

【公報種別】特許法第 17 条の 2 の規定による補正の掲載

【部門区分】第 6 部門第 3 区分

【発行日】平成25年10月24日 (2013.10.24)

【公開番号】特開2011-60290(P2011-60290A)

【公開日】平成23年3月24日 (2011.3.24)

【年通号数】公開・登録公報2011-012

【出願番号】特願2010-202137(P2010-202137)

【国際特許分類】

G 0 6 N 3/00 (2006.01)

G 0 6 Q 50/04 (2012.01)

G 0 5 B 13/02 (2006.01)

G 0 6 N 5/04 (2006.01)

【 F I 】

G 0 6 N 3/00 5 5 0 E

G 0 6 F 17/60 1 0 6

G 0 5 B 13/02 L

G 0 6 N 5/04 5 8 0 J

【手続補正書】

【提出日】平成25年9月9日 (2013.9.9)

【手続補正 1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項 1】

技術システムの制御および / または調整をコンピュータ支援により学習する方法であって、

技術システムの運転が、運転中の技術システムの状態 (s) と、技術システムの運転中に実行され、技術システムのそれぞれの状態 (s) を連続状態に移行させる活動とによって特徴付けられる方法において、

・技術システムの運転中に求められた、状態 (s)、活動 (a) および連続状態 (s') を含むトレーニングデータに基づいて、品質関数 (Q) と活動選択ルール ((s)) を学習するステップ；

ただし前記品質関数 (Q) は技術システムの最適運転をモデル化し、

前記活動選択ルール ((s)) は、技術システムの運転中に当該技術システムのそれぞれの状態 (s) に対して実行すべき活動 (a) を指示し、

・品質関数 (Q) および活動選択ルール ((s)) の学習中に、品質関数 (Q) の統計的不確定性に対する尺度 (Q) を、不確定性伝播によって求めるステップ；

・該統計的不確定性に対する尺度 (Q) と、品質関数 (Q) への統計的に緩和された要求に相当する確率パラメータ () とに基づいて、変形された品質関数を決定するステップ；

ただし前記不確定性伝播は、非対角要素が無視された共分散マトリクスを使用し、

・変形された品質関数に基づいて、活動選択ルール ((s)) を学習するステップ；
を有する方法。

【請求項 2】

前記品質関数 (Q) を、評価 (R) および状態活動確率 (P) を考慮して学習し、

それぞれの評価 (R) は、状態 (s)、当該状態で実行された活動 (a)、および連続

状態 (s') からなる組合せの品質を、技術システムの最適運転の観点で評価し、

それぞれの状態活動確率 (P) は、状態と当該状態で実行された活動 (a) に依存して、連続状態 (s') の確率 (P) を指示する請求項 1 記載の方法。

【請求項 3】

品質関数 (Q) と活動選択ルール ((s)) を、ベルマン反復式に基づいて学習し、各反復ステップで新たな品質関数 (Q) と、該品質関数 (Q) の統計的不確定性に対する新たな尺度を求め、それにより新たに変形された品質関数を決定し、

それぞれの反復ステップで共分散マトリクスを、前記品質関数 (Q) 、状態活動確率 (P) および評価 (R) に依存し、非対角要素を無視して求める請求項 2 記載の方法。

【請求項 4】

ベルマン反復法の m 番目の反復ステップで、活動選択ルールを以下の活動 s_{\max} に基づいて求め、

【数 1】

$$\forall s: a_{s, \max} = \arg \max_a [Q^m(s, a) - \xi \sigma Q^m(s, a)]$$

ここで

【数 2】

$$Q^m(s, a) = \sum_{s'} P(s'|s, a) [R(s, a, s') + \gamma V^{m-1}(s')]$$

は品質関数であり、

【数 3】

$$Q^m(s, a) - \xi \sigma Q^m(s, a)$$

は変形された品質関数であり、

$Q^m(s, a)$ は m 番目の反復ステップにおける品質関数 (Q) の統計的不確定性に対する尺度 (Q) であり、

ここで

【数 4】

$$\begin{aligned} (\sigma Q^m(s, a))^2 &= \sum_{s'} (D_{QQ})^2 (\sigma V^{m-1}(s'))^2 + \\ &\quad \sum_{s'} (D_{QP})^2 (\sigma P(s'|s, a))^2 + \\ &\quad \sum_{s'} (D_{QR})^2 (\sigma R(s, a, s'))^2, \end{aligned}$$

