

明 細 書

発明の名称：

機械学習装置、基板処理装置、学習済みモデル、機械学習方法、機械学習プログラム

技術分野

[0001] 本開示は、機械学習装置、基板処理装置、学習済みモデル、機械学習方法、機械学習プログラムに関する。

背景技術

[0002] 半導体装置の配線形成プロセスとして、配線溝およびビアホールに金属（配線材料）を埋め込むようにしたプロセス（いわゆる、ダマシンプロセス）が知られている。これは、層間絶縁膜に予め形成された配線溝やビアホールに、アルミニウムや銅、銀などの金属を埋め込んだ後、余分な金属を化学機械研磨（CMP）によって除去して平坦化するプロセス技術である。

[0003] 図1A～図1Dは、半導体装置における銅配線形成例を工程順に示す図である。まず、図1Aに示すように、半導体素子が形成された半導体基材1上の導電層1aの上に、たとえばSiO₂からなる酸化膜やLow-k材膜などの絶縁膜（層間絶縁膜）2を堆積し、この絶縁膜2の内部に、たとえばリソグラフィ・エッチング技術により、配線用の微細凹部としてのビアホール3と配線溝4を形成し、その上にTa₂N₅などからなるバリア層5、さらにその上に電界めっきにおける給電層としてのシード層6をスパッタリングなどにより形成する。

[0004] そして、図1Bに示すように、基板（研磨対象物）Wの表面に銅めっきを施すことで、基板Wのビアホール3および配線溝4内に銅を充填させるとともに、絶縁膜2上に銅膜7を堆積させる。その後、図1Cに示すように、化学機械研磨（CMP）などにより、バリア層5上のシード層6および銅膜7を除去してバリア層5の表面を露出させ、さらに、図1Dに示すように、絶縁膜2上のバリア層5、および必要に応じて、絶縁膜2の表層の一部を除去

して、絶縁膜 2 の内部にシード層 6 と銅膜 7 からなる配線（銅配線） 8 を形成する。

- [0005] 研磨プロセスにおけるスループットを向上させるため、2つの研磨ユニットと1つの洗浄ユニットとを備えた研磨装置が開発されている。このような研磨装置において、研磨後の基板（研磨対象物）は、2つの研磨ユニットから1つの洗浄ユニットに順次供給される。この場合、1枚の基板が洗浄工程に入ると、当該洗浄工程が終了するまで、他の基板は洗浄工程に入ることができない。そのため、研磨を終了した基板に対する洗浄を研磨直後に開始することができず、1つ前の基板の洗浄が終了まで待機する状況が発生する。
- [0006] ここで、金属膜研磨プロセス、たとえば銅配線形成プロセスにおける銅膜研磨プロセスにおいて、研磨後の基板が研磨終了後にそのままウェットな状態で放置されると、基板表面の銅配線を形成する銅の腐食が進行する。銅は、半導体回路において配線を形成するため、その腐食は配線抵抗の増大に繋がる。
- [0007] 研磨終了後、洗浄を開始するまでの間における、銅配線を構成する銅の腐食の進行を遅くするために、基板表面に純水を供給して、研磨後の基板表面が直接大気に晒されないようにすることが一般的に行われている。しかしながら、この方法では、銅の腐食を十分に抑制することはできない。銅の腐食をより効果的に抑制するためには、研磨終了から洗浄開始までの時間自体を極力短くすることが求められる。
- [0008] 従来、たとえば基板処理装置において、基板の搬送、処理および洗浄の工程を予め定められたタイムチャートに従って管理するスケジューラが提案されている。特許第5023146号公報では、第1研磨ユニットおよび第2研磨ユニットでの平均研磨時間と、搬送機構での平均搬送時間と、洗浄ユニットでの平均洗浄時間とを予め記憶しておき、タイムチャートの作成時に、基板に対する研磨終了から洗浄開始までの時間を最短にするように、予め記憶しておいた平均研磨時間、平均搬送時間および平均洗浄時間に基づいて、第1研磨ユニットおよび第2研磨ユニットでの研磨開始時刻を決定すること

が提案されている。

発明の概要

[0009] しかしながら、本件発明者の知見によれば、予め定められたタイムチャートに従って工程を管理する方法では、以下のような不都合がある。すなわち、研磨ユニットでの研磨時間は終点検出により決定されるため、研磨時間にはばらつきが存在する。これは、異なる製品であれば異なるレシピで終点検出するからであり、また、同じレシピであっても研磨時間と消耗部材の使用時間との間に相関があるからである。また、機械的なばらつきにより、各ユニットの動作時間にもばらつきが存在する。また、特定のユニット同士の動作にインターロックがあり、任意に動作できない場合がある。また、複数の処理ルートが混在する場合もある。また、特定のユニットが故障して突発的な通行止めが発生する場合もある。したがって、たとえば平均搬送時間がX秒であるのに対し、実際の動作時間が0.5秒遅くなった場合に、タイムチャートが後ろにずれることで、次の動作に大きな遅れが生じる状態となる可能性がある。

[0010] 基板の搬送開始のタイミングおよびその搬送ルートを装置内におけるその時その時の状態に応じて適切に決定することを可能にする機械学習装置、基板処理装置、学習済みモデル、機械学習方法、機械学習プログラムを提供することが望まれる。また、基板の搬送ルートが予め決められている場合に、基板の搬送開始のタイミングを装置内におけるその時その時の状態に応じて適切に決定することを可能にする機械学習装置、基板処理装置、学習済みモデル、機械学習方法、機械学習プログラムを提供することが望まれる。また、処理ユニットにおける表面処理時間を精度よく予測することを可能にする機械学習装置、基板処理装置、学習済みモデル、機械学習方法、機械学習プログラムを提供することが望まれる。

[0011] 本開示の一態様に係る機械学習装置は、
複数枚の基板を収容するカセットが載置される載置部と、
基板を表面処理する第1処理ユニットおよび第2処理ユニットと、

表面処理後の基板を洗浄する洗浄ユニットと、
前記載置部と前記第 1 処理ユニットおよび第 2 処理ユニットと前記洗浄ユニットとの間で基板を搬送する搬送部と、
前記第 1 処理ユニットおよび第 2 処理ユニットと前記洗浄ユニットと前記搬送部の動作を制御する制御部と、
を有する基板処理装置または当該基板処理装置のシミュレータに対して、機械学習を行う機械学習装置であって、
前記基板処理装置内における基板の位置および各ユニット内に位置する基板の当該ユニット内での経過時間を含む状態情報を取得する状態情報取得部と、
ある状態において、新たな基板をカセットから取り出すか否かおよび取り出す場合には第 1 処理ユニットおよび第 2 処理ユニットのどちらに搬送するかの行動を行うことに対する価値を予測する予測モデルを有し、前記状態情報取得部により取得された状態情報を入力として前記予測モデルに基づいて 1 つの行動を選択する行動選択部と、
前記行動選択部により選択された行動を行うように前記制御部に指示信号を送信する指示信号送信部と、
予め定められた枚数の基板処理終了後、単位時間あたりの処理枚数と、表面処理後の基板が前記洗浄ユニットにて洗浄開始となるまでに待たされた待ち時間とを含む動作結果を取得する動作結果取得部と、
前記処理枚数が多くかつ前記待ち時間が短いほど報酬が大きくなるように、前記動作結果取得部により取得された動作結果に基づいて報酬を計算し、当該報酬に基づいて前記予測モデルを更新する予測モデル更新部と、
を備える。

