

## マルコフ確率場を用いた調認識、自動和声付け、および自動対位法

米 田 隆 一<sup>†</sup> 西 本 卓 也<sup>†</sup> 嵯峨山 茂樹<sup>†</sup>

本研究では、Standard MIDI File, 楽譜等のシンボリックな音楽情報を入力として対旋律、和音、調等のラベルを付与する汎用的な手法を提案する。このような音楽のラベル付与問題は、音声認識における言語モデルとの類似性から、マルコフモデル、および生成モデルとしての HMM (hidden Markov models) を適用することが多かった。本研究でも確率モデルを踏襲するが、MIDI、楽譜などは、縦の和音、横の声部進行等、2 次元的な情報であり、マルコフ連鎖のような一次元的なモデリングでは不十分である。このような背景のもと、より広いコンテキストを重視すべきであるという考えが生まれる。マルコフ確率場 (Markov random field) モデルは、ノード間の関連をエッジであらわす一種の無向グラフモデルであり、エッジを設計することにより広いコンテキストを柔軟に設定できる。また、マルコフ確率場では我々の持つ音楽的知識を素性関数の設計 (エッジの設計) という操作に還元することが可能、という利点を持つ。さらに、識別モデルであるという点で HMM とは異なり、可能なすべての出力系列を入力系列の条件付き確率として求めることができるという点で、条件付き確率場 (conditional random fields) とも呼ばれる。本研究ではマルコフ確率場モデルに基いたさまざまな音楽情報処理に対するアプローチを論じる。本手法は汎用であり、応用範囲は多岐にわたるが、本稿では対旋律付け、和声付け、ドミナント定型句の同定、和声解析、調認識に対するアプローチを述べる。考えられる他の応用としては、リズム認識、楽曲のバージング、音楽情報検索など広範囲にわたる。

### Markov Random Fields Applied to Key Finding, Automatic Harmonization and Automatic Counterpoint

RYUICHI YONEDA,<sup>†</sup> TAKUYA NISHIMOTO<sup>†</sup> and SHIGEKI SAGAYAMA <sup>†</sup>

This paper describes a Markov random field approach for labeling tasks such as key finding, automatic harmonization, and automatic counterpoint. Most of our previous work deal with these labeling problems using Markov models, and hidden Markov models as a generative model, analogous to language models of automatic speech recognition. This work follows our previous probabilistic approach. 1-dimensional Markov chain, however, can not model large musical context like scores and SMFs (standard MIDI files) which have rich information in that chords and melodies spread 2-dimensionally. Thus the ability of dealing with large context is crucial in music modeling. Markov random field is a kind of undirected graphical models in which edges are regarded as relevances between nodes. The design of edges (feature functions) corresponds to our music knowledge in which nodes (symbols) have relevances with each other. Markov random field is also called as conditional random fields (CRFs) in that all possible output sequences have conditional probabilities given an input sequence. In other words, CRFs is a discriminative model. Although the proposed model can be applied to many purposes, this paper describes the approaches to five tasks: counterpoint, harmonization, dominant chunk detection, harmony analysis, and key finding. Other possible applications include rhythm recognition, musical parsing, and music information retrieval.

#### 1. はじめに

本稿では、MIDI、楽譜等のシンボリックな音楽情報を入力としてラベル (対旋律、和音、調等) を付与する汎用的な手法を提案する。我々の研究グループでは、このようなラベル付け問題を取り扱う際、音声認識に

おける言語モデルと同様、マルコフモデルを適用することが多かった。しかし、MIDI、楽譜などの 2 次元的な情報 (縦の和音、横の声部進行等) は、マルコフ連鎖のようなモデルでは不十分であり、より広いコンテキストの重視が必須である。このような背景のもと、広いコンテキストを柔軟に設定でき、かつ、我々の持つ音楽的知識を素性関数の設計という操作に還元することができるマルコフ確率場 (最大エントロピーモデル) に基いたアプローチを論じる。

<sup>†</sup> 東京大学大学院情報理工学系研究科  
Graduate School of Information Science and Technology,  
The University of Tokyo.