$$(D_{QQ}) = \gamma P(s'|s, a),$$

$$(D_{QP}) = R(s, a, s') + \gamma V^{m-1}(s'),$$

$$(D_{QR}) = P(s'|s, a)$$

[0 , 1] は非連続因子であり、
は確率パラメータであり、

【数 5】

$$V^m(s) = \max_a [Q^m(s, a) - \xi \sigma Q^m(s, a)];$$

$$(\sigma V^m(s))^2 = (\sigma Q(s, a_{s, \max}))^2 ;$$

が成り立ち、

$P(s' | s, a)$ は、状態 s で活動 a が実行された際の連続状態 s' に対する状態活動確率であり、

$R(s, a, s')$ は、状態 s で活動 a が実行された際の連続状態 s' の評価であり、

$P(s' | s, a)$ は、状態 - 活動確率の統計的不確定性であり、

$R(s, a, s')$ は、評価の統計的不確定性である請求項 3 記載の方法。

【請求項 5】

状態活動確率 (P) を状態活動確率分布としてモデル化し、および / または評価 (R) を評価確率分布としてモデル化する請求項 2 から 4 までのいずれか一項記載の方法。

【請求項 6】

状態活動確率 (P) の統計的不確定性 (P) を、モデル化した状態活動確率分布から求め、評価の統計的不確定性 (R) を、モデル化した評価確率分布から求める請求項 4 および 5 記載の方法。

【請求項 7】

状態活動確率分布および / または評価確率分布を、トレーニングデータからの相対的頻度としてモデル化し、

ここで状態活動確率分布は多項分布としてモデル化し、および / または評価確率分布は正規分布としてモデル化する請求項 5 または 6 記載の方法。

【請求項 8】

状態活動確率分布を、アприオリ分布とアポステリオリパラメータを用いたベイズの推定に基づいてモデル化し、ここでアポステリオリパラメータはトレーニングデータに依存する請求項 5 から 7 までのいずれか一項記載の方法。

【請求項 9】

アприオリ分布は、ディリクレ分布および / または正規分布である請求項 8 記載の方法。

【請求項 10】

ディリクレ分布のパラメータ (θ_{ijk}) は、連続状態 (s') の平均数と、トレーニングデータによる状態 (s) の総数の商に相当する請求項 9 記載の方法。

【請求項 11】

学習すべき活動選択ルールは、決定論的活動選択ルールである請求項 1 から 10 までのいずれか一項記載の方法。

【請求項 12】

ベルマン反復式の m 番目の反復ステップにおける活動選択ルール $\pi^m(s)$ は以下のとおりであり、

【数 6】

$$\pi^m(s) = \arg \max_a Q^m(s, a) - \xi \sigma Q^m(s, a)$$

ここで

【数 7】

$$\pi^m(s)$$

は、選択された活動である、請求項 4 にかかる請求項 11 記載の方法。

【請求項 13】

学習すべき活動選択ルールは、技術システムの状態 (s) のために実行可能な活動 (a) に対する確率分布を指示する確率論的活動選択ルール ($P(s, a)$) である請求項 1 から 10 までのいずれか一項記載の方法。

【請求項 14】

前記ベルマン反復式の各反復ステップにおいて、実行可能な活動 (a) に対する新たな確率分布として確率分布を求め、

該確率分布は、最後の反復ステップの確率分布を、変形された品質関数の値を最大にする活動 (a) に比較的高い確率が割り当てられるよう変形する、請求項 3 にかかる請求項

1 3 記載の方法。

【請求項 1 5】

当該方法により、タービンの制御および / または調整が学習される請求項 1 から 1 4 までのいずれか一項記載の方法。

【請求項 1 6】

当該方法により、風力発電設備の制御および / または調整が学習される請求項 1 から 1 4 までのいずれか一項記載の方法。

【請求項 1 7】

技術システムの運転方法であって、
該技術システムが、請求項 1 から 1 6 までのいずれか 1 項記載の方法により学習された制御および / または調整に基づいて運転され、学習された活動選択ルールにより技術システムのそれぞれの状態 (s) で実行すべき活動 (a) が選択される運転方法。

【請求項 1 8】

技術システムの運転中に、請求項 1 から 1 6 までのいずれか 1 項記載の方法が反復され、
各反復の際に、技術システムが取る新たな状態 (s) および / または実行すべき活動 (a) がトレーニングデータとして考慮される請求項 1 7 記載の方法。

