

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第4572304号
(P4572304)

(45) 発行日 平成22年11月4日 (2010. 11. 4)

(24) 登録日 平成22年8月27日 (2010. 8. 27)

(51) Int. Cl. F I
G 0 6 N 3/12 (2006. 01) G O 6 N 3/12
C 1 2 N 15/09 (2006. 01) C 1 2 N 15/00 A

請求項の数 5 (全 18 頁)

(21) 出願番号	特願2004-18802 (P2004-18802)	(73) 特許権者	301022471
(22) 出願日	平成16年1月27日 (2004. 1. 27)		独立行政法人情報通信研究機構
(65) 公開番号	特開2005-215786 (P2005-215786A)		東京都小金井市貫井北町4-2-1
(43) 公開日	平成17年8月11日 (2005. 8. 11)	(73) 特許権者	393031586
審査請求日	平成19年1月4日 (2007. 1. 4)		株式会社国際電気通信基礎技術研究所
			京都府相楽郡精華町光台二丁目2番地2
		(74) 代理人	100064746
			弁理士 深見 久郎
		(74) 代理人	100085132
			弁理士 森田 俊雄
		(74) 代理人	100083703
			弁理士 仲村 義平
		(74) 代理人	100096781
			弁理士 堀井 豊

最終頁に続く

(54) 【発明の名称】 化学的遺伝プログラミング装置および化学的遺伝プログラミング方法

(57) 【特許請求の範囲】

【請求項 1】

各個体ごとの遺伝情報と、ルール番号とシンボルとの対を複数組合むトランスファー情報のグループと、前記シンボルの書換の対応関係を含む書換情報のグループと、前記トランスファー情報と前記書換情報との組み合わせであって翻訳処理時の前記ルール番号と前記書換情報との対応関係を表す翻訳対応関係を複数個含む翻訳テーブルとを格納するための記憶手段を備え、

前記書換情報は、前記シンボルのうちの始点シンボルと前記シンボルのうちの少なくとも1つの書換シンボルとの間の書換の対応関係を特定する情報であり、

前記遺伝情報は、

関数を表す表現型実体を生成するための前記ルール番号の第1の系列を含む第1の情報と、

前記トランスファー情報として選択的に使用される、前記ルール番号と前記シンボルとの対を複数組合む第2の情報と、

前記書換情報を生成するための前記ルール番号の第2の系列を含む第3の情報とを含み、

前記シンボルは、

前記表現型実体により表現される前記関数を構成するシンボルであって、翻訳処理の終了を示す終端シンボルと、

1つまたは複数の他のシンボルに継続して翻訳可能であることに対応する非終端シンボ

10

20

ルとを含み、

前記第2の情報からの転写処理により選択された前記ルール番号と前記シンボルとの対を使用して前記トランスファー情報のグループを生成する転写処理手段と、

前記翻訳テーブルにおいて、前記第3の情報中の前記ルール番号に一致するルール番号を有する前記翻訳対応関係をランダムに選択して、前記翻訳対応関係により特定される前記書換情報を前記書換情報のグループに追加することで前記書換情報のグループを生成する第1の翻訳処理手段と、

前記書換情報のグループからランダムに選択した書換情報の始点シンボルと、前記トランスファー情報のグループからランダムに選択したトランスファー情報のシンボルとが一致する場合に、当該トランスファー情報と当該書換情報との組み合わせを追加することで、前記翻訳テーブルを更新するための更新処理手段と、

前記第1の系列に従って、当該第1の系列中の前記ルール番号に対応する前記翻訳対応関係により、前記シンボルの書換を前記終端シンボルに到達するまで継続することで、到達した前記終端シンボルの組み合わせとして表現される前記表現型実体を生成するための表現型実体生成処理手段と、

前記表現型実体により表現される前記関数の所定の目標に対する適応度に応じて、前記個体の選択処理を行うための選択処理手段と、

前記遺伝情報に対する突然変異処理を行うための突然変異処理手段とを備える、化学的遺伝プログラミング装置。

【請求項2】

前記遺伝情報、前記トランスファー情報、前記書換情報の交叉処理を行うための交叉処理手段をさらに備える、請求項1記載の化学的遺伝プログラミング装置。

【請求項3】

前記書換情報は、

前記非終端シンボルのうちの1つから前記非終端シンボルのうちの少なくとも1つの他の非終端シンボルへの継続書換情報と、

前記非終端シンボルのうちの1つから前記終端シンボルのうちの1つへの終端書換情報とを含む、請求項1記載の化学的遺伝プログラミング装置。

【請求項4】

前記個体間で、前記トランスファー情報と、前記書換情報と、前記翻訳テーブルに含まれる情報とを交換する処理を行う交換処理手段をさらに備える、請求項1記載の化学的遺伝プログラミング装置。

【請求項5】

関数を表す表現型実体を生成するためのルール番号の第1の系列を含む第1の情報と、トランスファー情報として選択的に使用される、前記ルール番号とシンボルとの対を複数組合む第2の情報と、書換情報を生成するための前記ルール番号の第2の系列を含む第3の情報とを含む遺伝情報に基づいて、遺伝的プログラミングを行うための化学的遺伝プログラミング方法であって、

現在の世代の各個体ごとの前記遺伝情報と、前記ルール番号と前記シンボルとの対を複数組合むトランスファー情報のグループと、前記シンボルの書換の対応関係を含む前記書換情報のグループと、前記トランスファー情報と前記書換情報との組み合わせであって翻訳処理時の前記ルール番号と前記書換情報との対応関係を表す翻訳対応関係を複数個含む翻訳テーブルとを記憶装置の記憶領域に格納する第1ステップを備え、

前記書換情報は、前記シンボルのうちの始点シンボルと前記シンボルのうちの少なくとも1つの書換シンボルとの間の書換の対応関係を特定する情報であり、

前記シンボルは、

前記表現型実体により表現される前記関数を構成するシンボルであって、翻訳処理の終了を示す終端シンボルと、

1つまたは複数の他のシンボルに継続して翻訳可能であることに対応する非終端シンボルとを含み、

10

20

30

40

50

前記第2の情報からの転写処理により選択された前記ルール番号と前記シンボルとの対を使用して前記トランスファー情報のグループを生成して前記記憶装置に格納する第2ステップと、

前記翻訳テーブルにおいて、前記第3の情報中の前記ルール番号に一致するルール番号を有する前記翻訳対応関係をランダムに選択して、前記翻訳対応関係により特定される前記書換情報を前記書換情報のグループに追加することで前記書換情報のグループを生成して前記記憶装置に格納する第3ステップと、

前記書換情報のグループからランダムに選択した書換情報の始点シンボルと、前記トランスファー情報のグループからランダムに選択したトランスファー情報のシンボルとが一致する場合に、当該トランスファー情報と当該書換情報との組み合わせを追加することで、前記翻訳テーブルを更新する第4ステップと、

前記第1の系列に従って、当該第1の系列中の前記ルール番号に対応する前記翻訳対応関係により、前記シンボルの書換を前記終端シンボルに到達するまで継続することで、到達した前記終端シンボルの組み合わせとして表現される前記表現型実体を生成する第5ステップと、