図面の簡単な説明

- [0012] [図1A]図 1 A は、半導体装置における銅配線形成例を工程順に示す図である。
。
[図1B]図 1 B は、半導体装置における銅配線形成例を工程順に示す図である

。

[図1C]図1 Cは、半導体装置における銅配線形成例を工程順に示す図である

。

[図1D]図1 Dは、半導体装置における銅配線形成例を工程順に示す図である

。

[図2]図2は、一実施の形態に係る基板処理装置の全体構成の概要を示す平面図である。

[図3]図3は、図2に示す基板処理装置の概要を示す構成図である。

[図4]図4は、スループットが最大となるように図2に示す基板処理装置を制御部により制御するときのタイムチャートである。

[図5]図5は、第1の実施形態に係る機械学習装置の構成を示すブロック図である。

[図6]図6は、第1の実施形態に係る予測モデルの構成の一例を説明するための模式図である。

[図7]図7は、第1の実施形態に係る機械学習方法の一例を示すフローチャートである。

[図8]図8は、第2の実施形態に係る機械学習装置の構成を示すブロック図である。

[図9]図9は、第2の実施形態に係る予測モデルの構成を説明するための模式図である。

[図10]図10は、第2の実施形態に係る機械学習方法の一例を示すフローチャートである。

[図11]図11は、第3の実施形態に係る機械学習装置の構成を示すブロック図である。

[図12]図12は、第3の実施形態に係る予測モデルの構成を説明するための模式図である。

[図13]図13は、第3の実施形態に係る機械学習方法の一例を示すフローチャートである。

発明を実施するための形態

- [0013] 実施形態の第1の態様に係る機械学習装置は、
- 複数枚の基板を収容するカセットが載置される載置部と、
- 基板を表面処理する第1処理ユニットおよび第2処理ユニットと、
- 表面処理後の基板を洗浄する洗浄ユニットと、
- 前記載置部と前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットとの間で基板を搬送する搬送部と、
- 前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットと前記搬送部の動作を制御する制御部と、
- を有する基板処理装置または当該基板処理装置のシミュレータに対して、機械学習を行う機械学習装置であって、
- 前記基板処理装置内における基板の位置および各ユニット内に位置する基板の当該ユニット内での経過時間を含む状態情報を取得する状態情報取得部と、
- ある状態において、新たな基板をカセットから取り出すか否かおよび取り出す場合には第1処理ユニットおよび第2処理ユニットのどちらに搬送するかの行動を行うことに対する価値を予測する予測モデルを有し、前記状態情報取得部により取得された状態情報を入力として前記予測モデルに基づいて1つの行動を選択する行動選択部と、
- 前記行動選択部により選択された行動を行うように前記制御部に指示信号を送信する指示信号送信部と、
- 予め定められた枚数の基板処理終了後、単位時間あたりの処理枚数と、表面処理後の基板が前記洗浄ユニットにて洗浄開始となるまでに待たされた待ち時間とを含む動作結果を取得する動作結果取得部と、
- 前記処理枚数が多くかつ前記待ち時間が短いほど報酬が大きくなるように、前記動作結果取得部により取得された動作結果に基づいて報酬を計算し、当該報酬に基づいて前記予測モデルを更新する予測モデル更新部と、
- を備える。

[0014] このような態様によれば、機械学習装置は、基板処理装置内におけるその時その時の基板の位置および各ユニット内に位置する基板の当該ユニット内での経過時間を含む状態情報に応じて、予測モデルに基づいて、新たな基板をカセットから取り出すか否かおよび取り出す場合には第1処理ユニットおよび第2処理ユニットのどちらに搬送するかの行動を選択することを試行錯誤し、あらかじめ定められた枚数の基板処理終了後、単位時間あたりの処理枚数が多くかつ表面処理後の基板が洗浄開始となるまでに待たされた待ち時間が短くなるほど大きな報酬を獲得し、当該報酬に基づいて予測モデルを更新することを繰り返すことにより、予測モデルの機械学習（強化学習）を行っている。そのため、このような機械学習装置により生成された学習済みの予測モデルを利用することにより、基板の搬送開始のタイミングおよびその搬送ルート、基板処理装置内におけるその時その時の状態に応じて、（単位時間あたりの処理枚数が多くかつ待ち時間が短くなるように）適切に決定することが可能になる。

[0015] 実施形態の第2の態様に係る機械学習装置は、第1の態様に係る機械学習装置であって、

前記第1処理ユニットおよび第2処理ユニットは、基板を研磨する研磨ユニットである。

[0016] 実施形態の第3の態様に係る機械学習装置は、第1または2の態様に係る機械学習装置であって、

前記状態情報は、前記第1処理ユニットおよび第2処理ユニットにて使用される消耗部材の使用時間をさらに含む。

[0017] 実施形態の第4の態様に係る機械学習装置は、第2の態様を引用する第3の態様に係る機械学習装置であって、

前記消耗部材は、回転テーブルに取り付けられた研磨パッド、トップリングに取り付けられて基板の外周を支持するリテーナリング、トップリングに取り付けられて基板の裏面を支持する弾性膜のうちの1つまたは2つ以上である。

- [0018] 実施形態の第5の態様に係る機械学習装置は、第1～4のいずれかの態様に係る機械学習装置であって、
前記状態情報は、前記カセット内に収容された基板に予め施されている処理のレシピ情報をさらに含む。
- [0019] 実施形態の第6の態様に係る機械学習装置は、第1～5のいずれかの態様に係る機械学習装置であって、
前記状態情報は、前記第1処理ユニットおよび第2処理ユニットの故障発生情報または連続運転時間をさらに含む。
- [0020] 実施形態の第7の態様に係る機械学習装置は、第1～6のいずれかの態様に係る機械学習装置であって、
前記状態情報は、前記第1処理ユニットおよび第2処理ユニットでの表面処理のレシピ情報をさらに含む。
- [0021] 実施形態の第8の態様に係る基板処理装置は、
複数枚の基板を収容するカセットが載置される載置部と、
基板を表面処理する第1処理ユニットおよび第2処理ユニットと、
表面処理後の基板を洗浄する洗浄ユニットと、
前記載置部と前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットとの間で基板を搬送する搬送部と、
前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットと前記搬送部の動作を制御する制御部と、
を備えた基板処理装置であって、
前記制御部は、第1～7のいずれかの態様に係る機械学習装置により生成された学習済みモデルを有し、当該基板処理装置内における基板の位置および各ユニット内に位置する基板の当該ユニット内での経過時間を含む状態情報を入力として、前記学習済みモデルに基づいて、新たな基板をカセットから取り出すか否かおよび取り出す場合には第1処理ユニットおよび第2処理ユニットのどちらに搬送するかの行動を選択し、選択した行動を行うように前記搬送部の動作を制御する。

