sum_{x_k}[r(x_k) + V^{*\mu}_{k+1,m}(\dot{y}_{<k}\dot{x}_{<k}x_k y_k)]\mu(x_k|\dot{y}_{<k}\dot{x}_{<k}y_k)}$$

  - $Best$ output
    $\dot{y}_k := \arg\max_{y_k} V^{\mu}_{km}(\dot{y}_{<k}\dot{x}_{<k}y_k)$
  - **Expectimax** sequence

$$\boxed{\dot{y}_k := \arg\max_{y_k} \sum_{x_k} \max_{y_{k+1}} \sum_{x_{k+1}} \ldots \max_{y_m} \sum_{x_m}[r(x_k) + \ldots + r(x_m)] \mu(x_{k:m}|\dot{y}_{<k}\dot{x}_{<k}y_{k:m})}$$

The AIμ Model in Recursive Form

Functional AIμ $\equiv$ Recursive AIμ

$$\mu(y_{1:k}x_{1:k}) \equiv \sum_{q:q(y_{1:k})=x_{1:k}} \mu(q)$$

The Universal AIξ Model

- **Task**: Replace the true but unknown prior probability μ with ξ

  - **In the Functional AIμ model**

$$\dot{y}_k := \underset{y_k}{argmax} \max_{p:p(\dot{x}_{<k}|\dot{y}_{<k})=\dot{y}_{<k}y_k} \sum_{q:q(\dot{y}_{<k})=\dot{x}_{<k}} \mu(q) V_{km}(p,q)$$

$$\Downarrow$$

$$\dot{y}_k := \underset{y_k}{argmax} \max_{p:p(\dot{x}_{<k}|\dot{y}_{<k})=\dot{y}_{<k}y_k} \sum_{q:q(\dot{y}_{<k})=\dot{x}_{<k}} 2^{-K(q)} V_{km}(p,q)$$

The Universal AIξ Model

- **Task**: Replace the true but unknown prior probability μ with ξ
  - **In the Recursive AIξ model**

$$\dot{y}_k := \underset{y_k}{argmax} \sum_{x_k} \underset{y_{k+1}}{max} \sum_{x_{k+1}} \ldots \underset{y_m}{max} \sum_{x_m} [r(x_k) + \ldots + r(x_m)]$$

$$\mu(x_{k:m} | \dot{y}_{<k} \dot{x}_{<k} y_{k:m})$$

$$\Downarrow$$

$$\dot{y}_k := \underset{y_k}{argmax} \sum_{x_k} \underset{y_{k+1}}{max} \sum_{x_{k+1}} \ldots \underset{y_m}{max} \sum_{x_m} [r(x_k) + \ldots + r(x_m)]$$

$$\xi(x_{k:m} | \dot{y}_{<k} \dot{x}_{<k} y_{k:m})$$

The Universal AIξ Model

- **Task**: Show the convergence of AIξ to AIμ
  - Utilize the Solomonoff's result generalized from N = 2 to an arbitrary alphabet
  - y_i are pure spectators

$$\sum_{k=1}^{\infty}\sum_{x_{1:k}}\mu(x_{1:k}|y_{1:k})(\xi(x_k|x_{<k}y_{<k})-\mu(x_k|x_{<k}y_{<k}))^2 \stackrel{+}{<} \frac{1}{2}ln2\cdot K(\mu)$$

$$\lim_{k\to\infty}\xi(x_k|x_{<k}y_{<k}) \to \mu(x_k|x_{<k}y_{<k})$$

- Outputs $\dot{y}_k$ of the AIξ model converge to the outputs $\dot{y}_k$ of the AIμ model atleast for the bounded horizon $h_k = m - k + 1$

Example Constants and Limits

$$\begin{array}{ccccccccc} 1 & \ll & l\langle y_k x_k\rangle & \ll & k & \leq & T & \ll & |Y\times X| \\ 1 & \overset{a}{\ll} & 2^{16} & \overset{b}{\ll} & 2^{24} & \overset{c}{\leq} & 2^{32} & \overset{d}{\ll} & 2^{65536} \end{array}$$

(a) The agents interface is wide

(b) The interface can be sufficiently explored

(c) The death is far away

(d) Most input/output combinations do not occur

These limits are *never used* in proofs but . . .
. . . we are only interested in theorems which do not degenerate under the above limits

Application

- Sequential Decision Theory (MDP)
  - Bellman equation for optimal policy

$$p^*(i) = \underset{a}{argmax} \sum_j M_{ij}^a U^*(j),\ U^*(i) = R(i) + \underset{a}{max} \sum_j M_{ij}^a U^*(j)$$

- Apply the AIμ model

$$S = (Y \times X),\ A = Y,\ a = y_k,\ M_{ij}^a = \mu(x_k | y_{<k} x_{<k} y_k)$$

$$i = y_{<k} x_{<k},\ R(i) = r(x_{k-1}),\ U^*(i) = V^*_{k-1,m}(y_{<k} x_{<k}) = r(x_{k-1}) + V^*_{km}(y_{<k} x_{<k})$$

$$j = y_{1:k} x_{1:k},\ R(j) = r(x_k),\ U^*(j) = V^*_{km}(y_{1:k} x_{1:k}) = r(x_k) + V^*_{k+1,m}(y_{1:k} x_{1:k})$$

Application

- Observations
  - We use the complete history as the environment state
  - The AIμ model does not assume
    - Markovian property
    - stationary environment
    - accessible environment
- Other applications
  - Game Playing
  - Function Minimization
  - Supervised Learning
- **Bold Claim**: AIμ is the most general model

Conclusion

- A parameterless model of AI based on Decision Theory and Algorithmic Probability is presented
- Makes minimal assumptions about the environment
- Is the AIξ model computable?
- Future Work
  - Derive value and reward bounds for AIξ model
  - Apply AIξ model to more problem classes