A Fourier AIXI Approximation (Universal AI)

Swapnil Shah, National Institute of Technology-Surat, India. *Member, GSTF.*
(email: swapnil.nitin.shah@gmail.com)

***Abstract*—AIXI is a pareto optimal theory of artificial intelligence and uses Theory of Universal Induction along with control theory to achieve this task. The Universal Induction theory presents the concept of the Solomonoff's Universal prior for sequence prediction, though it is incomputable. The challenge of making a computationally efficient version of the prior (if not optimal) and thereby using it in AIXI is addressed in this paper. The proposed model is then simulated for environments manifesting as *n*-order Markov sources and a suitable extension to general systems has been suggested. The superiority of this time and length bounded prior, in terms of computational complexity, over methods using purely binary strings has been presented.**

Index Terms— AIXI; Universal Induction; Fourier Series; Solomonoff's Prior; non binary representation; Markov Process.

I. INTRODUCTION

AIXI[1] is a universally optimal theory of artificial intelligence which combines the concepts of Universal induction[2,5] and control theory. It deals with producing an agent which gives the most optimal performance (in terms of accumulation of rewards) but in the context of infinite computational resources. This limitation makes it impossible to achieve such an agent with finite resources. Thus was devised the theory AIXI-tl[1] which is asymptotically optimal with the constraint of limited resources, though only up to a non universal additive constant. This constant can be huge in practice, thus making convergence to the true AIXI model extremely slow. Giving up computational optimality with gains in lowering of computational complexity is therefore a justifiable alternative. Nonetheless, the AIXI theory produces an agent that is pareto optimal in choice of its policy which makes it the most attractive option for achieving AGI. AIXI approach uses Solomonoff's universal prior[2,5] as prior bayesian probability for a given input data string to the Universal Turing machine and then determines the policy which gives the maximum expected value (expected future reward sum) with the use of generalized expectation maximization procedure. However, the prior itself is incomputable with limited resources. This paper discusses an approach based on multidimensional Fourier series representation of functions for determination of an approximation to the universal prior with the constraint of limited computational resources and low computational complexity. In this approach, a string is constructed from consecutive non-binary input/output pairs of a Universal Turing Machine as in the case of AIXI-tl model. But unlike AIXI-tl, the Universal prior is calculated by a weighted sum of time and length bounded probability distributions i.e., there exists a finite upper bound to their computational time and Kolmogorov complexity, generated using multidimensional Fourier series followed by determination of the optimal policy using importance sampling of the prior and a modification of the generalized expectation maximization procedure as used with the Monte Carlo[4] approach to AIXI. It can be argued that this approach converges to the true AIXI model in the limit of infinite computational resources.

Manuscript received November 23, 2012.

A. Universal Induction

The problem of predicting a string element x_i, given the previous string elements i.e., $x_{1:i-1}$ sampled from some unknown distribution can be solved using Bayesian inference provided the prior probability of all the possible distributions generating the string is known. In agreement with the Occam's Razor[1], the Solomonoff prior solves the problem of choosing the prior probability in an optimal way. The generalized expression of the prior can be put mathematically as:

$$\xi(x) \equiv \sum_p F(K(p)) \cdot p(x) \qquad (1)$$

The above summation is carried out over all possible distributions (the codes for which are prefix free) that generate the string starting with x with a certain probability $p(x)$ and F is a function of the Kolmogorov complexity $K(p)$ of p which associates a weight to each distribution (semi-measure) based on its (the distribution's) algorithmic probability.

B. AIXI

As illustrated in the figure below, any AI agent can be thought of as a chronological Turing machine p embedded in the environment which can also be modeled as a Turing machine q, where p is a policy which chooses action y_i for the input x_i (o_i below) and q is a policy that outputs reward r_i for every chosen y_i. Both the Turing machines have bidirectional worktapes and unidirectional input and output tapes.