[0022] 実施形態の第9の態様に係る学習済みモデル（チューニングされたニューラルネットワークシステム）は、

複数枚の基板を収容するカセットが載置される載置部と、

基板を表面処理する第1処理ユニットおよび第2処理ユニットと、

表面処理後の基板を洗浄する洗浄ユニットと、

前記載置部と前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットとの間で基板を搬送する搬送部と、

前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットと前記搬送部の動作を制御する制御部と、

を有する基板処理装置または当該基板処理装置のシミュレータに対して、機械学習を行うことにより生成された学習済みモデル（チューニングされたニューラルネットワークシステム）であって、

入力層と、入力層に接続された1または2以上の中間層と、中間層に接続された出力層とを有し、

前記基板処理装置内における基板の位置および各ユニット内に位置する基板の当該ユニット内での経過時間を含む状態情報が取得され、取得された状態情報が入力層に入力され、それにより出力層から出力される、新たな基板をカセットから取り出すか否かおよび取り出す場合には第1処理ユニットおよび第2処理ユニットのどちらに搬送するかの行動を行うことに対する価値に基づいて1つの行動が選択され、選択された行動を行うように前記搬送部の動作が制御され、予め定められた枚数の基板処理終了後、単位時間あたりの処理枚数と、表面処理後の基板が前記洗浄ユニットにて洗浄開始となるまでに待たされた待ち時間とを含む動作結果が取得され、前記処理枚数が多くかつ前記待ち時間が短いほど報酬が大きくなるように、取得された動作結果に基づいて報酬が計算され、当該報酬に基づいて各ノードのパラメータが更新される処理が繰り返されることにより、前記処理枚数が多くかつ前記待ち時間が短くなるような基板の搬送開始のタイミングおよびその搬送ルートを強化学習したものであり、

前記基板処理装置内における基板の位置および各ユニット内に位置する基板の当該ユニット内での経過時間を含む状態情報が入力層に入力されると、新たな基板をカセットから取り出すか否かおよび取り出す場合には第1処理ユニットおよび第2処理ユニットのどちらに搬送するかの行動を行うことに対する価値を予測して出力層から出力するよう、コンピュータを機能させるための学習済みモデル（チューニングされたニューラルネットワークシステム）である。

- [0023] 実施形態の第10の態様に係る機械学習方法は、
- 複数枚の基板を収容するカセットが載置される載置部と、
 - 基板を表面処理する第1処理ユニットおよび第2処理ユニットと、
 - 表面処理後の基板を洗浄する洗浄ユニットと、
 - 前記載置部と前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットとの間で基板を搬送する搬送部と、
 - 前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットと前記搬送部の動作を制御する制御部と、
 - を有する基板処理装置または当該基板処理装置のシミュレータに対して、コンピュータが実行する機械学習方法であって、
 - 前記基板処理装置内における基板の位置および各ユニット内に位置する基板の当該ユニット内での経過時間を含む状態情報を取得する状態情報取得ステップと、
 - 前記状態情報取得ステップにおいて取得された状態情報を入力として、ある状態において、新たな基板をカセットから取り出すか否かおよび取り出す場合には第1処理ユニットおよび第2処理ユニットのどちらに搬送するかの行動を行うことに対する価値を予測する予測モデルに基づいて、1つの行動を選択する行動選択ステップと、
 - 前記行動選択ステップにおいて選択された行動を行うように前記制御部に指示信号を送信する指示信号送信ステップと、
 - 予め定められた枚数の基板処理終了後、単位時間あたりの処理枚数と、表

面処理後の基板が前記洗浄ユニットにて洗浄開始となるまでに待たされた待ち時間とを含む動作結果を取得する動作結果取得ステップと、

前記処理枚数が多くかつ前記待ち時間が短いほど報酬が大きくなるように、前記動作結果取得ステップにおいて取得された動作結果に基づいて報酬を計算し、当該報酬に基づいて前記予測モデルを更新する予測モデル更新ステップと、

を含む。

- [0024] 実施形態の第1の態様に係る機械学習プログラムは、
- 複数枚の基板を収容するカセットが載置される載置部と、
- 基板を表面処理する第1処理ユニットおよび第2処理ユニットと、
- 表面処理後の基板を洗浄する洗浄ユニットと、
- 前記載置部と前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットとの間で基板を搬送する搬送部と、
- 前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットと前記搬送部の動作を制御する制御部と、
- を有する基板処理装置または当該基板処理装置のシミュレータに対して、機械学習を行うよう、コンピュータを機能させるための機械学習プログラムであって、
- 前記コンピュータを、
- 前記基板処理装置内における基板の位置および各ユニット内に位置する基板の当該ユニット内での経過時間を含む状態情報を取得する状態情報取得部と、
- ある状態において、新たな基板をカセットから取り出すか否かおよび取り出す場合には第1処理ユニットおよび第2処理ユニットのどちらに搬送するか of 行動を行うことに対する価値を予測する予測モデルを有し、前記状態情報取得部により取得された状態情報を入力として前記価値関数に基づいて1つの行動を選択する行動選択部と、
- 前記行動選択部により選択された行動を行うように前記制御部に指示信号

を送信する指示信号送信部と、

予め定められた枚数の基板処理終了後、単位時間あたりの処理枚数と、表面処理後の基板が前記洗浄ユニットにて洗浄開始となるまでに待たされた待ち時間とを含む動作結果を取得する動作結果取得部と、

前記処理枚数が多くかつ前記待ち時間が短いほど報酬が大きくなるように、前記動作結果取得部により取得された動作結果に基づいて報酬を計算し、当該報酬に基づいて前記予測モデルを更新する予測モデル更新部と、
として機能させる。

[0025] 実施形態の第12の態様に係る機械学習装置は、

複数枚の基板を収容するカセットが載置される載置部と、

基板を表面処理する第1処理ユニットおよび第2処理ユニットと、

表面処理後の基板を洗浄する洗浄ユニットと、

前記載置部と前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットとの間で基板を搬送する搬送部と、

前記カセットから取り出される基板の順番と前記第1処理ユニットおよび第2処理ユニットのどちらに搬送するかとの対応関係が規定された搬送ルールに従って、前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットと前記搬送部の動作を制御する制御部であって、

を有する基板処理装置または当該基板処理装置のシミュレータに対して、機械学習を行う機械学習装置であって、

前記基板処理装置内における基板の位置および各ユニット内に位置する基板の当該ユニット内での経過時間を含む状態情報を取得する状態情報取得部と、

ある状態において、新たな基板をカセットから取り出すか否かの行動を行うことに対する価値を予測する予測モデルを有し、前記状態情報取得部により取得された状態情報を入力として前記予測モデルに基づいて1つの行動を選択する行動選択部と、