前記表現型実体により表現される前記関数の所定の目標に対する適応度に応じて、前記個体の選択処理を行うための第6ステップと、

前記遺伝情報に対する突然変異処理を行い、前記遺伝情報の交叉処理を行うことで、前記遺伝情報の世代を更新する第7ステップと、

前記第1から第7ステップを所定回数繰り返すステップとを備える、化学的遺伝プログラミング方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、化学的遺伝アルゴリズムを用いた化学的遺伝プログラミング装置および化学的遺伝プログラミング方法に関する。

【背景技術】

【0002】

生物細胞において、シンボルから機能への翻訳、すなわち、遺伝子型から表現型への翻訳は、DNA (DeoxyriboNucleic Acid) によって決定されており、したがって、細胞が遺伝子の変更を受けると、このような翻訳プロセスも変更を受ける。DNAは、機能単位であるアミノ酸とトランスファーRNA (tRNA: transfer RiboNucleic Acid) とに翻訳される。翻訳されたアミノ酸とtRNAとは反応して、細胞の遺伝型の構造を決定するための翻訳テーブル、すなわち、アミノアシル-tRNAを生成する。遺伝型を決定するコードの構造に関する情報とコードの翻訳のための情報とが、ともにDNAに格納されているので、遺伝子の変異により、細胞は、その双方を進化させることができる。

【0003】

(化学的遺伝アルゴリズム)

このような現実の生物細胞中の機構を真似ることで、新たな遺伝アルゴリズムである「化学的遺伝アルゴリズム (CGA: Chemical Genetic Algorithms)」が提案されている (たとえば、非特許文献1および非特許文献2を参照)。

【0004】

CGAは、遺伝子型の情報と表現型の情報との間に動的な関係を持ち込み、遺伝子型から表現型への翻訳過程を自律的に最適化していくことができる。

【0005】

一般の遺伝的アルゴリズムにおいて、2進数で表わされるDNAと最適化したい関数のパラメータとの間の対応関係を定める「コード化」は、アルゴリズムの性能を左右する最も大事な要因である。「化学的遺伝アルゴリズム」では、この決定関係を自律的に行なわせる。

【0006】

10

20

30

40

50

このようなアルゴリズムは、最適化アルゴリズムとして使用することができる。最適化は、工学のさまざまな分野で最も要求度の高い処理の1つであり、化学的遺伝アルゴリズムは、生物類似にインテリジェント・システムを構築しようとするアプローチ方法の1つである。

【0007】

人工的に設計された進化システムの進化性は、一群の基本機能ユニットによって決定される。実生物における基本機能ユニットは20個のアミノ酸であり、DNAに書かれた遺伝情報は、上述した「アミノアシルtRNA」と呼ばれる一群の分子によってアミノ酸に変換される。生物システムにおいては、アミノアシルtRNAが、ゲノタイプ（遺伝子型：DNA上の塩基配列、遺伝コード）からフェノタイプ（表現型：たんぱく質の機能等）への対応関係を定義する基本的な分子である。

10

【0008】

実生物の機能ユニットであるアミノ酸や、それへの変換分子であるアミノアシルtRNAを選び取る処理は、正にDNA/RNAに書かれたゲノタイプ（コード）の進化と一緒に進められたのであって、その意味で、コードとその変換テーブルは「同時進化」もしくは広い意味での「共進化」をしたといえる。

【0009】

もしも、人工的な進化システムにおいても、コードと変換テーブルが共進化する仕組みを導入することができれば、進化実験をしながらシステムの進化性が自発的に高まるようにすることが可能となる。

20

【0010】

「化学的遺伝アルゴリズム」では、遺伝的アルゴリズムを生物代謝系との類推で発展させた手法である。

【0011】

通常の遺伝アルゴリズムでは、まず染色体集団が用意されて、それに突然変異・交叉などの遺伝的オペレーションが施されて処理が進む。しかしながら、「化学的遺伝アルゴリズム」（以下CGA）では、まず、染色体ではなく細胞の集団が用意される。

【0012】

細胞1個は、通常の染色体（DNA）に加えて、アミノ酸、tRNA、アミノアシルtRNAという3つの種類の小分子集団を含んでおり、DNA上の遺伝コードは、細胞内の転写・翻訳反応を通して実数値に変換され、適応度が計算される。

30

【0013】

遺伝的オペレーションとしては、通常のDNA突然変異、DNA交叉以外に、分子反応や細胞質交換も行なわれ、淘汰は、分子集団を含む細胞単位として行なわれる。

【0014】

（遺伝的プログラミング）

一方で、遺伝的プログラミングが、様々な工学的応用に適用されている（たとえば、非特許文献3を参照）。

【0015】

ランダムに生成された、所定のプログラミング構成要素からなるプログラムからスタートして、従来の遺伝的プログラミングは、遺伝的な操作を用いて各世代において最適なプログラムを、何世代にもわたって繁殖させていくことにより、その集団を進化させていく。

40

【非特許文献1】Suzuki, H., Sawai, H.: "Chemical genetic algorithms-Coevolution between codes and code translation." In: Standish, R.K., Bedau, M.A., Abbass, H. A. (eds.): Proceedings of the Eighth International Conference on Artificial Life (Artificial Life VIII)(2002) pp.162-172

【非特許文献2】Suzuki, H., Sawai, H., Piaseczny, W.: "Evolvability enhancement by the optimization of a chemical translation system-a case study." In: Dittrich, P., Kim, J.T.(eds.): The 7th European Conference on Artificial Life (ECAL) Wor

50

kshop Proceedings (2002) pp.146-155

【非特許文献3】Koza, R.: 1992. Genetic Programming : On the programming of Computers by Means of Natural Selection. London: MIT press (1992).

【発明の開示】

【発明が解決しようとする課題】

【0016】

しかしながら上述したようなCGAにおいては、関数パラメータの最適化の一手法については言及されているものの、これをプログラミングにどのように適用するのかについては明らかではないという問題点があった。

【0017】

本発明の目的は、化学的遺伝アルゴリズムを適用して、プログラミングを可能とする化学的遺伝プログラミング装置および化学的遺伝プログラミング方法を提供することである。

【課題を解決するための手段】

【0018】

この発明のある局面に従うと、各個体ごとの遺伝情報と、ルール番号とシンボルとの対を複数組合むトランスファー情報のグループと、シンボルの書換の対応関係を含む書換情報のグループと、トランスファー情報と書換情報との組み合わせであって翻訳処理時のルール番号と書換情報との対応関係を表す翻訳対応関係を複数個含む翻訳テーブルとを格納するための記憶手段を備え、書換情報は、シンボルのうちの始点シンボルとシンボルのうちの少なくとも1つの書換シンボルとの間の書換の対応関係を特定する情報であり、遺伝情報は、関数を表す表現型実体を生成するためのルール番号の第1の系列を含む第1の情報と、トランスファー情報として選択的に使用される、ルール番号とシンボルとの対を複数組合む第2の情報と、書換情報を生成するためのルール番号の第2の系列を含む第3の情報とを含み、シンボルは、表現型実体により表現される関数を構成するシンボルであって、翻訳処理の終了を示す終端シンボルと、1つまたは複数の他のシンボルに継続して翻訳可能であることに対応する非終端シンボルとを含み、第2の情報からの転写処理により選択されたルール番号とシンボルとの対を使用してトランスファー情報のグループを生成する転写処理手段と、翻訳テーブルにおいて、第3の情報中のルール番号に一致するルール番号を有する翻訳対応関係をランダムに選択して、翻訳対応関係により特定される書換情報を書換情報のグループに追加することで書換情報のグループを生成する第1の翻訳処理手段と、書換情報のグループからランダムに選択した書換情報の始点シンボルと、トランスファー情報のグループからランダムに選択したトランスファー情報のシンボルとが一致する場合に、当該トランスファー情報と当該書換情報との組み合わせを追加することで、翻訳テーブルを更新するための更新処理手段と、第1の系列に従って、当該第1の系列中のルール番号に対応する翻訳対応関係により、シンボルの書換を終端シンボルに到達するまで継続することで、到達した終端シンボルの組み合わせとして表現される表現型実体を生成するための表現型実体生成処理手段と、表現型実体により表現される関数の所定の目標に対する適応度に応じて、個体の選択処理を行うための選択処理手段と、遺伝情報に対する突然変異処理を行うための突然変異処理手段とを備える。