## 2. 音楽の確率定式化

### 2.1 一般化した数理構造

図1は、マルコフ確率場の概念を幾何学的に表現したものである。各ノードは音高、和音等のラベルを表し、黒のノードは現在付与しようとするラベルである。矢印はノード同士が関係を持っていることを表す。これは、MIDI、楽譜が持つ2次元的な情報によくマッチする。すなわち、旋律、和音等を作成する際、縦の和音構成、横の声部進行共に考慮にいれなければならない。さらに、音楽の持つ繰返し、模倣等の構造を考えると、より遠くのコンテキストまで影響が及んでいると考えられる。この点で、言語と比較して圧倒的に語彙サイズが小さいにもかかわらず音楽がリッチな表現力を持ちうる。

### 2.2 対位法、和声法の音楽モデル

対位法は、複数の旋律の組み合わせによる作曲法であり、単純な作曲法としては、与えられた旋律に適合する対旋律の付与(2声対位法)がある。2声の対位法においては、2声間の音程の制約、平行5度、同一音程の4度以上連続の禁止等の制約がある。つまり、与えられた旋律のある箇所の数音符前後の文脈により対旋律決定が可能だといえる。和声付けとは、与えられた旋律に適合し楽曲として適切な和音進行を付与することである。和声学においては、典型的な和声の終止定型が存在し、バス旋律より和音系列の決定が可能である(e.g. fa-so-so-do に対する II-I-V-I 曲末)。これも旋律の数音符前後の文脈により和音の決定が可能であるといえる。

また、入力は1次元的な旋律に限らず、2次元的なものもある。例えば、和声解析は複数旋律を入力として和音、非和声音等の種類を決定する。入力が複数旋律を持つという点で2次元的であり、非和声音の種類の決定は和音と同様、文脈依存である。また、楽曲のやや浅い構造解析法として、和音系列、和音機能(e.g. トニック、ドミナント)、カデンツをこの順に段階的に付与する。つまり、構造を持つ点で1次元的でない。ここでも、和音機能は和音系列に依存し、カデンツの箇所は和音機能に依存しているといえる。

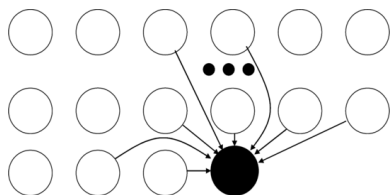


図1 音楽を一般化した数理的な構造

我々は、対旋律付け/和声付け問題を、定旋律が与えられた下での対旋律/和音系列の条件付き確率を最大にする問題として捉える。この問題における文脈依存性の設計は、マルコフ確率場モデルの素性関数の設計、つまりノード同士の関連の設計により可能である。

### 2.3 調認識の音楽モデル

調は、西洋音楽の音階に基づいて決められるものである。調を認知するには、音楽の旋律、分散和音などの音高情報を認識しその音高情報をもとに音楽的知識を用いて調を推測すると考えられる。従来から、tone profile による手法(Krumhansl-Schmuckler 法<sup>1)</sup>)等、音高のヒストグラムに基づく手法が主流であり、本研究でも同様のアプローチをとる。マルコフ確率場モデルでは、12種の音高に対応するノードと、調のノードを用意し、音高と調の間に関連があると考え(図2)。エッジは音高頻度に対応し、直観的には、音高頻度が高いほど関連が高いと考える。

## 3. マルコフ確率場モデル

### 3.1 マルコフ確率場

本研究で用いる確率モデルは、呼称は分野により異なり、概してコンピュータビジョンや画像処理の分野ではマルコフ確率場、自然言語処理やバイオインフォマティクスの分野では最大エントロピーモデル、conditional random fields、Markov network 等の呼称がある。

