東京大学 工学部 計数工学科/物理工学科

応用音響学：音声認識：Bayes の定理

嵯峨山 茂樹 ＜sagayama@hil.t.u-tokyo.ac.jp＞
東京大学 工学部 計数工学科
資料所在 http://hil.t.u-tokyo.ac.jp/～sagayama/applied-acoustics/

Bayes の定理：確率的逆問題の解法

- **原因、結果: それらが同時に起こる確率は**

$$P(\textbf{原因}, \textbf{結果}) = P(\textbf{結果}|\textbf{原因})P(\textbf{原因}) = P(\textbf{原因}|\textbf{結果})P(\textbf{結果})$$

- **事前確率と事後確率**

$$P(\textbf{原因}|\textbf{結果}) = \frac{P(\textbf{結果}|\textbf{原因})P(\textbf{原因})}{P(\textbf{結果})}$$

- **事後確率最大化:**
  ある結果に対して複数の原因が考えられる → 最も確率の高い原因は？

$$\hat{\textbf{原因}} = \underset{\textbf{原因}_i}{\operatorname{argmax}}\, P(\textbf{結果}|\textbf{原因}_i)P(\textbf{原因}_i)$$

- **音声認識の場合：**
  原因： 意図された発声内容
  結果： 音声信号

Bayes の定理：確率的逆問題の解法

- **原因W、結果Y: それらが同時に起こる確率は**

$$P(W,Y) = P(Y|W)P(W) = P(W|Y)P(Y)$$

- **事前確率と事後確率**

$$P(W|Y) = \frac{P(Y|W)P(W)}{P(Y)}$$

- **事後確率最大化**

$$\hat{W} = \underset{W_i}{\operatorname{argmax}}\, P(Y|W_i)P(W_i)$$

- **熊雄は、雨が降ると必ず傘を持って来るが、晴れた日でも２割の確率で持って来る。ところで本日の降水確率は30%であった。では、熊雄が傘を持って現れたら、外は雨が降っているのかそうでないのか。**

$$P(\textbf{雨}|\textbf{傘}) = \frac{1.0 \times 0.3}{P(\textbf{傘})} = \frac{0.3}{P(\textbf{傘})}$$

$$P(\textbf{晴}|\textbf{傘}) = \frac{0.2 \times 0.7}{P(\textbf{傘})} = \frac{0.14}{P(\textbf{傘})}$$

Thomas Bayes

Thomas Bayes : 初期の確率論の確立

Born: 1702 in London, England

Died: 17 April 1761 in Tunbridge Wells, Kent, England

Thomas Bayes was ordained, a Nonconformist minister like his father, and at first assisted his father in Holborn. In the late 1720s he became minister of the Presbyterian Chapel in Tunbridge Wells, 35 miles southeast of London. Bayes remained minister at Tunbridge Wells until 1752 when he retired, but continued to live in Tunbridge Wells.

Bayes set out his theory of probability in Essay towards solving a problem in the doctrine of chances published in the Philosophical Transactions of the Royal Society of London in 1764. The paper was sent to the Royal Society by Richard Price, a friend of Bayes'.

Bayes's conclusions were accepted by Laplace in a 1781 memoir, rediscovered by Condorcet (as Laplace mentions), and remained unchallenged until Boole questioned them in the Laws of Thought. Since then Bayes' techniques have been subject to controversy.

(George Boole: Born: 2 Nov 1815 in Lincoln, Lincolnshire, England / Died: 8 Dec 1864 in Ballintemple, County Cork, Ireland)

George Boole

George Boole

Born: 2 Nov 1815 in Lincoln, Lincolnshire, England
Died: 8 Dec 1864 in Ballintemple, County Cork, Ireland

In 1854 he published An investigation into the Laws of Thought, on Which are founded the Mathematical Theories of Logic and Probabilities. Boole approached logic in a new way reducing it to a simple algebra, incorporating logic into mathematics. He pointed out the analogy between algebraic symbols and those that represent logical forms. It began the algebra of logic called Boolean algebra which now finds application in computer construction, switching circuits etc.

Boole also worked on differential equations, the influential Treatise on Differential Equations appeared in 1859, the calculus of finite differences, Treatise on the Calculus of Finite Differences (1860), and general methods in probability. He published around 50 papers and was one of the first to investigate the basic properties of numbers, such as the distributive property, that underlie the subject of algebra.

Many honours were given to Boole as the genius in his work was recognised. He received honorary degrees from the universities of Dublin and Oxford and was elected a Fellow of the Royal Society (1857). However his career, which was started rather late, came to an unfortunately early end when he died at the age of 49.