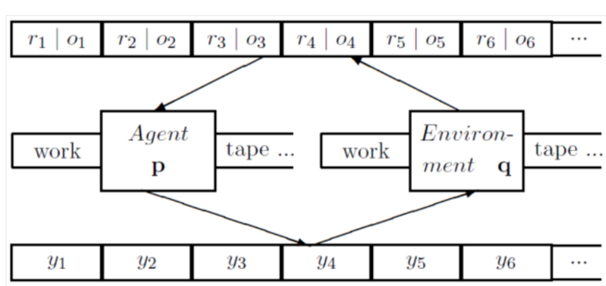


Fig 1. Model of an AIXI agent embedded in an environment where both the agent and environment behave as chronological Turing Machines.

DOI: 10.5176/2251-3043_2.4.201

Equation(2) described below is the mathematical formulation of the expected value of the cumulative future reward of an expectation maximizing agent for known probability distribution μ and the most optimal policy $p*$[1].

$$V_{k,m}^{*\mu}(yx_{<k}) = \max_{y_k} \sum_{x_k} [r(x_k) + V_{k+1,m}^{*\mu}(yx_{1:k})] \cdot \mu(yx_{<k}\underline{yx_k}) \quad (2)$$

The left hand side of the equation is the expected value of sum of future rewards from time step k to the agent's lifetime m with the string history $yx_{<k}$ for an optimal policy p^* and the right hand side expresses the same as the maximum of the sum, of the current reward and the future value multiplied by the probability that string elements y_k, x_k follow the given string history $yx_{<k}$ immediately, taken over all x_k for a particular y_k. AIXI theory works by replacing the known distribution μ with the Solomonoff's prior ξ as ξ converges to the true distribution automatically as the length of the predicted string is extended to infinity.

II. Probability Distributions (Density functions)

A. Use of Fourier basis as Generating Functions

It is known that any periodic function or the periodic repetitions of an aperiodic function can be represented as a weighted sum of all orthogonal basis functions defined over a period. The simplest orthogonal basis set is formed with the trigonometric functions *Sin(nx)* and *Cos(nx)*, and the resulting representation is called Fourier series. The probability distributions can be modeled using the Fourier series, however the required argument of the distribution in this case is a string which comprises of independent elements and hence each element can be treated as a dimension. This finally calls for the use of multi-dimensional Fourier series. The probability distributions (density functions) thus created are only structured for one period of each dimension; hence the string elements have to be scaled accordingly. Henceforth, 'one period of each dimension' will be referred to as a period without the loss of generality.

B. Probability Distribution bounds

In order to ensure that the hyper-geometric curve generated by the Fourier series is indeed a probability distribution function, it should satisfy the following criteria:

$$P(x_1, x_2, \dots x_n) \geq 0 \quad \forall (x_1, x_2, \dots x_n) \qquad \text{(c.1)}$$

$$\oint P(x_1, x_2, \dots x_n) = 1 \ \mid \forall (x_1, x_2, \dots x_n) \in S \qquad \text{(c.2)}$$

The criterion(c.2) can be easily satisfied as the hyper volume of any such curve over a period depends only on the constant term that appears in the Fourier series representation of the curve. The criterion(c.1) can be satisfied with a specific representation of the function which is:

$$H_n(x) = \sum_{k=0}^{n} (a_k \cdot \prod_{j=1}^{k} G_{1,j}(x)) \ \mid \{a_k > 0, G_{1,j} \geq 0\} \quad (3)$$

Here H_n is a non negative n^{th} harmonic function, a_k are arbitrary non negative constants and G_{1j} is itself a first harmonic probability distribution function which is strictly non negative and has only one root over a period which is denoted by the string x'. All functions H_n that satisfy the criterion(c.2) form a mutually exclusive and exhaustive set of n^{th} harmonic probability distributions P_n. The constant term in the Fourier series expansion of a probability distribution should be equal to $\frac{1}{(2\pi)^d}$ for the volume to be equal to 1. The distributions G_1 (Generating distributions) having exactly one root over a period can thus be modeled as:

$$G_1 \equiv \frac{1}{(2\pi)^d} + \sum_{i=1}^{d} a_i \cdot Sin(x_i + \lambda_i) \mid (a_i > 0, 0 \leq \lambda_i < 2\pi, \sum a_i = \frac{1}{(2\pi)^d}) \quad (4)$$

III. Prior in the Fourier basis

In this section, a construct of the universal prior based on the generated distributions will be hypothesized and its convergence to the true distribution is presented for certain simulation parameters in the next section. Before we construct the prior, it is very necessary to define the Kolmogorov complexity of each generated distribution.