マルコフ確率場<sup>2)</sup>は、周辺の文脈に依存する値をモデル化するのに適した手法である。図1において、近隣同士の関係をグラフのエッジとみなしたとき、クリーク(完全部分グラフ)集合 $\mathcal{C}$ が構成される。確率は、各クリーク $c$ に対応するポテンシャル関数 $V_c$ のlog-linearモデルになるといわれている(Hammersley-Clifford's theorem)。

$$P(f) = \exp \left( \sum_{c \in \mathcal{C}} V_c(f) \right) / Z$$

我々の音楽情報解析では、 $V_c$ を素性関数 $f_i$ とそれに対応する重み $\lambda_i$ の積であるとみなす。これは、最大エントロピーモデル<sup>3)</sup>(以下、ME)における確率分布関数の式と本質的に同一である。

$$p_{\lambda}(y|x) = \exp \left( \sum_i \lambda_i f_i(x, y) \right) / Z_{\lambda}(x) \quad (1)$$

(ただし $Z$ は正規化項 $Z_{\lambda}(x) = \sum_y \exp \left( \sum_i \lambda_i f_i(x, y) \right)$ )  
素性関数 $f_i$ は通常、事例の有無を表す2値関数を考える。式(2)は和声付けで考える素性関数のひとつであり、図1の黒のノードが現和音に、エッジで張られた近隣のノードのひとつが前和音に相当する。黒のノ

ドの候補 (この例では和音の候補) をすべて考慮するという点で、マルコフ確率場モデルは識別モデルであり、我々の研究グループが扱ってきた生成源 (隠れ状態) を仮定する生成モデルとは異なる。

$$f(\text{前和音}, \text{現和音}) \equiv \begin{cases} 1 & \text{if 前和音} = V \text{ の和音 \&\& \\ & \text{現和音} = I \text{ の和音} \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

### 3.2 MEMM, CRF

Maximum Entropy Markov Model<sup>4)</sup> (以下、MEMM) は ME の順次適用により、確率の積が最大となるラベル系列を最適解とみなす手法である。最適解の探索にはビームサーチ等が適用できる。Conditional Random Fields<sup>5)</sup> (以下、CRF) は入力系列そのものを入力とするグローバルな最適解を求める手法であり、ME の特殊形、かつ、HMM の一般化となっている (附録参照)。図 2 は、概念を幾何学的に表現したものである。

CRF では、可能な解系列の数は入力系列に対し指数的に増えるが、動的計画法の一種である forward-backward アルゴリズムにより効率よい計算が実現可能である<sup>6)</sup>。また、式 (1) におけるパラメータ  $\lambda$  の推定には反復スケリング法<sup>7)</sup> や準ニュートン法<sup>8)</sup> の一種 (L-BFGS<sup>8)</sup>) が適用できる。実験には、汎用の分類器、チャンカー を用いた。両者ともパラメータ推定に L-BFGS が使用されている。なお、式 (1) におけるパラメータ  $\lambda$  の事前分布として、通常、正規分布を仮定<sup>9)</sup> し、ハイパーパラメータ (正規分布の分散) の値を変化させながら交差検定し、最適な値を選択する。

## 4. 評価実験

4.1 節、4.2 節、4.3 節において、Humdrum Toolkit<sup>10)</sup> に付属する和声ラベル付きのバッハのコラール 16 曲 (humdrum-kern フォーマット) をすべて階名に変換したものを評価実験に用いた。音符を持たない弱拍部 (連続する 8 分音符をもつソプラノ旋律に対応する 4 分音符のバス旋律の弱拍部) 等には、ダミーのラベルを与えた。また、曲頭、曲末にもまた別のダミーのラベルを与えた。