【0019】

好ましくは、化学的遺伝プログラミング装置は、遺伝情報、トランスファー情報、書換情報の交叉処理を行うための交叉処理手段をさらに備える。

【0020】

好ましくは、書換情報は、非終端シンボルのうちの1つから非終端シンボルのうちの少なくとも1つの他の非終端シンボルへの継続書換情報と、非終端シンボルのうちの1つから終端シンボルのうちの1つへの終端書換情報とを含む。

【0021】

好ましくは、化学的遺伝プログラミング装置は、個体間で、トランスファー情報と、書換情報と、翻訳テーブルに含まれる情報とを交換する処理を行う交換処理手段をさらに備

10

20

30

40

50

える。

【0022】

この発明の他の局面に従うと、関数を表す表現型実体を生成するためのルール番号の第1の系列を含む第1の情報と、トランスファー情報として選択的に使用される、ルール番号とシンボルとの対を複数含む第2の情報と、書換情報を生成するためのルール番号の第2の系列を含む第3の情報とを含む遺伝情報に基づいて、遺伝的プログラミングを行うための化学的遺伝プログラミング方法であって、現在の世代の各個体ごとの遺伝情報と、ルール番号とシンボルとの対を複数含むトランスファー情報のグループと、シンボルの書換の対応関係を含む書換情報のグループと、トランスファー情報と書換情報との組み合わせであって翻訳処理時のルール番号と書換情報との対応関係を表す翻訳対応関係を複数個含む翻訳テーブルとを記憶装置の記憶領域に格納する第1ステップを備え、書換情報は、シンボルのうちの始点シンボルとシンボルのうちの少なくとも1つの書換シンボルとの間の書換の対応関係を特定する情報であり、シンボルは、表現型実体により表現される関数を構成するシンボルであって、翻訳処理の終了を示す終端シンボルと、1つまたは複数の他のシンボルに継続して翻訳可能であることに対応する非終端シンボルとを含み、第2の情報からの転写処理により選択されたルール番号とシンボルとの対を使用してトランスファー情報のグループを生成して記憶装置に格納する第2ステップと、翻訳テーブルにおいて、第3の情報中のルール番号に一致するルール番号を有する翻訳対応関係をランダムに選択して、翻訳対応関係により特定される書換情報を書換情報のグループに追加することで書換情報のグループを生成して記憶装置に格納する第3ステップと、書換情報のグループからランダムに選択した書換情報の始点シンボルと、トランスファー情報のグループからランダムに選択したトランスファー情報のシンボルとが一致する場合に、当該トランスファー情報と当該書換情報との組み合わせを追加することで、翻訳テーブルを更新する第4ステップと、第1の系列に従って、当該第1の系列中のルール番号に対応する翻訳対応関係により、シンボルの書換を終端シンボルに到達するまで継続することで、到達した終端シンボルの組み合わせとして表現される表現型実体を生成する第5ステップと、表現型実体により表現される関数の所定の目標に対する適応度に応じて、個体の選択処理を行うための第6ステップと、遺伝情報に対する突然変異処理を行い、遺伝情報の交叉処理を行うことで、遺伝情報の世代を更新する第7ステップと、第1から第7ステップを所定回数繰り返すステップとを備える。

【発明の効果】

【0023】

この発明にかかる化学的遺伝プログラミング装置および化学的遺伝プログラミング方法によれば、遺伝情報の翻訳テーブルを持ち、遺伝情報から作られたトランスファー情報および書換情報が新しい翻訳テーブルを生成するというフィードバック機構を備え、「コードと翻訳テーブルの共進化」を実現して、最適なプログラミング結果を短時間に得ることが可能である。

【発明を実施するための最良の形態】

【0024】

以下、図面を参照して本発明の実施の形態について説明する。

【0025】

(本発明のシステム構成)

図1は、本発明の化学的遺伝プログラミング装置100の一例の外観を示す図である。

【0026】

すなわち、図1に示した例では、化学的遺伝プログラミング装置100は、一例として、汎用のコンピュータにより実現されるものとしている。

【0027】

図1を参照して、この化学的遺伝プログラミング装置(コンピュータ)100は、CD-ROM(Compact Disc Read-Only Memory)上の情報を読込むためのCD-ROMドライブ108およびフレキシブルディスク(Flexible Disk、以下FD)116に情報を読

み書きするためのFDドライブ106を備えたコンピュータ本体102と、コンピュータ本体102に接続された表示装置としてのディスプレイ104と、同じくコンピュータ本体102に接続された入力装置としてのキーボード110およびマウス112を含む。

【0028】

図2は、この化学的遺伝プログラミング装置100のハードウェア構成をブロック図形式で示す図である。

【0029】

図2に示されるように、この化学的遺伝プログラミング装置100を構成するコンピュータ本体102は、CD-ROMドライブ108およびFDドライブ106に加えて、それぞれバスBSに接続されたCPU (Central Processing Unit) 120と、ROM (Read Only Memory) およびRAM (Random Access Memory) を含むメモリ122と、直接アクセスメモリ装置、たとえば、ハードディスク124と、ネットワーク200とデータの授受を行うための通信インタフェース128とを含んでいる。CD-ROMドライブ108にはCD-ROM 118が装着される。FDドライブ106にはFD 116が装着される。

【0030】

なお、CD-ROM 118は、コンピュータ本体に対してインストールされるプログラム等の情報を記録可能な媒体であれば、他の媒体、たとえば、DVD-ROM (Digital Versatile Disc) やメモリカードなどでもよく、その場合は、コンピュータ本体102には、これらの媒体を読取ることが可能なドライブ装置が設けられる。

【0031】

本発明の化学的遺伝プログラミング装置の主要部は、コンピュータハードウェアと、CPU 120により実行されるソフトウェアとにより構成される。一般的にこうしたソフトウェアはCD-ROM 118、FD 116等の記憶媒体に格納されて流通し、CD-ROMドライブ108またはFDドライブ106等により記憶媒体から読取られてハードディスク124に一旦格納される。または、当該装置がネットワークに接続されている場合には、ネットワーク上のサーバから一旦ハードディスク124にコピーされてもよい。そうしてさらにハードディスク124からメモリ122中のRAMに読出されてCPU 120により実行される。なお、ネットワーク接続されている場合には、ハードディスク124に格納することなくRAMに直接ロードして実行するようにしてもよい。