Boole's wife (Mary - niece of Sir George Everest, after whom the mountain is named) believed that a remedy should resemble the cause. She put Boole to bed and threw buckets of water over the bed since his illness had been caused by getting wet.

音声認識の原理

音声認識とは？

許されるあらゆる発声内容の仮説の中で、発声された音声をもっともよく説明できるものを探す問題

構成要素

- 音声分析 — 音声波形から有用な音声の特徴量のパターンを抽出する
- 音素モデル — 言語の単位(音素など)の音響的性質をモデル化する
- 言語モデル — 探索すべき解(認識結果)の探索空間を規定する
- 解探索過程 — 定義された探索空間中で最適解を効率良く探す

音声認識の構成要素

- 音声認識
  - 音声分析：MFCC
  - 音響モデル：HMM
  - 言語モデル：Word n-gram
  - 探索過程：Viterbi Decoder
- 特徴
  - 日本(NTT)が活躍した分野
  - 迫江(NEC): DPマッチング、板倉(NTT): Itakura-Saito距離

連続音声認識の構成/原理

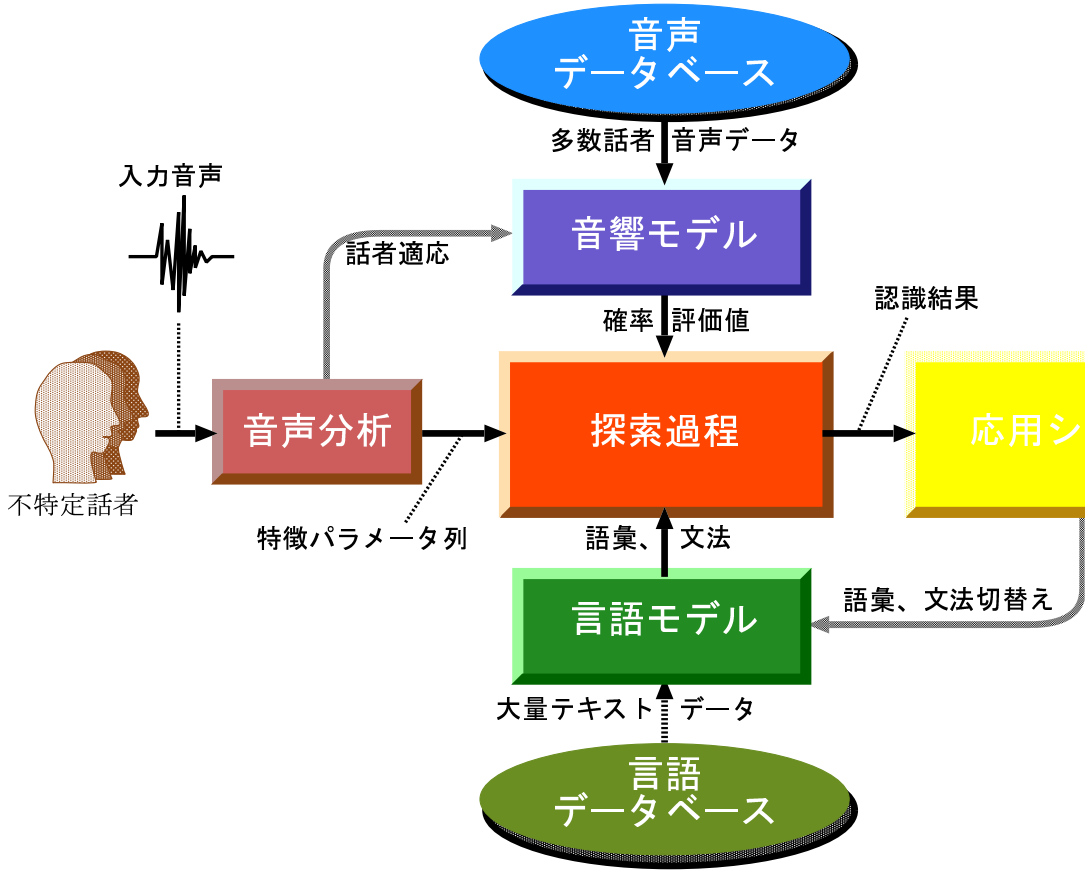

図1. 連続音声認識の構成要素。音声分析により音声特徴を抽出し、語彙や文法によって音素の並びを規定する言語モデルのもとで、音素を確率モデルにより表現する音響モデルによって入力の確率評価をして、事後確率最大の経路を効率良く探索する。