### 4.1 対旋律付け

ソプラノ音自身と前後のソプラノ音をコンテキストとし、バス旋律を CRF で推定するバス生成タスクを行った。学習データはソプラノ・バスの旋律ペアとし

た。テスト旋律 (e.g. ベートーヴェン第 9 交響曲第 4 楽章の主旋律) に対旋律付けしたところ、妥当な音符系列が確認された (図 3)。

### 4.2 和声付け

ソプラノ音自身と前後のソプラノ音をコンテキストとし、和音系列を CRF で推定する和声付けタスクを行った。学習データはソプラノ旋律と和音のペアとした。テスト旋律を和声付けしたところ、妥当な和音系列が確認された (図 4)。

正解は一意ではないので、定量的評価として、前和音、現在音、前音をコンテキストとする現在和音の推定を ME で行なった。評価は、1 曲とそれ以外に分割する交差検定で行った。ソプラノ音と和音のペアの学習、バス音と和音のペアの学習それぞれにおいて、和音の正解率は 61%, 64% であった。

### 4.3 和声解析

各声部の音自身とその前後音、前和音をコンテキストとし、和音系列を CRF で推定する和声解析タスクを行った。学習データは 4 声部すべての旋律と和音のペアとした。1 曲とそれ以外に分割する交差検定で評価を行ったところ、正解率は 75% であった。

### 4.4 ドミナント定型句の同定

文献<sup>11)</sup> の pp. 114-115 のバス課題において、ドミナント定型、終止定式の決定にあたり、階名を入力とし、定型句の始まりを B, 終了までを I, その他を O とラベル付けする方法を考えた (BIO ラベル法<sup>12)</sup>、図 5)。現在音、前後 2 音、前音+現在音の組み合わせ、



図 3 第 9 の主旋律への対旋律付け。音符を持たない弱拍部 (ダミーのラベルを付与) を入力に含めるか否かで、4 分音符の多い旋律 (1-4 小節)、8 分音符の多い旋律 (5-8 小節) とともに出力可能である。



図 4 第 9 の主旋律への和声付け。和音記号中の “b” は第 2 転回形をあらわす。得られる出力は和音系列のみなので、和音パートとバスパートは、音域が 1 オクターブにおさまるよう手与えた。

[http://homepages.inf.ed.ac.uk/s0450736/maxent\\_toolkit.html](http://homepages.inf.ed.ac.uk/s0450736/maxent_toolkit.html)  
<http://chasen.org/~taku/software/CRF++/>

非和声音の種類同定はこのタスクに含まれない。

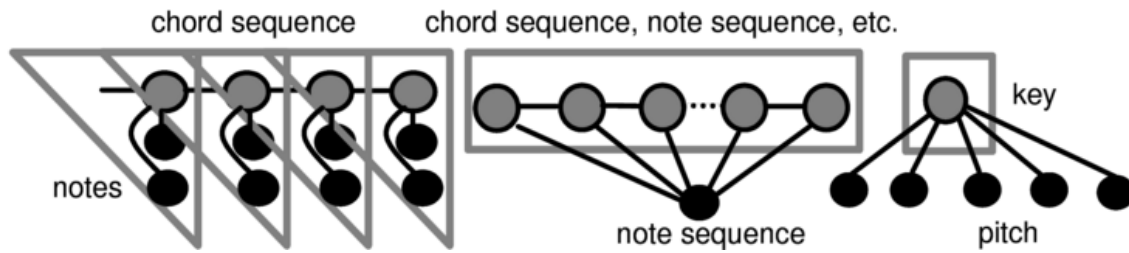


図 2 マルコフ確率場のグラフによる表現: 黒のノードは入力、グレイのノードは出力、グレイの線で囲まれた領域は最適解を表す。左図: 和声解析の概念図 (MEMM)。各声部の音高と前和音から最尤の和音を求める操作を繰返し適用していく。中図: 対旋律付け、和声付け等の概念図 (CRF)。入力系列そのものから、グローバルな最適解を求める。右図: 調認識の概念図。楽曲全体の音高頻度により調を決定する。