【0032】

図1および図2に示したコンピュータのハードウェア自体およびその動作原理は一般的なものである。したがって、本発明の最も本質的な部分は、FD 116、CD-ROM 118、ハードディスク124等の記憶媒体に記憶されたソフトウェアである。

【0033】

なお、一般的傾向として、コンピュータのオペレーティングシステムの一部として様々なプログラムモジュールを用意しておき、アプリケーションプログラムはこれらモジュールを所定の配列で必要な時に呼び出して処理を進める方式が一般的である。そうした場合、当該化学的遺伝プログラミング装置100を実現するためのソフトウェア自体にはそうしたモジュールは含まれず、当該コンピュータでオペレーティングシステムと協働してはじめて化学的遺伝プログラミング装置100が実現することになる。しかし、一般的なプラットフォームを使用する限り、そうしたモジュールを含ませたソフトウェアを流通させる必要はなく、それらモジュールを含まないソフトウェア自体およびそれらソフトウェアを記録した記録媒体（およびそれらソフトウェアがネットワーク上を流通する場合のデータ信号）が実施の形態を構成すると考えることができる。

【0034】

(生物システムにおける転写・翻訳機構)

まず、本発明の化学的遺伝プログラミング装置の動作を説明する前提として、実際の生物細胞における代謝の機構について概説しておく。

【0035】

図3は、このような実際の生物細胞における代謝の機構を説明するための概念図である。

【0036】

図3を参照して、生物細胞における遺伝情報の翻訳は、一群の分子による代謝反応によって行なわれる。まず、生物は生合成反応により無機物からアミノ酸を作り出す。その主役となるのは植物が行なう光合成反応であり、この反応が作り出す糖を出発点とした解糖系の副産物としてのアミノ酸は合成される。また動物は、植物が合成したアミノ酸を摂取・消化することによってアミノ酸を獲得する。

【0037】

一方、DNAが持つ遺伝情報は、大きく2つの流れをたどる。

10

【0038】

まず1つは、DNAからポリメラーゼによる転写によって作られたtRNAがアミノ酸とカップリングしてアミノアシルtRNAを作る。tRNAは、約80個の塩基からなるポリヌクレオチドであり、細胞内では特殊なクローバ構造に折畳まれる。

【0039】

その一方の端にはmRNAと相補的にマッチングするアンチコドンが、もう一方の端はアミノ酸とマッチングする認識部位が含まれており、tRNAはこの特殊な構造から、コドンからアミノ酸へのアダプタとして働くことができる。

【0040】

このtRNAをアミノ酸とカップリングする役目を果たするのがアミノアシルtRNA合成酵素（アミノアシルtRNAシンセターゼ）と呼ばれる一群の酵素である。アミノアシルtRNA合成酵素は、tRNAクローバ構造に含まれるアミノ酸認識部位に取付き、それと照合するアミノ酸を見つけ出してはそれらを結合させる。各アミノ酸には1つずつ別のアミノアシルtRNA合成酵素が用意されており、したがって細胞内では20種類のアミノアシルtRNA合成酵素が使われている。

20

【0041】

DNAの遺伝情報がたどるもう1つの流れは、DNAからポリメラーゼによる転写によって作られたmRNAの翻訳である。mRNAは、コドンの配列情報を保持した情報高分子であり、タンパク質のアミノ酸配列のもとになる。巨大タンパク質構造であるリボソームは、このmRNAの中に含まれるコドンを1つ1つ順番に読取っては、対合するアンチコドンを有するアミノアシルtRNAを見つけ出し、そこにカップリングしているアミノ酸を鎖状に結合していく。こうして得られたポリペプチド鎖（タンパク質）が、また酵素として細胞のさまざまな代謝反応を触媒する。

30

【0042】

以上述べた生物遺伝情報の翻訳機構において、生物においては、ゲノタイプからフェノタイプへの対応関係（翻訳テーブル）が可変になっているということが特徴的である。生物では、コドン（ゲノタイプ・ユニット）とアミノ酸（フェノタイプ・ユニット）が直接結合した分子であるアミノアシルtRNAが翻訳テーブルを表わす。アミノアシルtRNAの生成は、DNAを転写したtRNAがもとになるから、生物における翻訳テーブルは、DNA情報の改変によって変化し得る仕組みを持つ。

40

【0043】

加えて、生物の翻訳機構が持つもう1つの重要な特徴は、代謝マップが一方向性のものではなく、フィードバックを含んでいるという点である。

【0044】

すなわち、翻訳を含めた細胞内のすべての代謝反応は、酵素によって触媒されているが、それら酵素自体は代謝反応の産物であり、細胞内の各分子は複雑に相互依存している。特に、遺伝情報の翻訳については、アミノアシルtRNAを合成する酵素であるアミノアシルtRNA合成酵素は、たんぱくとしてアミノアシルtRNAをコードしたコードの翻訳によってリボソームで合成されるから、コード（アミノアシルtRNA合成酵素を表現するコドンシーケンス）と翻訳テーブル（アミノアシルtRNA）が相互に依存しており

50

、共進化する仕組みになっている。

【0045】

(化学的遺伝プログラミング処理の概要)

以下では、このような実際の生物の代謝反応をモデル化した本発明にかかる「化学的遺伝プログラミング」の構成の概略について説明する。

【0046】

(細胞内代謝反応)

図4は、化学的遺伝プログラミング処理の基本的な流れの概略を示す概念図である。

【0047】

図4を参照して、まず、各々が次の4つの分子を内部に持つ複数の人工細胞(個体)を考える。すなわち、1個のDNA、 L_1 個のtRNA、 M_1 個のアミノ酸、 R_1 個のアミノアシルtRNAである。

10

【0048】

これら人工の分子は、次のような代謝反応を毎世代行なう。

【0049】

各DNAは、第1の部分、第2の部分、第3の部分を含んでいるものとする。

【0050】

DNAの第1の部分は、後に説明するように化学的遺伝プログラミングにより生成される「表現型ツリー」の翻訳のための情報(ルール番号)を有する。

【0051】

20

DNAの第2の部分は、tRNAテーブルを転写により生成するための情報を含んでいる。この第2の部分は、ルール番号とシンボルとの対を複数組合む。ルール番号とシンボルとの対は、実際の生物におけるtRNAが、その一方の端にはmRNAと相補的にマッチングするアンチコドン有し、もう一方の端はアミノ酸とマッチングする認識部位を有することをモデル化したことに相当する。ただし、以下の説明で明らかになるように、本発明にかかる化学的遺伝プログラミング処理では、mRNAに相当する構成は含まれないため、tRNAテーブルのルール番号は、後に説明するアミノアシルtRNAテーブルを介してDNAの第1の部分のルール番号と対応することになる。ここでは、同一のルール番号同士が対応する。

【0052】

30

ここで、「シンボル」とは、表現型ツリーを構成するアルファベット文字、数学記号、数字等である。シンボルとしては、必要に応じて、他の記号を含ませてもよい。また、シンボルには、終端シンボルと非終端シンボルとが含まれるものとする。プログラミングにより最終的に生成するのが、たとえば、ある関数形である場合を考えると、この最終的な関数形は、変数を表現するためのアルファベット文字(たとえば、 x)と数学記号(演算記号、たとえば、 $+$ 、 $-$ など)により表現される。したがって、このようなプログラミングの最終形に含まれるシンボルは、終端シンボルであり、それ以外のシンボルは非終端シンボルである。

