Partial Operator Induction with Beta Distribution

Nil Geisweiller

AGI-18
Prague

Problem:

Combining Models from Different Contexts

Theory:

Solomonoff Operator Induction and Beta Distribution

Practice:

Inference Control Meta-Learning

Outline

Problem:

Combining Models from Different Contexts

Theory:

Solomonoff Operator Induction and Beta Distribution

Practice:

Inference Control Meta-Learning

Outline

Problem:

Combining Models from Different Contexts

Theory:

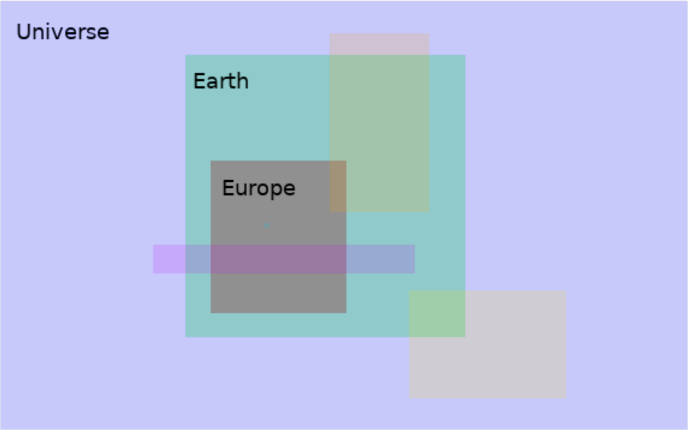
Solomonoff Operator Induction and Beta Distribution

Practice:

Inference Control Meta-Learning

Problem: Models from different contexts

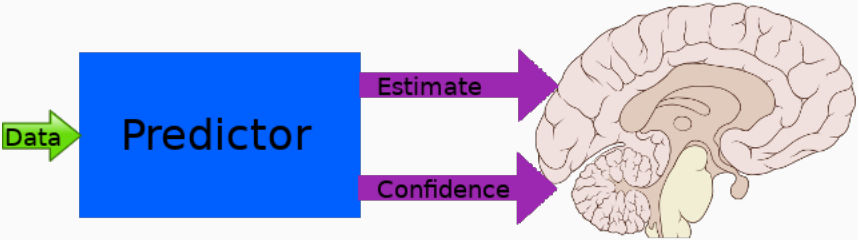
How to combine models obtained from different contexts?



Large Contexts → Underfit

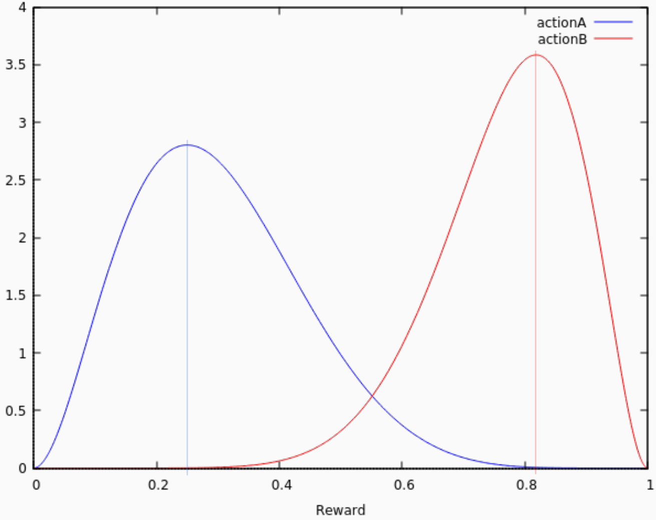
Small Contexts → Overfit

Problem: Preserve Uncertainty



Problem: Preserve Uncertainty

Exploration vs Exploitation (Thompson Sampling)



Problem: ImplicationLink

```
ImplicationLink &lt;TV>
  R
  S
```

$\equiv$

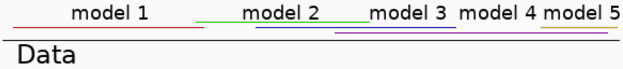
Second Order
$P(S|R)$

Beta Distribution in disguise

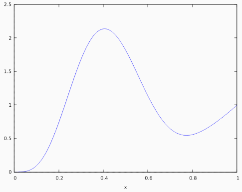
Solution

Bayesian Model Averaging / Solomonoff Operator Induction, modified to:

1. Support partial models



2. Produce a probability distribution estimate, rather than probability estimate.



3. Specialize for Beta distributions

Outline

Problem:

Combining Models from Different Contexts

Theory:

Solomonoff Operator Induction and Beta Distribution

Practice:

Inference Control Meta-Learning

Outline

Problem:

Combining Models from Different Contexts

Theory:

Solomonoff Operator Induction and Beta Distribution

Practice:

Inference Control Meta-Learning

Solomonoff Operator Induction

Bayesian Model Averaging + Universal Distribution

Probability Estimate:

$$\hat{P}(A_{n+1}|Q_{n+1}) = \sum_j a_0^j \prod_{i=1}^{n+1} O^j(A_i|Q_i)$$

where:

- Q_i = i^{th} question
- A_i = i^{th} answer
- O^j = j^{th} operator
- a_0^j = prior of j^{th} operator

Beta Distribution Operator

Specialization of Solomonoff Operator Induction

OpenCog implication link

```
ImplicationLink <TV>
  R
  S
```

$$\equiv$$

Class of parameterized operators

$$O_p^j(A_i|Q_i) = \text{if } R^j(Q_i) \text{ then } \begin{cases} p, & \text{if } A_i = A_{n+1} \\ 1-p, & \text{otherwise} \end{cases}$$

Beta Distribution

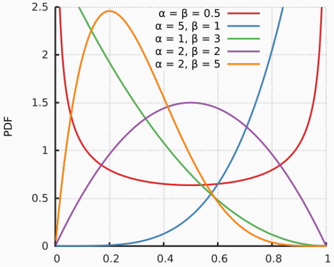
Probability Density Function:

$$pdf_{\alpha,\beta}(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$$

Beta Function:

$$B_x(\alpha,\beta) = \int_0^x p^{\alpha-1}(1-p)^{\beta-1} dp$$

$$B(\alpha,\beta) = B_1(\alpha,\beta)$$



Conjugate Prior:

$$pdf_{m+\alpha,n-m+\beta}(x) \propto x^m(1-x)^{n-m} pdf_{\alpha,\beta}(x)$$

Artificial Completion

$$O^j_p \quad (A_i|Q_i) = \quad \text{if } R^j(Q_i) \text{ then } \begin{cases} p, & \text{if } A_i = A_{n+1} \\ 1-p, & \text{otherwise} \end{cases}$$

Data

Artificial Completion

$$O^{j}_{p,C}(A_i|Q_i) = \begin{array}{l} \text{if } R^{j}(Q_i) \text{ then } \begin{cases} p, & \text{if } A_i = A_{n+1} \\ 1-p, & \text{otherwise} \end{cases} \\ \text{else } C(A_i|Q_i) \end{array}$$

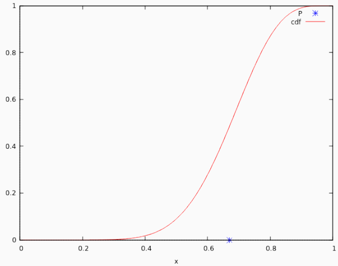
Second Order Solomonoff Operator Induction

Probability Estimate:

$$\hat{P}(A_{n+1}|Q_{n+1}) = \sum_j a_0^j \prod_{i=1}^{n+1} O^j(A_i|Q_i)$$

Probability Distribution Estimate:

$$\hat{cdf}_{(A_{n+1}|Q_{n+1})}(x) = \sum_{O^j(A_{n+1}|Q_{n+1}) \leq x} a_0^j \prod_{i=1}^{n} O^j(A_i|Q_i)$$



Combing Solomonoff Operator Induction and Beta Distributions

$$\hat{cdf}_{(A_{n+1}|Q_{n+1})}(x) \propto \sum_j a_0^j r^j B_x(m^j+\alpha, n^j-m^j+\beta)B(m^j+\alpha, n^j-m^j+\beta)$$

where

- n^j = number of observations explained by j^{th} model
- m^j = number of true observations explained by j^{th} model
- r^j = likelihood of the unexplained data

$r^j =$???

Combing Solomonoff Operator Induction and Beta Distributions

$$\hat{cdf}_{(A_{n+1}|Q_{n+1})}(x) \propto \sum_j a_0^j r^j B_x(m^j+\alpha, n^j-m^j+\beta)B(m^j+\alpha, n^j-m^j+\beta)$$

where

- n^j = number of observations explained by j^{th} model
- m^j = number of true observations explained by j^{th} model
- r^j = likelihood of the unexplained data

$r^j = ??? \approx 2^{-v^{(1-c)}}$

- $v = n - n^j$ = number of unexplained observations
- c = compressability parameter
  - $c = 1 \rightarrow$ explains remaining data
  - $c = 0 \rightarrow$ can't explain remaining data