図 5 文献<sup>[11]</sup>の p. 113 のバス課題の例題。書の説明によると半終止 (第 4 小節) 決定後、終止定型 (6 小節 2 拍め～最後)、ドミナント定型 (1 小節 2 拍め～2 小節 1 拍め、2 小節 2 拍め～3 小節 1 拍め、5 小節 2 拍め～6 小節 1 拍め) の順に決定する。本研究では、終止定型箇所、ドミナント定型箇所の決定を、ドミナント定型同定タスクと呼ぶことにする。

表 1 実験の概観

|        | 入力形式    | 定量的評価     |
|--------|---------|-----------|
| 対旋律付け  | humdrum | 60%～70%程度 |
| 和声付け   | humdrum | 60%～70%程度 |
| D 定型同定 | 独自形式    | 99%       |
| 和声解析   | humdrum | 75%       |
| 調認識    | MIDI    | 82%       |

現在音+次音の組み合わせを素性とする CRF で学習した。1 曲とそれ以外に分割する交差検定で、総音符数 174 中、誤りは 1 箇所 (精度 99%) であった。

#### 4.5 調 認 識

MIDI の音高 (mod 12) の相対頻度を素性関数として調を決定する手法を用いた。これは、テキスト分類タスクの定式化<sup>[13]</sup>において単語頻度を音高頻度におきかえたものと同一である。すなわち、式 (2) において素性関数を 12 個用意し、値を楽曲全体から得られる相対頻度とした。学習には、24 の調を 4 曲ずつ含む 96 曲の Standard MIDI File (30 秒程度) を用いた。自身の調以外のあらゆる調に移調する (例えば、二長調の学習データがあった場合、曲全体の音高を 2 半音下げることにより、その曲が八長調であるとしたときの学習データが得られることになる) ことにより、学習データを増やした。すべての調が 1 セット中に含まれるよう 4 分割し交差検定したところ、誤りは 17 曲 (精度 82%) であった。

## 5. 考 察

### 5.1 本手法についての考察

従来から、最適系列を求める問題では、より遠くのコンテキストを考慮するため bigram, trigram の導入が主であったが、これらが素性関数の設計により容易である点で進歩しているといえる。さらに、様々な要素を素性関数として還元できるという長所もある。例えば、図 3、図 4 では平行 5 度等の禁則を犯しているが、階名のみを素性としているためと考えられ、旋律間の音程の導入で改善される可能性がある。また、データスパースネスに関しても、エントロピーを最大にすることによってスムージングの効果が自動的に達成されると考えられ、少量の学習データでの効果が期待できる。

一方で、同時の複数出力 (e.g. 和声解析における複数旋律の非和声音の種別の同定) のモデル化が困難と思われる。また、むやみに入力を複雑にすることも危険であり、例えば、入力が高次の情報を利用していると思われる作曲者推定、ジャンル同定のようなタスクには不向きであると思われる。

### 5.2 関連研究との比較

Pickens ら<sup>[14]</sup>は周辺の文脈として MIDI の音高を用い、現在の音高を推定することにより音楽情報検索に応用している。これに対し、我々は様々なタスクに適用可能であることを示した。

和声付け、対位法について、川上ら<sup>[15]</sup>、中瀬ら<sup>[16]</sup>による研究は、HMM, マルコフモデルに基づくため、基本的にとり同士の和音間/音符間の状態遷移のみをみる。よって和声常套句の導入、リズム単語の導入等、HMM の階層化によるスムージングを必要とした。これに対し本手法では、より遠くのコンテキストを取り入れることが容易であり、さらに、任意のシンボル (音符の種別 (4 分音符、8 分音符,...) 等) を素性関数