前記行動選択部により選択された行動を行うように前記制御部に指示信号

を送信する指示信号送信部と、

予め定められた枚数の基板処理終了後、単位時間あたりの処理枚数を含む動作結果を取得する動作結果取得部と、

前記処理枚数が多いほど報酬が大きくなるように、前記動作結果取得部により取得された動作結果に基づいて報酬を計算し、当該報酬に基づいて前記予測モデルを更新する予測モデル更新部と、
を備える。

[0026] このような態様によれば、機械学習装置は、基板処理装置内におけるその時その時の基板の位置および各ユニット内に位置する基板の当該ユニット内での経過時間を含む状態情報に応じて、予測モデルに基づいて、新たな基板をカセットから取り出すか否かの行動を選択することを試行錯誤し、あらかじめ定められた枚数の基板処理終了後、単位時間あたりの処理枚数が多くなるほど大きな報酬を獲得し、当該報酬に基づいて予測モデルを更新することを繰り返すことにより、予測モデルの機械学習（強化学習）を行っている。そのため、このような機械学習装置により生成された学習済みの予測モデルを利用することにより、基板の搬送開始のタイミングを、装置内におけるその時その時の状態に応じて、（単位時間あたりの処理枚数が多くなるように）適切に決定することが可能になる。

[0027] 実施形態の第13の態様に係る機械学習装置は、第12の態様に係る機械学習装置であって、

前記第1処理ユニットおよび第2処理ユニットは、基板を研磨する研磨ユニットである。

[0028] 実施形態の第14の態様に係る機械学習装置は、第12または13の態様に係る機械学習装置であって、

前記状態情報は、前記第1処理ユニットおよび第2処理ユニットにて使用される消耗部材の使用時間をさらに含む。

[0029] 実施形態の第15の態様に係る機械学習装置は、第13の態様を引用する第14の態様に係る機械学習装置であって、

前記消耗部材は、回転テーブルに取り付けられた研磨パッド、トップリングに取り付けられて基板の外周を支持するリテーナリング、トップリングに取り付けられて基板の裏面を支持する弾性膜のうちの1つまたは2つ以上である。

[0030] 実施形態の第16の態様に係る機械学習装置は、第12～15のいずれかの態様に係る機械学習装置であって、

前記状態情報は、前記カセット内に収容された基板に予め施されている処理のレシピ情報をさらに含む。

[0031] 実施形態の第17の態様に係る機械学習装置は、第12～16のいずれかの態様に係る機械学習装置であって、

前記状態情報は、前記第1処理ユニットおよび第2処理ユニットの連続運転時間をさらに含む。

[0032] 実施形態の第18の態様に係る機械学習装置は、第12～17のいずれかの態様に係る機械学習装置であって、

前記状態情報は、前記第1処理ユニットおよび第2処理ユニットでの表面処理のレシピ情報をさらに含む。

[0033] 実施形態の第19の態様に係る基板処理装置は、

複数枚の基板を収容するカセットが載置される載置部と、

基板を表面処理する第1処理ユニットおよび第2処理ユニットと、

表面処理後の基板を洗浄する洗浄ユニットと、

前記載置部と前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットとの間で基板を搬送する搬送部と、

前記カセットから取り出される基板の順番と前記第1処理ユニットおよび第2処理ユニットのどちらに搬送するかとの対応関係が規定された搬送ルールに従って、前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットと前記搬送部の動作を制御する制御部と、

を備えた基板処理装置であって、

前記制御部は、第12～18のいずれかの態様に係る機械学習装置により

生成された学習済みモデルを有し、当該基板処理装置内における基板の位置および各ユニット内に位置する基板の当該ユニット内での経過時間を含む状態情報を入力として、前記学習済みモデルに基づいて、新たな基板をカセットから取り出すか否かの行動を選択し、選択した行動を行うように前記搬送部の動作を制御する。

[0034] 実施形態の第20の態様に係る学習済みモデル（チューニングされたニューラルネットワークシステム）は、

複数枚の基板を収容するカセットが載置される載置部と、

基板を表面処理する第1処理ユニットおよび第2処理ユニットと、

表面処理後の基板を洗浄する洗浄ユニットと、

前記載置部と前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットとの間で基板を搬送する搬送部と、

前記カセットから取り出される基板の順番と前記第1処理ユニットおよび第2処理ユニットのどちらに搬送するかとの対応関係が規定された搬送ルールに従って、前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットと前記搬送部の動作を制御する制御部と、

を有する基板処理装置または当該基板処理装置のシミュレータに対して、機械学習を行うことにより生成された学習済みモデル（チューニングされたニューラルネットワークシステム）であって、

入力層と、入力層に接続された1または2以上の中間層と、中間層に接続された出力層とを有し、

前記基板処理装置内における基板の位置および各ユニット内に位置する基板の当該ユニット内での経過時間を含む状態情報が取得され、取得された状態情報が入力層に入力され、それにより出力層から出力される、新たな基板をカセットから取り出すか否かの行動を行うことに対する価値に基づいて1つの行動が選択され、選択された行動を行うように前記搬送部の動作が制御され、予め定められた枚数の基板処理終了後、単位時間あたりの処理枚数を含む動作結果が取得され、前記処理枚数が多いほど報酬が大きくなるように

、取得された動作結果に基づいて報酬が計算され、当該報酬に基づいて各ノードのパラメータが更新される処理が繰り返されることにより、前記処理枚数が多くなるような基板の搬送開始のタイミングを強化学習したものであり、

前記基板処理装置内における基板の位置および各ユニット内に位置する基板の当該ユニット内での経過時間を含む状態情報が入力層に入力されると、新たな基板をカセットから取り出すか否かの行動を行うことに対する価値を予測して出力層から出力するよう、コンピュータを機能させるための学習済みモデル（チューニングされたニューラルネットワークシステム）である。

- [0035] 実施形態の第21の態様に係る機械学習方法は、
- 複数枚の基板を収容するカセットが載置される載置部と、
 - 基板を表面処理する第1処理ユニットおよび第2処理ユニットと、
 - 表面処理後の基板を洗浄する洗浄ユニットと、
 - 前記載置部と前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットとの間で基板を搬送する搬送部と、
 - 前記カセットから取り出される基板の順番と前記第1処理ユニットおよび第2処理ユニットのどちらに搬送するかとの対応関係が規定された搬送ルールに従って、前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットと前記搬送部の動作を制御する制御部と、
 - を有する基板処理装置または当該基板処理装置のシミュレータに対して、コンピュータが実行する機械学習方法であって、
 - 前記基板処理装置内における基板の位置および各ユニット内に位置する基板の当該ユニット内での経過時間を含む状態情報を取得する状態情報取得ステップと、
 - 前記状態情報取得ステップにおいて取得された状態情報を入力として、ある状態において、新たな基板をカセットから取り出すか否かの行動を行うことに対する価値を予測する予測モデルに基づいて、1つの行動を選択する行動選択ステップと、