A. Kolmogorov Complexity of Fourier distributions

The Kolmogorov complexity of a distribution can be defined as the length of the shortest program that can describe it completely where each program is itself a string of certain alphabet and each such code is prefix-free[5]. In this case, the number of harmonics in the distribution can be called its complexity and the alphabet is thus made of functions G_1. It is well known that each block code is prefix-free. Hence, for each such program to be prefix-free, it is in interest to consider distributions only belonging to a particular harmonic n and to form a combined semi measure of all such distributions with a particular string of G_1's (bearing same multiplicative constants but different phases) before its algorithmic probability can be evaluated. It can be quite easily shown that such semi measures of lower harmonics are nothing but the normalized sum of the semi measures of higher harmonics bearing the same prefix (the lower harmonic semi measure). The string G' thus serves the purpose of a prefix free code for the n^{th} harmonic.

B. The Ultimate Dimension 'D'

From equation(5) and from laws of permutations and combinations, the number of solutions for $\sum a_i = \frac{1}{(2\pi)^d}, a_i = n\varepsilon$, (n=1,2,3..) and hence the number of alphabets (Generating distributions) G_1 possible for dimensions d is given by:

$$N_d = \binom{m-1}{d-1} \mid \frac{1}{(2\pi)^d} = m\varepsilon + k, k < \varepsilon \quad (5)$$

where m is an integer. The convergence entails the following relation between ε and D (because π is irrational):

$$\varepsilon = \frac{p}{\lfloor p \cdot D(2\pi)^D + 1 \rfloor}, p > \frac{1}{(2\pi)^D} \quad (6)$$

where D is the approximation to infinity for a computer with given ε. The dimensions cannot be increased beyond the limiting dimension D for a given ε and p.

C. Choice of semi measure

It is apparent that functional groupings of the generated distributions can also behave as distributions themselves. Such a grouping of all distributions of harmonic n and a particular string sequence G' of functions G_1 is of prime interest. Each such grouping is termed as a semi measure $I_{i,n}$. The algorithmic probability of each $I_{i,n}$ according to Solomonoff's Induction theory[2,5] is given by:

$$F(n) = \frac{1}{(N_d)^n} \qquad (7)$$

The formula for the (harmonic bounded) prior can be now be written as (normalized sum of probability distributions, hence satisfies Kraft's inequality to be a valid prior):

$$\xi_{n,d}(x) = \sum_i \frac{1}{(N_d)^n} \cdot I_{i,n}(x) \qquad (8)$$

The value of N_d converges to 1 as the number of dimensions is increased to $d = D$. As the number of functions G_1 becomes 1 at the limiting dimension, only one semi measure $I_{D,n}$ can be constructed. The algorithmic probability of $I_{D,n}$ is equal to 1 according to equation(5). This semi measure $I_{D,n}$ is bounded both in length (complexity) and time. As ε tends to zero, $I_{D,n}$ reaches the true prior and hence the true distribution. However, its value cannot be computed, as the true prior is not bounded either in length or time. Nonetheless, the measure $I_{D,n}$ is a close approximation to the true prior in the limit of finite resources, i.e., D, n and ε.