Outline

Problem:

Combining Models from Different Contexts

Theory:

Solomonoff Operator Induction and Beta Distribution

Practice:

Inference Control Meta-Learning

Outline

Problem:

Combining Models from Different Contexts

Theory:

Solomonoff Operator Induction and Beta Distribution

Practice:

Inference Control Meta-Learning

Inference Control Meta-learning

Learn how to reason efficiently

Inference Control Meta-learning

Learn how to reason efficiently

Methodology:

1. Solve sequence of problems (via reasoning)

Inference Control Meta-learning

Learn how to reason efficiently

Methodology:

1. Solve sequence of problems (via reasoning)
2. Store inference traces

Inference Control Meta-learning

Learn how to reason efficiently

Methodology:

1. Solve sequence of problems (via reasoning)
2. Store inference traces
3. Mine traces to discover patterns

Inference Control Meta-learning

Learn how to reason efficiently

Methodology:

1. Solve sequence of problems (via reasoning)
2. Store inference traces
3. Mine traces to discover patterns
4. Build control rules

```
Implication <TV>
  And
    <inference-pattern>
    <rule>
  <good-inference>
```

Inference Control Meta-learning

Learn how to reason efficiently

Methodology:

1. Solve sequence of problems (via reasoning)
2. Store inference traces
3. Mine traces to discover patterns
4. Build control rules

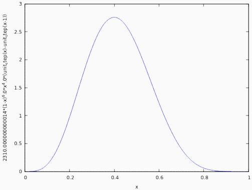
```
Implication <TV>
  And
    <inference-pattern>
    <rule>
  <good-inference>
```

5. Combine control rules to guide future reasoning

Combine Control Rules

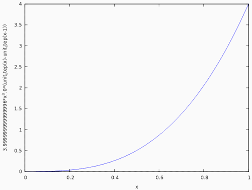
```
Implication <TV1>
  And
    <inference-pattern-1>
    deduction-rule
  <good-inference>
```

```
Implication <TV2>
  And
    <inference-pattern-2>
    deduction-rule
  <good-inference>
```

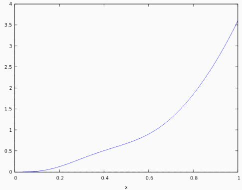


$$\oplus$$

$$c = 1$$



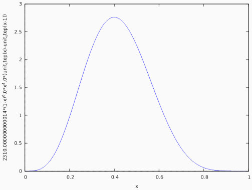
$$=$$



Combine Control Rules

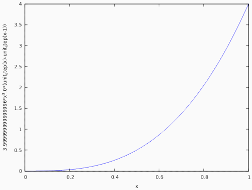
```
Implication <TV1>
  And
    <inference-pattern-1>
    deduction-rule
  <good-inference>
```

```
Implication <TV2>
  And
    <inference-pattern-2>
    deduction-rule
  <good-inference>
```

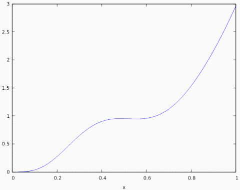


$\oplus$

$c = 0.5$

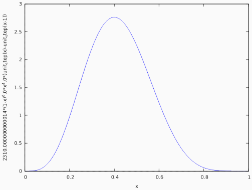


$=$



Combine Control Rules

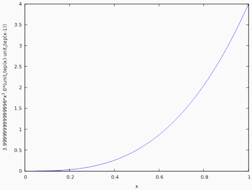
```
Implication <TV1>
  And
    <inference-pattern-1>
    deduction-rule
  <good-inference>
```



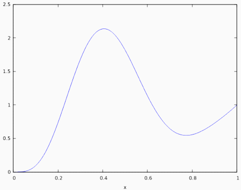
$\oplus$

$c = 0.1$

```
Implication <TV2>
  And
    <inference-pattern-2>
    deduction-rule
  <good-inference>
```



$=$



Conclusion

Contribution:

- Second Order Solomonoff Operator Induction
- Specialized for Beta Distribution
- Attempt to Deal with Partial Models

Conclusion

Contribution:

- Second Order Solomonoff Operator Induction
- Specialized for Beta Distribution
- Attempt to Deal with Partial Models

Future Work:

- Improve Likelihood of Unexplained Data
- More Experiments (Inference Control Meta-learning)

Conclusion

Contribution:

- Second Order Solomonoff Operator Induction
- Specialized for Beta Distribution
- Attempt to Deal with Partial Models

Future Work:

- Improve Likelihood of Unexplained Data
- More Experiments (Inference Control Meta-learning)

Thank you!