前記行動選択ステップにおいて選択された行動を行うように前記制御部に指示信号を送信する指示信号送信ステップと、

予め定められた枚数の基板処理終了後、単位時間あたりの処理枚数を含む動作結果を取得する動作結果取得ステップと、

前記処理枚数が多いほど報酬が大きくなるように、前記動作結果取得ステップにおいて取得された動作結果に基づいて報酬を計算し、当該報酬に基づいて前記予測モデルを更新する予測モデル更新ステップと、
を含む。

- [0036] 実施形態の第2の態様に係る機械学習プログラムは、
複数枚の基板を収容するカセットが載置される載置部と、
基板を表面処理する第1処理ユニットおよび第2処理ユニットと、
表面処理後の基板を洗浄する洗浄ユニットと、
前記載置部と前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットとの間で基板を搬送する搬送部と、
前記カセットから取り出される基板の順番と前記第1処理ユニットおよび第2処理ユニットのどちらに搬送するかとの対応関係が規定された搬送ルールに従って、前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットと前記搬送部の動作を制御する制御部と、
を有する基板処理装置または当該基板処理装置のシミュレータに対して、機械学習を行うよう、コンピュータを機能させるための機械学習プログラムであって、
前記コンピュータを、
前記基板処理装置内における基板の位置および各ユニット内に位置する基板の当該ユニット内での経過時間を含む状態情報を取得する状態情報取得部と、
ある状態において、新たな基板をカセットから取り出すか否かの行動を行うことに対する価値を予測する予測モデルを有し、前記状態情報取得部により取得された状態情報を入力として前記予測モデルに基づいて1つの行動を

選択する行動選択部と、

前記行動選択部により選択された行動を行うように前記制御部に指示信号を送信する指示信号送信部と、

予め定められた枚数の基板処理終了後、単位時間あたりの処理枚数を含む動作結果を取得する動作結果取得部と、

前記処理枚数が多いほど報酬が大きくなるように、前記動作結果取得部により取得された動作結果に基づいて報酬を計算し、当該報酬に基づいて前記予測モデルを更新する価値関数更新部と、

として機能させる。

[0037] 実施形態の第23の態様に係る機械学習装置は、

基板を表面処理する処理ユニットにおける表面処理のレシピ情報と、基板情報と、前記処理ユニット内にて使用される消耗部材の使用時間と、前記処理ユニットの連続運転時間と、前記処理ユニットにおける実際の表面処理時間との関係性を機械学習する機械学習装置であって、

前記処理ユニットでの表面処理のレシピ情報と、基板情報と、前記処理ユニット内にて使用される消耗部材の使用時間と、前記処理ユニットの連続運転時間とを入力情報として取得する入力情報取得部と、

前記処理ユニットでの表面処理のレシピ情報と、基板情報と、前記処理ユニット内にて使用される消耗部材の使用時間と、前記処理ユニットの連続運転時間とに基づいて、前記処理ユニットにおける表面処理時間を予測する予測モデルを有し、前記入力情報取得部により取得された入力情報を入力として、前記予測モデルに基づいて、前記処理ユニットにおける表面処理時間を予測して出力する予測部と、

前記処理ユニットにおける実際の表面処理時間を取得する実表面処理時間取得部と、

前記実表面処理時間取得部により取得された実際の表面処理時間と前記予測部により予測された表面処理時間との誤差に応じて前記予測モデルを更新する予測モデル更新部と、

を備える。

[0038] このような態様によれば、機械学習装置は、処理ユニットでの表面処理のレシピ情報と、基板情報と、処理ユニット内にて使用される消耗部材の使用時間と、処理ユニットの連続運転時間と、処理ユニットにおける実際の表面処理時間との対応関係を教師データとして、予測モデルの機械学習（教師あり学習）を行っている。そのため、このような機械学習装置により生成された学習済みの予測モデルを利用することにより、処理ユニットでの表面処理のレシピ情報と、基板情報だけでなく、処理ユニット内にて使用される消耗部材の使用時間と、処理ユニットの連続運転時間をも考慮して、処理ユニットにおける表面処理時間を精度よく予測することが可能となり、これにより、タイムチャートの作成時に、当該予測された表面処理時間に基づいて、基板の搬送開始のタイミングを精度よく決定することが可能になる。

[0039] 実施形態の第24の態様に係る機械学習装置は、第23の態様に係る機械学習装置であって、

前記処理ユニットは、基板を研磨する研磨ユニットである。

[0040] 実施形態の第25の態様に係る機械学習装置は、第24の態様に係る機械学習装置であって、

前記消耗部材は、回転テーブルに取り付けられた研磨パッド、トップリングに取り付けられて基板の外周を支持するリテーナリング、トップリングに取り付けられて基板の裏面を支持する弾性膜のうちの1つまたは2つ以上である。

[0041] 実施形態の第26の態様に係る基板処理装置は、

複数枚の基板を収容するカセットが載置される載置部と、

基板を表面処理する第1処理ユニットおよび第2処理ユニットと、

表面処理後の基板を洗浄する洗浄ユニットと、

前記載置部と前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットとの間で基板を搬送する搬送部と、

前記カセットから取り出される基板の順番と前記第1処理ユニットおよび

第2処理ユニットのどちらに搬送するか、およびその搬送開始時刻との対応関係が規定された搬送ルールに従って、前記第1処理ユニットおよび第2処理ユニットと前記洗浄ユニットと前記搬送部の動作を制御する制御部と、を備えた基板処理装置であって、

前記制御部は、第23～25のいずれかの態様に係る機械学習装置により生成された学習済みモデルを有し、前記カセットに収容された各基板に対して、前記第1処理ユニットまたは第2処理ユニットでの表面処理のレシピ情報と、基板情報と、前記第1処理ユニットまたは第2処理ユニット内にて使用される消耗部材の使用時間と、前記第1処理ユニットまたは第2処理ユニットの連続運転時間とを入力として、前記学習済みモデルに基づいて、前記第1処理ユニットまたは第2処理ユニットにおける表面処理時間を予測し、予測した表面処理時間に基づいて、前記搬送開始時刻を決定する。

[0042] 実施形態の第27の態様に係る学習済みモデル（チューニングされたニューラルネットワークシステム）は、

基板を表面処理する処理ユニットにおける表面処理のレシピ情報と、基板情報と、前記処理ユニット内にて使用される消耗部材の使用時間と、前記処理ユニットの連続運転時間と、前記処理ユニットにおける実際の表面処理時間との関係性を機械学習することにより生成された学習済みモデル（チューニングされたニューラルネットワークシステム）であって、

入力層と、入力層に接続された1または2以上の中間層と、中間層に接続された出力層とを有し、

前記処理ユニットでの表面処理のレシピ情報と、基板情報と、前記処理ユニット内にて使用される消耗部材の使用時間と、前記処理ユニットの連続運転時間とが入力層に入力され、それにより出力層から出力される出力結果と、前記処理ユニットにおける実際の表面処理時間とが比較され、その誤差に応じて各ノードのパラメータが更新される処理が繰り返されることにより、前記処理ユニットでの表面処理のレシピ情報と、基板情報と、前記処理ユニット内にて使用される消耗部材の使用時間と、前記処理ユニットの連続運転