【0053】

DNAの第3の部分は、アミノ酸を翻訳により生成するための情報を含んでいる。

40

【0054】

個々の「アミノ酸」は、上記シンボルの書換の対応関係に対応する。ここで、この書換の対応関係を「 $A \rightarrow B$ 」のように表し、この対応関係において「A」を始点シンボル、「B」を書換シンボルと呼ぶことにする。

【0055】

ここで、「転写」とは、DNA中の情報をそのままコピーする処理に相当し、「翻訳」とは、DNA中の情報(ルール番号)を、後に説明するアミノアシルtRNAテーブルにしたがって、対応するシンボルの書換を行っていく処理を意味する。

【0056】

すなわち、所定の非終端シンボルから、DNAの第1の部分の情報に従って、書換を順

50

次行っていく。この書換は、途中で生成される分岐のためにツリー構造となる。このようなツリーを「表現型ツリー」とよぶ。そして、この表現型ツリーの全ての枝の端部が、終端シンボルとなった時点で、そのDNAの第1の部分についての翻訳処理が終了する。このようにして翻訳処理が終了することで、生成された表現型ツリーの枝の端部により表現される結果を、「表現型実体」と呼ぶ。

【0057】

以下、図4で示される原則的な処理について、簡単に説明する。

【0058】

まず、DNAの第2の部分のL個のルール番号—シンボルのペアから、L個のtRNAが転写によって作られ、tRNAテーブルの集団に加えられる。ここで、tRNAの数が、 L_1 個をオーバーしたら、tRNAをランダムに選んで消去する。

10

【0059】

次に、M個のアミノ酸をDNAの第3の部分のM個のルール番号から翻訳によって作る。翻訳は、存在しているアミノアシルtRNAの集団（アミノアシルtRNAテーブル）を用いて行なわれ、DNA上のルール番号に一致するルール番号を持つアミノアシルtRNAがランダムに選ばれる。それを持つアミノ酸がアミノ酸集団（アミノ酸テーブル）に加えられる。

【0060】

このとき同様に、アミノ酸数が M_1 個をオーバーしたら、アミノ酸をランダムに選んで消す。

20

【0061】

次に、tRNAとアミノ酸との間の分子反応が起きる。ランダムに選んだtRNAのシンボルとランダムに選んだアミノ酸の始点シンボルが一致していたなら、2つを結合して1つのアミノアシルtRNAを作る。これを繰返して、 R_2 個の新しいアミノアシルtRNAが、アミノアシルtRNA集団に加えられたら、アミノアシルtRNAをランダムに選んで消し、アミノアシルtRNA数が R_1 個以内になるようにする。この結果、図4に示すように、同一のルール番号に対して、異なる書換の対応関係（アミノ酸）が対応することも許容するものとする。

【0062】

次に、DNAの第1の部分のルール番号から、アミノアシルtRNAを用いた翻訳により、表現型ツリーを生成する。

30

【0063】

表現型ツリーの最終形により表現されるプログラミング結果、たとえば、関数形を外部に用意したターゲット関数と比較し、細胞の適応度を計算する。後に説明するように、この適応度に応じて、細胞の選択が行われる。さらに、遺伝的プログラミングのために、DNAには、突然変異処理と、DNAの交叉処理、細胞間の分子交換処理などが行われて、1世代分の処理が終了する。

【0064】

以上の反応は、もちろん生物の翻訳機構を完全に模倣したものではない。DNAは、mRNAの完全な情報を持っているわけではないし、したがって酵素として働くタンパク質はこのモデルでは作られない。しかしながら、この簡略化された代謝系は、アミノアシルtRNAが遺伝情報の翻訳テーブルを持ち、翻訳によって作られた分子が新しいアミノアシルtRNAを生成するというフィードバック機構を備えており、生物が行なう「コードと翻訳テーブルの共進化」を実現している。

40

【0065】

（化学的遺伝プログラミング装置の機能構成）

図5は、図1および図2で説明した化学的遺伝プログラミング装置100の機能ブロック図である。

【0066】

図4で説明したような処理を実現するために、ハードディスク124中には、各細胞ご

50

とのDNA情報を格納するためのDNA情報テーブル1242と、各細胞において転写により生成されたtRNAテーブルを格納するためのtRNA情報テーブル1244と、各細胞において翻訳により生成されたアミノ酸テーブルを格納するためのアミノ酸情報テーブル1246と、各細胞において反応により生成されたアミノアシルtRNAテーブルを格納するアミノアシルtRNA情報テーブル1248とが、記憶されている。

【0067】

また、CPU120が、ハードディスク124等に格納されたプログラムに基づいて行う処理として、演算処理部は、DNAの第2の部分からtRNAの転写処理を行うtRNA転写処理部1201と、DNAの第3の部分とアミノアシルtRNAテーブルとに基づいて、アミノ酸の翻訳処理を行うアミノ酸翻訳処理部1202と、アミノ酸とtRNAとの反応処理を行ってアミノアシルtRNAテーブルを更新するためのアミノアシルtRNA更新処理部1203と、アミノアシルtRNAテーブルとDNAの第1の部分とに基づく翻訳により表現型ツリーを生成するための表現型ツリー生成処理部1205と、適応度に応じて細胞の選択処理を行うための選択処理部1205と、DNAに対する突然変異処理を行うための突然変異処理部1206と、DNAの交叉および分子交換処理を行うための交叉・分子交換処理部1207とを含む。

10

【0068】

(遺伝的オペレーション)

以下、具体的な化学的遺伝プログラミングの処理について説明する。

【0069】

図6および図7は、このような化学的遺伝プログラミングの処理の流れを説明するためのフローチャートである。

20

【0070】

まず、図6を参照して、処理が開始されると、初期化処理が行われる(ステップS100)。

【0071】

ここで、初期化処理では、代謝機構を持った細胞を表現するため変数およびテーブルの記憶領域を、たとえば、ハードディスク124のような記憶装置中に細胞N個分用意する。

【0072】

初め、各細胞のDNAの第1および第3の部分にはランダムなルール番号を与え、DNAの第2の部分には、ランダムなルール番号および所定のシンボルのうちのランダムな非終端シンボルを対として与え、tRNAテーブル、アミノアシルtRNAは、空の状態であるとする。世代を表す変数GNの値が1に設定される。また、各細胞にアミノ酸をM₁個用意する。

30

【0073】

図8は、このように初期的に準備されるアミノ酸の例を示す図である。

【0074】

図8に示した例では、シンボルは6つの非終端シンボルの”I, N, O, R, Z, D”と、16個の終端シンボルの”x, ^, +, -, *, /, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9”を含んでおり、アミノ酸は、これらシンボル間の書換の対応関係を示す。ここでは、全ての翻訳は、シンボル”I”から開始され、全ての非終端シンボルが終端シンボルに置き換えられたときに、翻訳が終了する。すなわち、現在存在している表現のうちの非終端シンボルに対して、DNA中に現われている左から右の順序に従って翻訳(書換)が行われ、すべての非終端シンボルが置き換えられてしまうまでこの処理が続けられる。非終端シンボルは、さらに翻訳可能なシンボルのことを意味する。