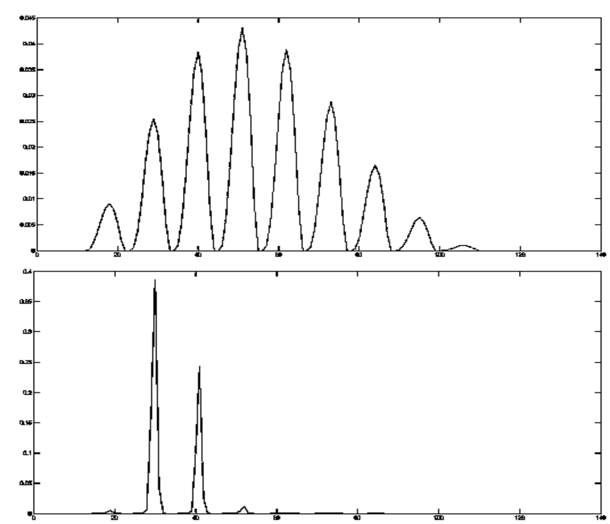
IV. Simulation Results

The implementation of above algorithm for environments sampling from an i.i.d source is obvious. It is interesting to implement the proposed algorithm for n order Markov sources and extend it further to general non-Markovian environments. The results for order-1 source show how with increase in number of event samples, the prior calculated as per eq. 8 attains the conditional probability of the true distribution. Results for higher order Markov sources have not been shown due to space constraints.

A. Simulation Parameters

1. Environment sampling from a 1st order Markov Source.
2. "2 state 2 action" agent environment system.
3. Further discretization of state and action spaces for computability using n-point DFT instead of continuous Fourier series. (2 point DFT used for example simulation i.e., without any interpolation)
4. True sequence of events generated by Monte-Carlo sampling from the true sampling distribution.

B. Results



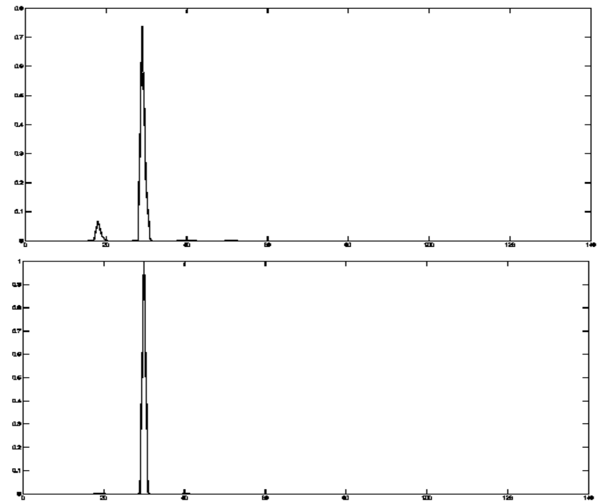


Fig 2. Fraction of conditional probabilities corresponding to each Fourier distribution in the conditional probability of the harmonic bounded prior after 6, 20, 200 and 400 samples respectively.

V. Application to AIXI

The prior $\xi_{d,n}$ can be substituted in (2) of the AIXI model for a finite lifetime of the agent $m = d$ and the normal expectation maximization procedure can be carried out to determine the best policy p_*. However, the computational time grows exponentially with m rendering it almost impossible to achieve a good policy p_* in finite time with given resources for all but very small m. This calls for the use of a procedure called Importance Sampling used in the Monte Carlo approach to AIXI[4]. The outputs and inputs have to be learnt equally in this case which means now the argument of the prior is a string xy and not just $xy_{<0}\underline{xy}$. It can be seen that this procedure will converge to the true AIXI as the sampling probability $P_{i,m}$ tends to Infinity and the ε is reduced to zero. The detailed procedure is as follows:

At each time step k, the strings xy are sampled from all possible strings with the sampling probability $P_{i,m}$ proportional to $\xi_{d,n}(xy)$ using some MCMC(Markov Chain Monte Carlo) Multivariate sampling algorithm such as Metropolis Hastings, Splice sampling, etc. By choosing appropriate parameters, localization of samples can be avoided which is a drawback of MCMC algorithms. The i^{th} sampled string is $xy^{(i)}$and its expected value is given by equation:

$$V^{+}_{k,m}\left(xy^{(i)}\right) = \left[\sum_t r(x_t)\right] \quad \forall\, x_t \in (\, xy^{(i)}) \qquad (9)$$

The following steps are then followed recursively for $k <= j <= m$ starting with $j = m$ to determine the value of y^*:

(a) $y^*_j = \arg\max_{y_j} V^{+}_{k,m}(xy_{k:j})$

(b) $P_{i,j-1} = \begin{cases} P_{i,j}, & if\ y^{(i)}_j = y^*_j \\ 0, & otherwise \end{cases}$