時間と、前記処理ユニットにおける実際の表面処理時間との関係性を機械学習したものであり、

前記処理ユニットでの表面処理のレシピ情報と、基板情報と、前記処理ユニット内にて使用される消耗部材の使用時間と、前記処理ユニットの連続運転時間とが入力層に入力されると、前記処理ユニットにおける表面処理時間を予測して出力層から出力するよう、コンピュータを機能させるための学習済みモデル（ニューラルネットワークシステム）である。

[0043] 実施形態の第28の態様に係る機械学習方法は、

基板を表面処理する処理ユニットにおける表面処理のレシピ情報と、基板情報と、前記処理ユニット内にて使用される消耗部材の使用時間と、前記処理ユニットの連続運転時間と、前記処理ユニットにおける実際の表面処理時間との関係性を機械学習する、コンピュータが実行する機械学習方法であって、

前記処理ユニットでの表面処理のレシピ情報と、基板情報と、前記処理ユニット内にて使用される消耗部材の使用時間と、前記処理ユニットの連続運転時間とを入力情報として取得する入力情報取得ステップと、

前記処理ユニットでの表面処理のレシピ情報と、基板情報と、前記処理ユニット内にて使用される消耗部材の使用時間と、前記処理ユニットの連続運転時間とに基づいて、前記処理ユニットにおける表面処理時間を予測する予測モデルを利用して、前記入力情報取得ステップにおいて取得された入力情報を入力として、前記予測モデルに基づいて、前記処理ユニットにおける表面処理時間を予測する予測ステップと、

前記処理ユニットにおける実際の表面処理時間を取得する実表面処理時間取得ステップと、

前記実表面処理時間取得ステップにおいて取得された実際の表面処理時間と前記予測ステップにおいて予測された表面処理時間との誤差に応じて前記予測モデルを更新する学習モデル更新ステップと、
を含む。

[0044] 実施形態の第29の態様に係る機械学習プログラムは、

基板を表面処理する処理ユニットにおける表面処理のレシピ情報と、基板情報と、前記処理ユニット内にて使用される消耗部材の使用時間と、前記処理ユニットの連続運転時間と、前記処理ユニットにおける実際の表面処理時間との関係性を機械学習するよう、コンピュータを機能させるための機械学習プログラムであって、

前記コンピュータを、

前記処理ユニットでの表面処理のレシピ情報と、基板情報と、前記処理ユニット内にて使用される消耗部材の使用時間と、前記処理ユニットの連続運転時間とを入力情報として取得する入力情報取得部と、

前記処理ユニットでの表面処理のレシピ情報と、基板情報と、前記処理ユニット内にて使用される消耗部材の使用時間と、前記処理ユニットの連続運転時間とに基づいて、前記処理ユニットにおける表面処理時間を予測する予測モデルを有し、前記入力情報取得部により取得された入力情報を入力として、前記学習モデルに基づいて、前記処理ユニットにおける表面処理時間を予測して出力する予測部と、

前記処理ユニットにおける実際の表面処理時間を取得する実表面処理時間取得部と、

前記実表面処理時間取得部により取得された実際の表面処理時間と前記予測部により予測された表面処理時間との誤差に応じて前記予測モデルを更新する学習モデル更新部と、

として機能させる。

[0045] 以下に、添付の図面を参照して、実施の形態の具体例を詳細に説明する。

なお、以下の説明および以下の説明で用いる図面では、同一に構成され得る部分について、同一の符号を用いるとともに、重複する説明を省略する。

[0046] 以下に説明する実施の形態では、図1Bに示すように、表面に銅膜7が成膜された基板Wに対し、図1Cに示すように、バリア層5の上の銅膜7およびシード層6を研磨除去（第1研磨）してバリア層7を露出させ、次いで、

図1Dに示すように、絶縁膜2上のバリア層5および必要に応じて絶縁膜2の表層の一部を研磨除去（第2研磨）する、2段研磨を行う例を説明するが、2段研磨はあくまで一例であり、本実施の形態は、このような2段研磨に限定されないことは言うまでもない。

[0047] 図2は、一実施の形態に係る基板処理装置10の全体構成の概要を示す平面図であり、図3は、図2に示す基板処理装置10の概要を示す構成図である。

[0048] 図2に示すように、本実施の形態に係る基板処理装置10は、研磨装置であり、略矩形形状のハウジング11と、複数枚の基板（研磨対象物）を収容する複数（図示された例では3つ）のカセット12が載置される載置部14と、基板を表面処理（研磨）する第1処理ユニット20および第2処理ユニット30と、表面処理（研磨）後の基板を洗浄する洗浄ユニット40と、載置部14と第1処理ユニット20および第2処理ユニット30と洗浄ユニット40との間で基板を搬送する搬送部50と、第1処理ユニット20および第2処理ユニット30と洗浄ユニット40と搬送部50の動作を制御する制御部70と、を有している。

[0049] このうち載置部14に載置されるカセット12は、たとえばSMIF（Standard Manufacturing Interface）ポッドまたはFOUP（Font Opening Unified Pod）からなる密閉容器内に収容される。

[0050] 図2に示すように、第1処理ユニット20および第2処理ユニット30は、ハウジング11の内部のうち、その長手方向に沿った一側（図2における上側）に配置されている。本実施の形態では、第1処理ユニット20および第2処理ユニット30は、いずれも、基板を研磨する研磨ユニットである。

[0051] 第1処理ユニット20は、第1研磨部22と第2研磨部24とを有している。第1処理ユニット20の第1研磨部22は、基板Wを着脱自在に保持するトップリング22aと、表面に研磨面を有する研磨パッドが取り付けられた回転テーブル22bとを有しており、第2研磨部24は、基板Wを着脱自在に保持するトップリング24aと、表面に研磨面を有する研磨パッドが取

り付けられた回転テーブル24bとを有している。同様に、第2処理ユニット30は、第1研磨部32と第2研磨部34とを有している。第2処理ユニット30の第1研磨部32は、トップリング32aと回転テーブル32bとを有しており、第2研磨部34は、トップリング34aと回転テーブル34bとを有している。

[0052] 図2に示すように、洗浄ユニット40は、ハウジング10の内部のうち、その長手方向に沿った他側（図2における下側）に配置されている。図示された例では、洗浄ユニット40は、第1洗浄機42aと、第2洗浄機42bと、第3洗浄機42cと、第4洗浄機42dと、搬送機構44（図3参照）とを有している。第1～4洗浄機42a～42dは、ハウジング10の長手方向に沿って、この順に直列に配置されている。搬送機構44（図3参照）は、洗浄機42a～42dと同じ数（図示された例では4つ）のハンドを有し、洗浄機42a～42dの並び（すなわちハウジング10の長手方向）に沿って往復移動可能である。

[0053] 図3に示すように、搬送機構44の往復移動によって、基板Wは、第1洗浄機42a→第2洗浄機42b→第3洗浄機42c→第4洗浄機42dと順次搬送されながら洗浄される。この洗浄タクト（洗浄時間）は、洗浄機42a～42dのうちの最も洗浄時間の長い洗浄機における洗浄時間にて設定され、最も洗浄時間の長い洗浄機における洗浄工程が終了したのち、搬送機構44が駆動されて基板Wが搬送される。