40

【0075】

また、たとえば、”N→(NON)”というような対応関係は、ここで、1つのシンボル”N”から3つのシンボル”N”、”O”、”N”への分岐が生じることを意味する。図8に示したアミノ酸のライブラリは、あくまでも一例であって、たとえば、さらに、多

50

くのシンボルや他の対応関係が含まれていてもよいし、翻訳が開始されるシンボルも複数個存在してもよい。

【0076】

なお、DNAの第2の部分中にランダムに割当てられるシンボルは、この初期的なアミノ酸のライブラリの非終端シンボルに対応するようになっているものとする。また、第1および第3の部分のランダムなルール番号は、この第2の部分のルール番号の範囲で割当てが行われているものとする。

【0077】

再び、図6に戻って、続いて、各細胞において、 L_1 個のtRNAをDNAの第2部分からの転写により生成する（ステップS102）。

10

【0078】

さらに、ランダムに選んだtRNAのシンボルとランダムに選んだアミノ酸の始点シンボルが一致していたなら、2つを結合して1つのアミノアシルtRNAを作るとの処理を繰返すことで、 R_1 個のアミノアシルtRNAからなるテーブルを生成する（ステップS104）。

【0079】

次に、生成されたアミノアシルtRNAテーブルを用いて、遺伝子の第1の部分の翻訳して、表現型ツリーを生成する（ステップS106）。

【0080】

図9は、このようにして生成される表現型ツリーの一例を示す図である。

20

【0081】

図9に示した例では、初期シンボル" I "から始まって、途中で3つの分岐が生じ、各分岐がそれぞれ、" 4 "、" + "、" x "という終端シンボルに翻訳（書換）が行われることで、翻訳処理が終了している。この場合は、この表現型ツリーにより、" 4 + x "との関数形（表現型実体）が生成されたことになる。より一般的には、このような最終形は、多変数の関数であってもよい。

【0082】

再び、図6もどって、次に、表現型ツリーにより生成された関数とターゲット関数との比較に基づいて、細胞の選択処理を行う（ステップS108）。

【0083】

30

すなわち、表現型ツリーにより表される関数とターゲット関数とを比較して一致度（適応度）を算出する。たとえば、所定の変数の変域において、複数の所定のサンプル点での表現型ツリーにより表される関数の値とターゲット関数の値の差の総和の平方に依存して、この平方に対して単調に減少する関数を評価関数とし、この評価関数の値を一致度とすればよい。つまり、一致の度合いが低くなるほど、一致度は小さくなる。

【0084】

次に、この一致度に応じた確率で選択されるような選択規則、たとえば、ルーレット選択により細胞の選択を行う。したがって、次の世代の細胞の個体数は、このようにして選択されたものとなる。

【0085】

40

次に、DNAの突然変異処理が行われる（ステップS110）。一旦、細胞が次の世代に対して選択されると、各細胞のDNA中のルール番号の各々は、ランダムに確率 p_m により、他の正当なコードと置き換えられる。

【0086】

続いて、DNAの交叉処理と分子交換処理が行われる（ステップS112）。

【0087】

すなわち、突然変異が終わった後に、選択された細胞の個数をNとするとき、すべての細胞を $N/2$ 組のペアにする。各ペアは、確率 p_c で、交叉処理を受ける。つまり、確率 p_c で、ランダムに選択されたDNAの対の間で、一点における交叉が行なわれる。

【0088】

50

さらに、それぞれのペアにおいて、分子交換を行う。たとえば、分子交換では、tRNA、アミノ酸、アミノアシルtRNAの各々について、50%の分子をランダムに選んで交換するものとする。

【0089】

このような進化の過程によって、最適なアミノ酸が、世代間を通して伝搬していくが、適切でないアミノ酸は除去されていく。アミノ酸およびアミノアシルtRNAの複雑さは、時間とともに大きくなっていき、アミノ酸の生産され得る組合せは、1つのアミノアシルtRNAテーブルに組み込まれていくことになる。

【0090】

次に、世代変数GNの値が1だけインクリメントされる（ステップS114）。 10

【0091】

そして、世代変数GNが所定の最大世代数Gmaxを超えているかが判断される（ステップS116）。もしも、世代変数GNが所定の最大世代数Gmaxを超えていれば、処理は終了する。一方、世代変数GNが所定の最大世代数Gmax以下であるときは、処理はステップS118に移行する。

【0092】

続いて、DNAの第2の部分のL個のルール番号—シンボルのペアから、L個のtRNAが転写によって作られ、tRNAテーブルの集団に加えられる（ステップS118）。ここで、tRNAの数が、L₁個をオーバーしたら、tRNAをランダムに選んで消去する。 20

【0093】

次に、M個のアミノ酸をDNAの第3の部分のM個のルール番号から翻訳によって生成する（ステップS120）。翻訳は、存在しているアミノアシルtRNAの集団（アミノアシルtRNAテーブル）を用いて行なわれ、DNA上のルール番号に一致するルール番号を持つアミノアシルtRNAがランダムに選ばれる。それを持つアミノ酸がアミノ酸集団（アミノ酸テーブル）に加えられる。

【0094】

このとき同様に、アミノ酸数がM₁個をオーバーしたら、アミノ酸をランダムに選んで消す。

【0095】

次に、tRNAとアミノ酸との間の分子反応処理により、アミノアシルtRNAテーブルの更新を行う（ステップS122）。すなわち、ランダムに選んだtRNAのシンボルとランダムに選んだアミノ酸の始点シンボルが一致していたなら、2つを結合して1つのアミノアシルtRNAを作る。これを繰返して、R₂個の新しいアミノアシルtRNAが、アミノアシルtRNA集団に加えられたら、アミノアシルtRNAをランダムに選んで消し、アミノアシルtRNA数がR₁個以内になるようにする。この結果、同一のルール番号に対して、異なる書換の対応関係（アミノ酸）が対応することも許容する。 30

【0096】

次に、DNAの第1の部分のルール番号から、アミノアシルtRNAを用いた翻訳により、表現型ツリーを生成する（ステップS124）。 40

【0097】

さらに、ステップS108、S110、S112とそれぞれ同様にして、細胞の選択処理を行い（ステップS126）、DNAの突然変異処理（ステップS128）、DNAの交叉と分子交換処理（ステップS130）を行って、処理はステップS114に復帰する。

【0098】

（具体例）

化学的遺伝プログラミング処理の有用性を検証するために、単一変数の数学的な関数形をこの手続きで獲得する処理を行った。特に、ターゲット関数としては、以下のものとする。 50

【0099】

$$f(x) = 2x^6 + 3x^4 + 4x^2 + 100$$

このような関数を、領域（-5, 5）において最適化を行なう。各個体の一致度（適応度）は、以下の関数によって定義される。

【0100】

【数1】

$$\text{一致度: } f = \exp(6 \times \exp(-d/25000000))$$

【0101】

ここで、 d は、上記領域において所定の間隔をもった有限個の値の組について、生成された関数の各値と、ターゲット関数の各値との間の差の和の平方である。これらの実験のためのパラメータの値は、以下のとおりである。