(c) $V^{+}_{k,m}\left(xy_{k:j}\right) = \dfrac{\sum_{i \mid xy^{(i)}_{k:j-1} = xy_{k:j-1}} V^{+}_{k,m}(xy^{(i)}_{k:j}) \cdot P_{i,j-1}}{\sum_{i \mid xy^{(i)}_{k:j-1} = xy_{k:j-1}} P_{i,j-1}}$

VI. ADVANTAGES OVER BINARY REPRESENTATION

The computational time complexity of the prior constructed using binary prefix codes bounded by length l' and time t' (and also that of Levin's Universal search) is of the order:

$$t\left(\xi^{t',l'}(xy_{1:k})\right) = O(2^{l'} \cdot t') \tag{10}$$

Hypothesizing that the true environment follows a harmonic bounded Fourier distribution (binary axiomatic systems being the quantized counterpart), out of all these $2^{l'}$ programs, only the ones expressible in Fourier series representation using exactly n harmonics are generated for constructing the prior which reduces the time for computing the prior by a considerable margin. Even more reduction in time can be had by further reducing n, though at the cost of accuracy. The time complexity for the prescribed method is of the order:

$$t\left(\xi^{t',n}(xy_{1:k})\right) = O((N_d)^n \cdot t') \tag{11}$$

where t' in this case is the maximum amongst the time required for evaluating each $I_{i,n}$. Further combining this method with the Importance Sampling procedure suggested by the Monte Carlo approach[4] reduces the policy computation time by a great extent.

VII. CONCLUSIONS

This paper presents a novel method for computation of the time and length bounded Universal prior using a non binary Fourier representation of the same and also provides a proof for its convergence to the true prior in limit of infinite computational resources. Simulation results for n order Markov processes are seen to validate our hypothesis and can be generalized in the limit $n \to \infty$. Further, an approach called importance sampling is presented, as inspired from[4], and is applied to the prior thus generated for efficient policy computation in AIXI. It is premised to have a considerable advantage over the methods employing binary representation of programs in terms of computational time complexity as understood from equations (10) and (11). The implementation of the method is not presented here and the results of the same will be published soon. An immediate goal would be to use this approach in real world environments and with limited computational resources in terms of ε, n and D.

REFERENCES

[1] M. Hutter. Universal Algorithmic Intelligence, *Technical Report IDSIA-01-03*, 2003, pp. 16-32, 47-52.

[2] R. J. Solomonoff. A formal theory of inductive inference: parts 1 and 2. *Information and Control*, 1964, pp. 7:1-22, 224-254.

[3] M. Hutter. The fastest and shortest algorithm for all well-defined problems. *International Journal of Foundations of Computer Science*, 2002, pp. 13(3):431-443.

[4] S. Pankov. A computational approximation to the AIXI model. *AGI-08,* 2008.

[5] S. Legg. Solomonoff Induction. 2004.

[6] M. Hutter. New error bounds for Solomonoff prediction. *J. Computer and System Science,* 2001, pp. 62(4):653-667.

[7] S. Shah. Time an Length bounded AIXI for non binary strings using multidimensional Fourier series, *ATAI 2012,* 2012.

Shah Swapnil (M'2012), born April 1990 at Anand in the state of Gujarat, India received his bachelors degree in electronics and communication from National Institute of Technology, Surat, India in August 2012. He currently functions as a business technology analyst at Deloitte Touche Tohmatsu, India Pvt. Ltd.

Right from his undergraduate years, he has worked on numerous projects in generic robot control, neural learning, image processing, machine vision, distributed intelligence and swarm behavior. Quite recently, he shifted his focus to developing computable solutions to the general AI problem or Strong AI i.e., of human-like cognition by a machine and has published papers with international conferences held by organizations such as Inter Science research network, GSTF, and World Science and Engineering Academy and Society(WSEAS) on topics ranging from automated simulation of sketched circuits using computer vision to extension of deterministic agent – environment AI model to the probabilistic framework of Quantum Physics.