[0054] 図2および図3に示すように、搬送部50は、載置部14と第1処理ユニット20および第2処理ユニット30と洗浄ユニット40とにより挟まれた領域に配置されている。図示された例では、搬送部50は、研磨前の基板Wを180°反転させる第1反転機52aと、研磨後の基板Wを180°反転させる第2反転機52bと、第1反転機52aと載置部14との間に配置された、第1搬送ロボット54aと、第2反転機52bと洗浄ユニット40との間に配置された、第2搬送ロボット54bとを有している。

[0055] 図2および図3に示すように、第1処理ユニット20と洗浄ユニット40

との間には、載置部 1 4 側から順に、第 1 リニアトランスポート 5 6 a、第 2 リニアトランスポート 5 6 b、第 3 リニアトランスポート 5 6 c および第 4 リニアトランスポート 5 6 d が配置されている。このうち第 1 リニアトランスポート 5 6 a の上方には、上述した第 1 反転機 5 2 a が配置されており、その下方には、上下に昇降可能なリフト 5 8 a が配置されている。また、第 2 リニアトランスポート 5 6 b の下方には、上下に昇降可能なプッシャ 6 0 a が配置されており、第 3 リニアトランスポート 5 6 c の下方には、上下に昇降可能なプッシャ 6 0 b が配置されている。第 4 トランスポート 5 6 d の下方には、上下に昇降可能なリフト 5 8 b が配置されている。

[0056] 図 2 および図 3 に示すように、第 2 処理ユニット 4 0 側には、載置部 1 4 側から順に、第 5 リニアトランスポート 5 6 e、第 6 リニアトランスポート 5 6 f および第 7 リニアトランスポート 5 6 g が配置されている。このうち第 5 リニアトランスポート 5 6 e の下方には、上下に昇降可能なリフト 5 8 c が配置されている。また、第 6 リニアトランスポート 5 6 f の下方には、上下に昇降可能なプッシャ 6 0 c が配置されており、第 7 リニアトランスポート 5 6 g の下方には、上下に昇降可能なプッシャ 6 0 d が配置されている。

[0057] 次に、このような構成からなる基板処理装置（研磨装置） 1 0 を用いて基板 W を表面処理（研磨）する工程の一例について説明する。

[0058] まず、載置部 1 4 に載置されたカセット 1 2 の 1 つから第 1 搬送ロボット 5 4 a により奇数枚目に取り出された基板（1 枚目、3 枚目…の基板）は、第 1 反転機 5 2 a → 第 1 リニアトランスポート 5 6 a → トップリング 2 2 a（第 1 処理ユニット 2 0 の第 1 研磨部 2 2）→ 第 2 リニアトランスポート 5 6 b → トップリング 2 4 a（第 1 処理ユニット 2 0 の第 2 研磨部 2 4）→ 第 3 リニアトランスポート 5 6 c → 第 2 搬送ロボット 5 4 b → 第 2 反転機 5 2 b → 第 1 洗浄機 4 2 a → 第 2 洗浄機 4 2 b → 第 3 洗浄機 4 2 c → 第 4 洗浄機 4 2 d → 第 1 搬送ロボット 5 4 a という経路（搬送ルート）で搬送されて、元のカセット 1 2 に戻される。

- [0059] また、載置部 1 4 に載置されたカセット 1 2 の 1 つから第 1 搬送ロボット 5 4 a により偶数枚目に取り出された基板（2 枚目、4 枚目…の基板）は、第 1 反転機 5 2 a → 第 4 リニアトランスポータ 5 6 d → 第 2 搬送ロボット 5 4 b → 第 5 リニアトランスポータ 5 6 e → トップリング 3 2 a（第 2 処理ユニット 3 0 の第 1 研磨部 3 2）→ 第 6 リニアトランスポータ 5 6 f → トップリング 3 4 a（第 2 処理ユニット 3 0 の第 2 研磨部 3 4）→ 第 7 リニアトランスポータ 5 6 g → 第 2 搬送ロボット 5 4 b → 第 2 反転機 5 2 b → 第 1 洗浄機 4 2 a → 第 2 洗浄機 4 2 b → 第 3 洗浄機 4 2 c → 第 4 洗浄機 4 2 d → 第 1 搬送ロボット 5 4 a という経路（搬送ルート）で搬送されて、元のカセット 1 2 に戻される。
- [0060] ここで、第 1 処理ユニット 2 0 の第 1 研磨部 2 2 および第 2 処理ユニット 3 0 の第 1 研磨部 3 2 では、上述したように、バリア層 5 の上の銅膜 7 およびシード層 6 が研磨除去（第 1 研磨）され、第 1 処理ユニット 2 0 の第 2 研磨部 2 4 および第 2 処理ユニット 3 0 の第 2 研磨部 3 4 では、絶縁膜 2 上のバリア層 5 および必要に応じて絶縁膜 2 の表層の一部が研磨除去（第 2 研磨）される。そして、第 2 研磨後の基板は、洗浄機 4 2 a ～ 4 2 d にて順次洗浄され、乾燥されたのち、カセット 1 2 に戻される。
- [0061] 洗浄ユニット 4 0 では、第 1 処理ユニット 2 0 にて研磨された 1 枚目の基板が第 1 洗浄機 4 2 a にて洗浄されたのち、1 枚の基板と第 2 処理ユニット 3 0 にて研磨された 2 枚目の基板が搬送機構 4 4 にて同時に把持され、1 枚目の基板が第 2 洗浄機 4 2 b に、2 枚目の基板が第 1 洗浄機 4 2 a に同時に搬送され、2 枚の基板が同時に洗浄される。そして、1 枚目の基板および 2 枚目の基板が洗浄されたのち、1 枚目および 2 枚目の基板と第 1 処理ユニット 2 0 にて研磨された 3 枚目の基板が搬送機構 4 4 にて同時に把持され、1 枚目の基板が第 3 洗浄機 4 2 c に、2 枚目の基板が第 2 洗浄機 4 2 b に、3 枚目の基板が第 1 洗浄機 4 2 a に同時に搬送され、3 枚の基板が同時に洗浄される。このような動作が順次繰り返されることで、2 つの処理ユニット 2 0、3 0 に対して、1 つの洗浄ユニット 4 0 にて対処することができる。