として設計することが可能である。ただし、HMM を階層的にすることと、CRF で素性関数を増設することは同一ではないため、単純には比較できない。

調認識については、本報告の手法の入力は Standard MIDI File であるのに対し、齊藤ら<sup>17)</sup>の手法は入力音が音響であるため、単純に比較はできない。しかし、和声などの知識は必要とせず、音高のヒストグラムなどの情報で同程度の精度を得られる点で本研究と比較できるレベルだと考える。

### 5.3 展 望

近年、隠れ CRF モデル<sup>18)</sup>が音響入力音声認識タスクに適用可能であることが示され、高精度を得られることが報告されている。これにより、単にシンボルの系列だけではなく、値を持つ系列でも扱えることがわかる。例えば、我々の研究グループにおいて、リズム認識<sup>19)</sup>ではスカラー値 (MIDI の inter-onset-interval) の時系列を、調認識<sup>17)</sup>では specmurt<sup>20)</sup>から加工されたベクトル値 (12 次元 chroma ベクトル) の時系列を扱っている。素性設計に対する柔軟性というマルコフ確率場モデルの利点を活かし、隠れ CRF モデルによる高性能化を期待できる。

4.4 節で述べた BIO ラベル法には様々な変種が提案されている<sup>21)</sup>が、特に、既存ラベルに Begin, Inside ラベルを追加する方法は音楽においても有用である。音楽大学の教科書では、カデンツを 3 種に分類できるとされている (e.g. K1:T-D-T)。それぞれのカデンツに対し B, I 付きラベルを用意する (e.g. K1 については K1-B, K1-I)。そして、B, I 付きのシンボルの最適系列を求めることにより、カデンツ認識も可能である。

### 5.4 統計的手法の限界についての考察

一般に、統計的な学習による手法には、規則を列挙するなどの無意味な作業から解放される、どのような入力に対しても解をだす、等の利点があるが、音楽的な文法的説明が欠如する、等の限界もある。よって、100 年以上前に確立した対位法、和声法はもちろんのこと、それらを計算機が扱いやすいよう知識化されたアイデア (e.g. Winograd による和声解析、generative theory of tonal music、以下 GTTM) を利用すべきであると考ええる。

音楽において、制約に基づく文法 (e.g. head-driven phrase structure grammar、以下 HPSG) による定式化もすでに提案されている<sup>22)</sup>。この定式化に、GTTM による構造解析、Winograd による和声解析等の知識を取り込みつつ、マルコフ確率場モデルのような統計モデルを取り込むことも視野にいたいと考える。実際、自然言語のパージングにおいては、マルコフ確率

場モデルに基づく統計的手法が最先端のひとつである<sup>23)</sup>一方で、制約に基づく文法に確率モデルを導入する手法<sup>24)</sup>もひとつの主流であり、近年ではマルコフ確率場モデルと HPSG の融合が提案されている<sup>25)</sup>。

## 6. ま と め

周辺の文脈に依存するラベル列のモデル化に適したマルコフ確率場モデルを音楽情報の解析に適用した。そして、シンボリックな音楽情報を入力とする種々のタスクを統一的な枠組みで扱えることを示した。具体的には、対旋律付け、和声付け、ドミナント定型の同定、和声解析、調認識への適用を示した。今後の展開としては、音程等、音楽のモデル化に有効な素性を吟味したい。また、ドミナント定型同定から構文解析 (カデンツ同定等の楽曲構造解析) へ拡張したい。