10

【0102】

突然変異確率 p_m は、0.007であり、交叉確率 p_c は0.7である。

【0103】

個体の数は143とした。

【0104】

図10は、以上の条件において、最適個体と平均個体の一致度の各世代ごとの変化を表す図である。

【0105】

この条件では、最適な個体は、第58世代において現われていることがわかった。

20

【0106】

また、図11は、最終的に得られた関数とターゲット関数との比較結果を示す図である。さらに、図12は、細胞の中に進化したアミノ酸のヒストグラムを示す図である。

【図面の簡単な説明】

【0107】

【図1】本発明の化学的遺伝プログラミング装置100の一例の外観を示す図である。

【図2】化学的遺伝プログラミング装置100のハードウェア構成をブロック図形式で示す図である。

【図3】実際の生物細胞における代謝の機構を説明するための概念図である。

30

【図4】化学的遺伝プログラミング処理の基本的な流れの概略を示す概念図である。

【図5】図1および図2で説明した化学的遺伝プログラミング装置100の機能ブロック図である。

【図6】化学的遺伝プログラミングの処理の流れを説明するための第1のフローチャートである。

【図7】化学的遺伝プログラミングの処理の流れを説明するための第2のフローチャートである。

【図8】初期的に準備されるアミノ酸の例を示す図である。

【図9】生成される表現型ツリーの一例を示す図である。

【図10】最適個体と平均個体の一致度の各世代ごとの変化を表す図である。

40

【図11】最終的に得られた関数とターゲット関数との比較結果を示す図である。

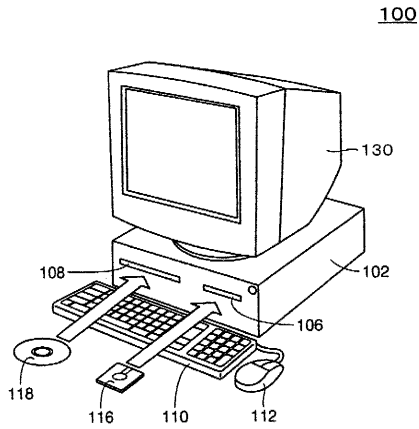
【図12】細胞の中に進化したアミノ酸のヒストグラムを示す図である。

【符号の説明】

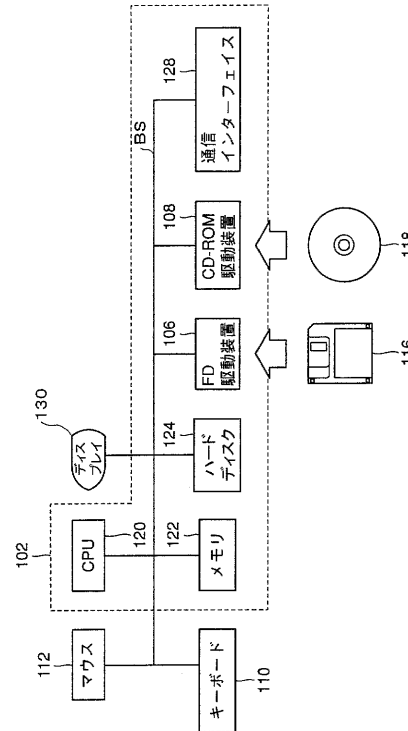
【0108】

100 化学的遺伝プログラミング装置、102 コンピュータ本体、104 ディスプレイ、106 FDドライブ、108 CD-ROMドライブ、110 キーボード、112 マウス、118 CD-ROM、120 CPU、122 メモリ、124 ハードディスク、128 通信インタフェース、130 ディスプレイ。