【請求項 1 9】

コンピュータに、
・技術システムの運転中に求められた、状態 (s)、活動 (a) および連続状態 (s ') を含むトレーニングデータに基づいて、品質関数 (Q) と活動選択ルール ((s)) を学習するステップ；
ただし前記品質関数 (Q) は技術システムの最適運転をモデル化し、
前記活動選択ルール ((s)) は、技術システムの運転中に当該技術システムのそれぞれの状態 (s) に対して実行すべき活動 (a) を指示し、
・品質関数 (Q) および活動選択ルール ((s)) の学習中に、品質関数 (Q) の統計的不確定性に対する尺度 (Q) を、不確定性伝播によって求めるステップ；
・該統計的不確定性に対する尺度 (Q) と、品質関数 (Q) への統計的に緩和された要求に相当する確率パラメータ () とに基づいて、変形された品質関数を決定するステップ；
ただし前記不確定性伝播は、非対角要素が無視された共分散マトリクスを使用し、
・変形された品質関数に基づいて、活動選択ルール ((s)) を学習するステップ；
を実行させるためのコンピュータプログラム。

【請求項 2 0】

コンピュータを備えた、技術システムの制御装置であって、
技術システムの運転が、運転中の技術システムの状態 (s) と、技術システムの運転中に実行され、技術システムのそれぞれの状態 (s) を連続状態に移行させる活動とによって特徴付けられる、制御装置において、
前記コンピュータは、
・技術システムの運転中に求められた、状態 (s)、活動 (a) および連続状態 (s ') を含むトレーニングデータに基づいて、品質関数 (Q) と活動選択ルール ((s)) を学習し、
ただし前記品質関数 (Q) は技術システムの最適運転をモデル化し、
前記活動選択ルール ((s)) は、技術システムの運転中に当該技術システムのそれぞれの状態 (s) に対して実行すべき活動 (a) を指示し、
・品質関数 (Q) および活動選択ルール ((s)) の学習中に、品質関数 (Q) の統計的不確定性に対する尺度 (Q) を、不確定性伝播によって求め、
・該統計的不確定性に対する尺度 (Q) と、品質関数 (Q) への統計的に緩和された要求に相当する確率パラメータ () とに基づいて、変形された品質関数を決定し、
ただし前記不確定性伝播は、非対角要素が無視された共分散マトリクスを使用し、

・変形された品質関数に基づいて、活動選択ルール（ (s) ）を学習することを特徴とする制御装置。

【請求項 21】

前記コンピュータは、
前記品質関数（ Q ）を、評価（ R ）および状態活動確率（ P ）を考慮して学習し、
それぞれの評価（ R ）は、状態（ s ）、当該状態で実行された活動（ a ）、および連続状態（ s' ）からなる組合せの品質を、技術システムの最適運転の観点で評価し、
それぞれの状態活動確率（ P ）は、状態と当該状態で実行された活動（ a ）に依存して、連続状態（ s' ）の確率（ P ）を指示する請求項 20 記載の制御装置。

【請求項 22】

前記コンピュータは、
品質関数（ Q ）と活動選択ルール（ (s) ）を、ベルマン反復式に基づいて学習し、
各反復ステップで新たな品質関数（ Q ）と、該品質関数（ Q ）の統計的不確定性に対する新たな尺度を求め、それにより新たに変形された品質関数を決定し、
それぞれの反復ステップで共分散マトリクスを、前記品質関数（ Q ）、状態活動確率（ P ）および評価（ R ）に依存し、非対角要素を無視して求める請求項 21 記載の制御装置。

【請求項 23】

前記コンピュータは、状態活動確率（ P ）を状態活動確率分布としてモデル化し、および / または評価（ R ）を評価確率分布としてモデル化する請求項 21 または 22 記載の制御装置。

【請求項 24】

前記コンピュータは、
状態活動確率（ P ）の統計的不確定性（ P ）を、モデル化した状態活動確率分布から求め、評価の統計的不確定性（ R ）を、モデル化した評価確率分布から求める請求項 23 記載の制御装置。

【請求項 25】

前記コンピュータは、
状態活動確率分布および / または評価確率分布を、トレーニングデータからの相対的頻度としてモデル化し、
ここで状態活動確率分布は多項分布としてモデル化し、および / または評価確率分布は正規分布としてモデル化する請求項 23 または 24 記載の制御装置。

【請求項 26】

前記コンピュータは、
状態活動確率分布を、アприオリ分布とアポステリオリパラメータを用いたベイズの推定に基づいてモデル化し、ここでアポステリオリパラメータはトレーニングデータに依存する請求項 23 から 25 までのいずれか 1 項記載の制御装置。

【請求項 27】

学習すべき活動選択ルールは、決定論的活動選択ルールである請求項 20 から 26 までのいずれか 1 項記載の制御装置。

【請求項 28】

学習すべき活動選択ルールは、技術システムの状態（ s ）のために実行可能な活動（ a ）に対する確率分布を指示する確率論的活動選択ルール（ (s) ）である請求項 20 から 26 までのいずれか 1 項記載の制御装置。

【請求項 29】

タービンまたは風力発電設備を制御する、
請求項 20 から 28 までのいずれか 1 項記載の制御装置。