## 参 考 文 献

- 1) Krumhansl, Carol, *Cognitive Foundations of Musical Pitch*, Oxford University Press, 1990.
- 2) S. Z. Li, *Markov random field modeling in computer vision*, Springer Verlag, 1995.
- 3) A. L. Berger, S. A. Della Pietra, and V. J. Della Pietra, "A maximum entropy approach to natural language processing," *Computational Linguistics*, 1996.
- 4) Andrew McCallum, Dayne Freitag, Fernando Pereira, "Maximum Entropy Markov Models for Information Extraction and Segmentation," *ICML 2000*, 2000.
- 5) J. Lafferty, A. McCallum, and F. Pereira, "Conditional Random Fields: Probabilistic models for segmenting and labeling sequence data," *Proc. of ICML*, 2001.
- 6) Hanna M. Wallach, "Conditional Random Fields: An Introduction," Technical Report MS-CIS-04-21, Department of Computer and Information Science, University of Pennsylvania, 2004.
- 7) J. Darroch and D. Ratcliff, "Generalized iterative scaling for log-linear models," In *Ann. Math. Statistics*, 43, pp. 1470-1480, 1972.
- 8) Dong C. Liu and Jorge Nocedal, "On the limited memory BFGS method for large scale optimization," *Math. Programming* 45(3, (Ser. B)), pp.503-528, 1989.
- 9) Stanley F. Chen and Ronald. Rosenfeld. "A gaussian prior for smoothing maximum entropy models," Technical report, Carnegie Mellon University, 1999.
- 10) D. Huron, The Humdrum Toolkit: Software for Music Research, <http://dactyl.som.ohio->

state.edu/Humdrum/.

- 11) 島岡譲, “音楽の理論と実習 I,” 音楽之友社, 1982.
- 12) Lance A. Ramshaw and Mitchell P. Marcus, “Text chunking using transformation-based learning,” In *Proceedings of the Third ACL Workshop on Very Large Corpora*, 1995
- 13) Kamal Nigam, John Lafferty and Andrew McCallum, “Using Maximum Entropy for Text Classification,” In *IJCAI-99 Workshop on Machine Learning for Information Filtering*, pp. 61-67, 1999.
- 14) Jeremy Pickens and Costas Iliopoulos, “Markov random fields and maximum entropy modeling for music information retrieval,” *Proc. of IS-MIR 2005*, 2005.
- 15) 川上 隆, 中井 満, 下平 博, 嵯峨山 茂樹, “隠れマルコフモデルを用いた旋律への自動和声付け,” 情報処理学会研究報告 (MUS), 99-MUS-34, pp.59-66, 2000.
- 16) 中渕昌平, 西本卓也, 嵯峨山 茂樹, “動的計画法と音列出現確率を用いた対位法の対旋律の自動生成,” 情報処理学会研究報告 (MUS), 2004-MUS-56, pp. 65-70, 2004.
- 17) 齊藤 翔一郎, 武田 晴登, 西本 卓也, 嵯峨山 茂樹, “specmurt 分析と chroma vector を用いた HMM による音楽音響信号の調認識,” 情報処理学会研究報告 (MUS), 2005-MUS-61, pp. 85-90, 2005.
- 18) Asela Gunawardana, Milind Mahajan, Alex Acero, John C. Platt, “Hidden conditional random fields for phone classification,” *Proc. of Eurospeech 2005*, 2005.
- 19) 齋藤 直樹, 中井 満, 下平 博, 嵯峨山 茂樹, “隠れマルコフモデルによる音楽演奏からの音符列の推定,” 情報処理学会音楽情報科学研究会, 99-MUS-33, pp.27-32, Dec. 1999.
- 20) 高橋佳吾, 西本卓也, 嵯峨山 茂樹, “対数周波数逆畳み込みによる多重音の基本周波数解析,” 情報処理学会研究報告, 2003-MUS-53, pp. 61-66, 2003.
- 21) E. F. Tjong and K. Sang, “Text chunking by system combination,” In *Proceedings of Conference on Computational Natural Language Learning*, 2000.
- 22) Satoshi Tojo, “Harmony Analysis in Head Phase Structure Grammar,” *Seventh International Conference on Principles and Practice of Constraint Programming, Musical Constraints Workshop*, 2001.
- 23) E. Charniak, “A maximum-entropy-inspired parser,” In *Proc. NAACL-2000*, pages 132-139, 2000.
- 24) S. P. Abney, “Stochastic attribute-value grammars,” *Computational Linguistics*, 23(4), 1997.
- 25) Yusuke Miyao, Jun'ichi Tsujii, “Probabilistic disambiguation models for wide-coverage

HPSG parsing,” *Proceedings of the 43rd Annual Meeting of the ACL*, 2005.