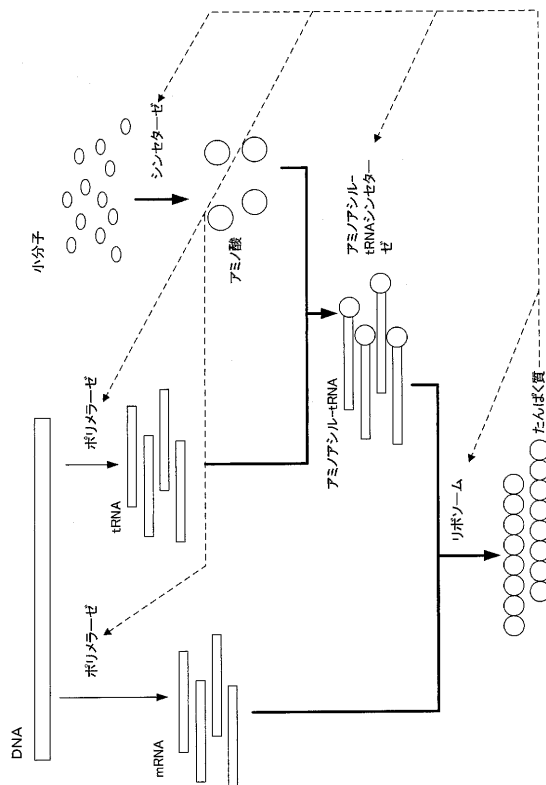
【図 1】



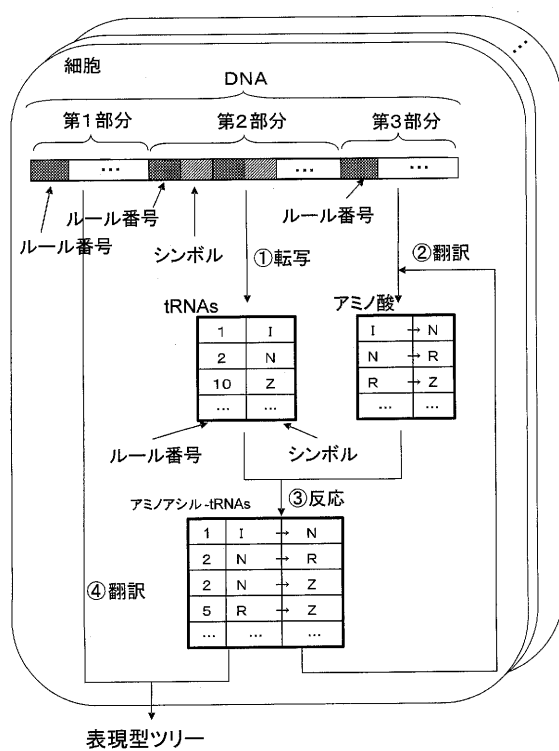
【図 2】



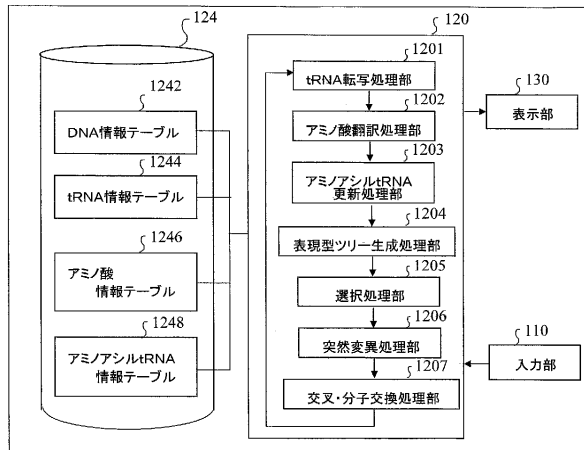
【図 3】



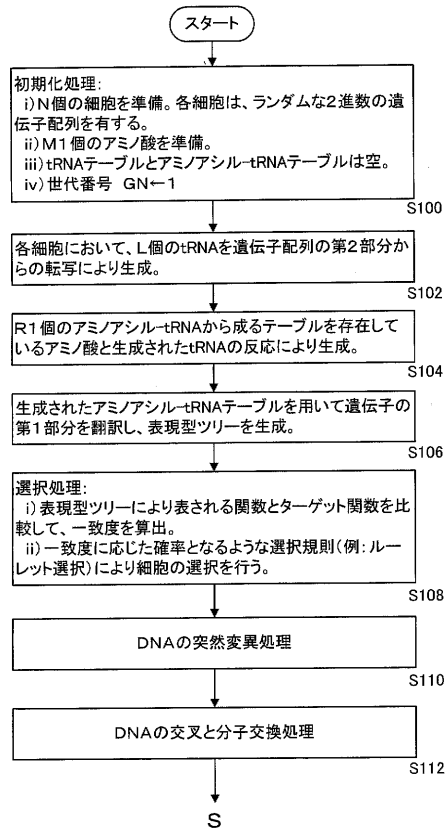
【図 4】



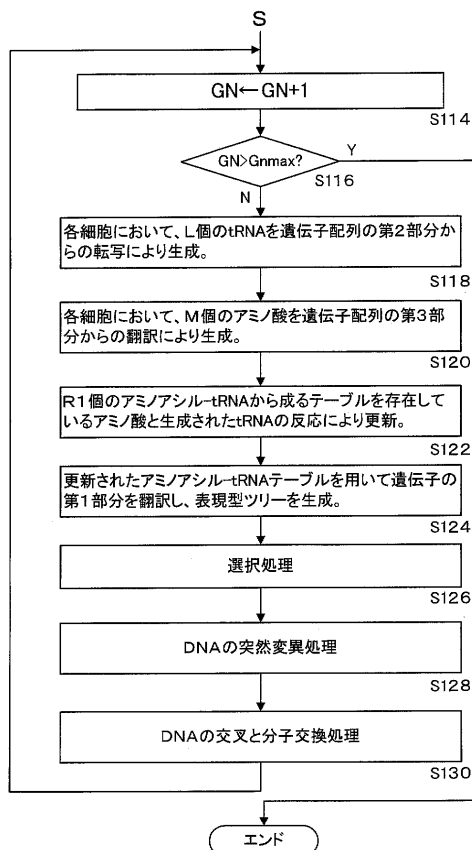
【図 5】



【図 6】



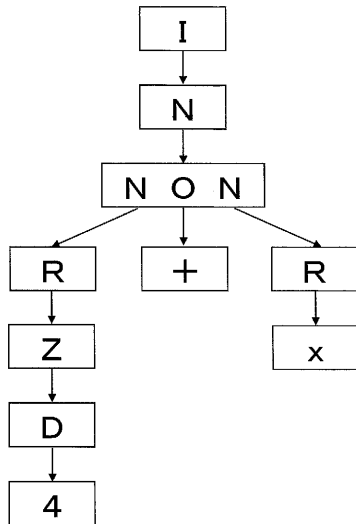
【圖 7】



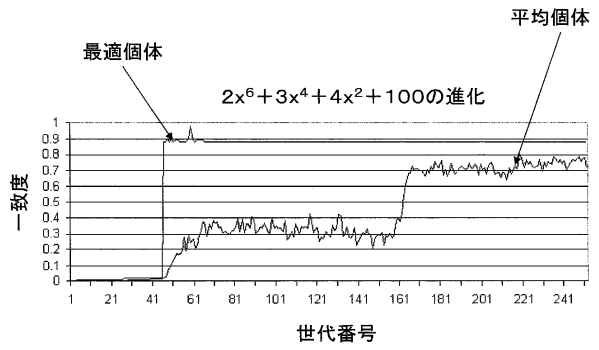
【図 8】

I	N	R	Z	O	D
$I \rightarrow N$	$N \rightarrow (NON)$ $N \rightarrow R$ $N \rightarrow Z$	$R \rightarrow x$ $R \rightarrow Z$ $R \rightarrow ZZ$ $R \rightarrow N'N$	$Z \rightarrow D$ $Z \rightarrow ZZ$	$O \rightarrow +$ $O \rightarrow -$ $O \rightarrow *$ $O \rightarrow /$	$D \rightarrow 0$ $D \rightarrow 1$ $D \rightarrow 2$ $D \rightarrow 3$ $D \rightarrow 4$ $D \rightarrow 5$ $D \rightarrow 6$ $D \rightarrow 7$ $D \rightarrow 8$ $D \rightarrow 9$

【図 9】

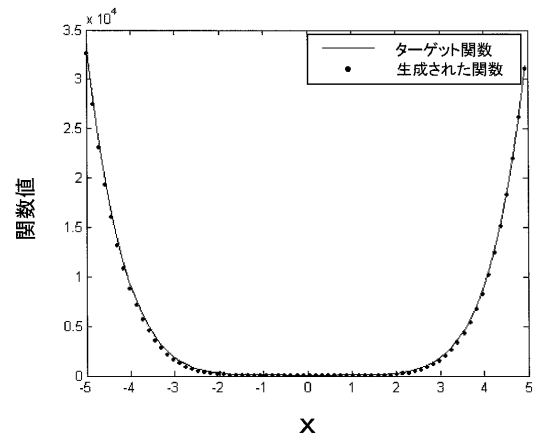


【図 10】



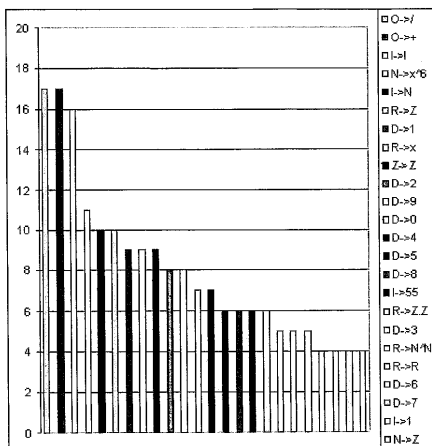
【図 11】

最終的に生成された関数とターゲット関数



【図 12】

細胞中のアミノ酸のヒストグラム



フロントページの続き

(74)代理人 100098316

弁理士 野田 久登

(74)代理人 100109162

弁理士 酒井 將行

(72)発明者 澤井 秀文

東京都小金井市貫井北町4-2-1 独立行政法人通信総合研究所内

(72)発明者 ボイツェック ピアセチェニー

京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内

(72)発明者 鈴木 秀明

京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内

審査官 稲垣 良一

(56)参考文献 特表2000-507940 (JP, A)

鈴木 秀明 他, 化学的GAの提案 -- DNAコードとその翻訳方法との共進化, 産業システム情報化研究会資料, 社団法人電気学会, 2002年12月4日, IIS-02-29, pp.33-40

Hidefumi Sawai, et al., Chemical Genetic Algorithms - Coevolutionary Genotype-Phenotype Mapping by Modelling of Metabolism in Cell, Proceedings of the IEEE 2003 Congress on Evolutionary Computation(CEC-2003), IEEE, 2003年12月12日, Vol.1, pp.639-646

新美 礼彦 他, オブジェクト指向を用いた決定木と ADF GP の組み合わせによる学習, 知識ベースシステム研究会資料, 社団法人人工知能学会, 1999年1月27日, SIG-KBS-9802, pp.19-24

(58)調査した分野(Int.Cl., DB名)

G06N 3/12

G06N 3/00

JSTPlus (JDreamII)