## 付 録

### A.1 HMM (hidden Markov models) が CRF (conditional random fields) の特殊形であることの略証

CRF は、入力系列  $\mathbf{x}$  に対する各出力系列  $\mathbf{y}$  の条件付き確率を表現する識別モデルである。HMM では  $\mathbf{x}$  が観測系列に、 $\mathbf{y}$  が隠れ状態系列に対応することに注意されたい。入力系列を  $\mathbf{x} = w_1, \dots, w_{\#\mathbf{x}}$ , それに対応するラベル系列を  $\mathbf{y} = t_1, \dots, t_{\#\mathbf{x}}$  とすると、HMM におけるコストは以下のように書ける。ここで、 $\delta(t = t_i)$  は、 $t = t_i$  のときに 1 を、それ以外のときに 0 を返す関数である。

$$\begin{aligned} & \log\left(\prod_{i=1}^{\#\mathbf{x}} p(w_i|t_i)p(t_i|t_{i-1})\right) \\ &= \sum_{i=1}^{\#\mathbf{x}} \{\log p(w_i|t_i) + \log p(t_i|t_{i-1})\} \\ &= \sum_{i=1}^{\#\mathbf{x}} \left\{ \sum_{\langle w, t \rangle} \delta(w = w_i) \delta(t = t_i) \log p(w_i|t_i) \right. \\ & \quad \left. + \sum_{\langle t, t' \rangle} \delta(t = t_i) \delta(t' = t_{i-1}) \log p(t_i|t_{i-1}) \right\} \quad (3) \end{aligned}$$

ここで、 $\mathbf{F}(\mathbf{x}, \mathbf{y}) = \{F_1(\mathbf{x}, \mathbf{y}), \dots, F_K(\mathbf{x}, \mathbf{y})\}$  (ただし  $F_k(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{\#\mathbf{x}} f_k(\langle w_{i-1}, t_{i-1} \rangle, \langle w_i, t_i \rangle)$ ) と定義される素性関数ベクトル  $\mathbf{F}$  を導入する。各素性関数  $f$  は (2) のように定義される。すると、(3) は、

$$\begin{aligned} &= \sum_{w, t} F_{\langle w, t \rangle}(\mathbf{x}, \mathbf{y}) \cdot \lambda_{\langle w, t \rangle} + \sum_{t, t'} F_{\langle t, t' \rangle}(\mathbf{x}, \mathbf{y}) \cdot \lambda_{\langle t, t' \rangle} \\ &= \mathbf{\Lambda} \cdot \mathbf{F}(\mathbf{x}, \mathbf{y}) \end{aligned}$$

と書ける。ただし  $F_{\langle w, t \rangle}(\mathbf{x}, \mathbf{y})$ ,  $F_{\langle t, t' \rangle}(\mathbf{x}, \mathbf{y})$  は  $\mathbf{F}$  の 1 つの次元に相当し、

$$\begin{aligned} F_{\langle w, t \rangle}(\mathbf{x}, \mathbf{y}) &= \sum_i \delta(w = w_i) \delta(t = t_i) \\ F_{\langle t, t' \rangle}(\mathbf{x}, \mathbf{y}) &= \sum_i \delta(t = t_i) \delta(t' = t_{i-1}) \end{aligned}$$

である。すなわち、HMM における観測コスト  $\log p(w_i|t_i)$  状態遷移コスト  $\log p(t_i|t_{i-1})$  が CRF におけるパラメータ  $\lambda$  に対応し、観測と状態遷移のみを素性関数とするという点で HMM は CRF の特殊型であるといえる。