- [0062] この場合、スループットが最大となるように基板処理装置 10 を制御部 70 により制御すると、図 4 のタイムチャートで示すように、2 枚目の基板が研磨されたのち第 1 洗浄機 42 a にて洗浄されるまでの間に洗浄待ち時間 S_1 が生じる。また、3 枚目の基板が研磨されたのち第 1 洗浄機 42 a にて洗浄されるまでの間に洗浄待ち時間 S_2 が生じる。さらに、4 枚目の基板については、研磨されたのち第 1 洗浄機 42 a にて洗浄されるまでの間に洗浄待ち時間 S_3 、 S_4 が生じる。このように、研磨終了後に洗浄が開始されるまでの間に洗浄待ち時間が生じると、たとえば銅配線形成プロセスにあっては、銅の腐食が懸念される。
- [0063] 研磨終了から洗浄開始までの待ち時間を短くするために、特許第 5023146 号公報では、第 1 研磨ユニットおよび第 2 研磨ユニットでの平均研磨時間と、搬送機構での平均搬送時間と、洗浄ユニットでの平均洗浄時間とを予め記憶しておき、タイムチャートの作成時に、基板に対する研磨終了から洗浄開始までの時間を最短にするように、平均研磨時間、平均搬送時間および平均洗浄時間に基づいて、第 1 研磨ユニットおよび第 2 研磨ユニットでの研磨開始時刻を決定することが提案されている。
- [0064] しかしながら、本件発明者の知見によれば、予め定められたタイムチャートに従って工程を管理する方法では、以下のような不都合がある。すなわち、研磨ユニットでの研磨時間は終点検出により決定されるため、研磨時間にばらつきが存在する。これは、異なる製品であれば異なるレシピで終点検出するからであり、また、同じレシピであっても研磨時間と消耗部材の使用時間との間に相関があるからである。また、機械的なばらつきにより、各ユニットの動作時間にもばらつきが存在する。また、特定のユニット同士の動作にインターロックがあり、任意に動作できない場合がある。また、複数の処理ルートが混在する場合もある。また、特定のユニットが故障して突発的な通行止めが発生する場合もある。したがって、たとえば平均搬送時間が X 秒であるのに対し、実際の動作時間が 0.5 秒遅くなった場合に、タイムチャートが後ろにずれることで、次の動作に大きな遅れが生じる状態となる可能

性がある。

[0065] (第1の実施形態)

以下に説明する第1の実施形態に係る機械学習装置80は、以上のような点を考慮してなされたものであり、基板Wの搬送開始のタイミングおよびその搬送ルートを、基板処理装置10内におけるその時その時の状態に応じて、(単位時間あたりの処理枚数が多くかつ待ち時間が短くなるように)適切に決定することを可能にできるものである。

[0066] 図5は、第1の実施形態に係る機械学習装置80の構成を示すブロック図である。機械学習装置80の少なくとも一部は、1つのコンピュータまたは量子コンピューティングシステム、もしくは互いにネットワークを介して接続された複数のコンピュータまたは量子コンピューティングシステムによって構成されている。

[0067] 図5に示すように、機械学習装置80は、通信部81と、制御部82と、記憶部83とを有している。各部81～83は、バスやネットワークを介して通信可能に接続されている。

[0068] このうち通信部81は、基板処理装置10の制御部70に対する通信インターフェースである。通信部81は、基板処理装置10の制御部70に有線で接続されていてもよいし、無線で接続されていてもよい。

[0069] 記憶部83は、たとえばフラッシュメモリなどの不揮発性データストレージである。記憶部83には、制御部82が取り扱う各種データが記憶される。

[0070] 図5に示すように、制御部82は、状態情報取得部82aと、行動選択部82bと、指示信号送信部82cと、動作結果取得部82dと、予測モデル更新部82eとを有している。これらの各部は、機械学習装置80内のプロセッサが所定のプログラムを実行することにより実現されてもよいし、ハードウェアで実装されてもよい。

[0071] 本実施の形態において、制御部82は、単位時間あたりの処理枚数が多く、かつ、表面処理後の基板が洗浄ユニット40にて洗浄開始となるまでに待

たされる待ち時間が短くなるような基板の搬送開始のタイミングおよびその搬送ルートを、基板処理装置10内におけるその時その時の状態に応じた試行錯誤を繰り返すことで、強化学習するものである。強化学習のアルゴリズムは、特に限定されるものではないが、たとえばQ学習、SARSA法、方策勾配法、Actor-Critic法などが用いられ得る。

[0072] 状態情報取得部82aは、基板処理装置10内における基板Wの位置および各ユニット20、30、40内に位置する基板Wの当該ユニット内での経過時間を含む状態情報を、基板処理装置10の制御部70から所定の時間間隔（たとえば0.1sごと）で繰り返し取得する。

[0073] 状態情報取得部82aが基板処理装置10の制御部70から取得する状態情報は、第1処理ユニット20および第2処理ユニット30にて使用される消耗部材の使用時間をさらに含んでもよい。本件発明者が鋭意検討を重ねた結果、第1処理ユニット20および第2処理ユニット30での処理時間（たとえば終点検出により決定される研磨時間）は、第1処理ユニット20および第2処理ユニット30にて使用される消耗部材の使用時間と相関関係があることが見出された。したがって、後述する予測モデル85に入力される状態情報が、第1処理ユニット20および第2処理ユニット30にて使用される消耗部材の使用時間を含んでいる場合には、予測モデル85による予測精度をさらに向上させることができる。消耗部材は、たとえば、回転テーブル22b、24b、32b、34bに取り付けられた研磨パッド、トップリング22a、24a、32a、34aに取り付けられて基板Wの外周を支持するリテーナリング、トップリング22a、24a、32a、34aに取り付けられて基板Wの裏面を支持する弾性膜のうちの1つまたは2つ以上であってよい。

[0074] 状態情報取得部82aが基板処理装置10の制御部70から取得する状態情報は、カセット12内に収容された基板Wに予め施されている処理のレシピ情報（たとえば図1Bに示す基板W表面の銅膜7の成膜条件）をさらに含んでもよい。本件発明者が鋭意検討を重ねた結果、第1処理ユニット2

0および第2処理ユニット30での処理時間（たとえば終点検出により決定される研磨時間）は、カセット12内に收容された基板Wに予め施されている処理のレシピ情報と相関関係があることが見出された。したがって、後述する予測モデル85に入力される状態情報が、カセット12内に收容された基板Wに予め施されている処理のレシピ情報を含んでいる場合には、予測モデル85による予測精度を向上させることができる。

[0075] 状態情報取得部82aが基板処理装置10の制御部70から取得する状態情報は、第1処理ユニット20および第2処理ユニット30の故障発生情報または連続運転時間をさらに含んでもよい。本件発明者が鋭意検討を重ねた結果、第1処理ユニット20および第2処理ユニット30にて運転間隔が空くと水が滞留したりして一回洗い直すことによりコンデションが大きく変わることから、第1処理ユニット20および第2処理ユニット30での処理時間（たとえば終点検出により決定される研磨時間）は、第1処理ユニット20および第2処理ユニット30の連続運転時間と相関関係があることが見出された。したがって、後述する予測モデル85に入力される状態情報が、第1処理ユニット20および第2処理ユニット30の連続運転時間を含んでいる場合には、予測モデル85による予測精度を向上させることができる。また、後述する予測モデル85に入力される状態情報が、第1処理ユニット20および第2処理ユニット30の故障発生情報を含んでいる場合にも、予測モデル85による予測精度を向上させることができる。これは、一方のユニットに故障が発生した場合には、その状況に応じて故障が発生していないユニットへと搬送ルートを変更することで、通行止めによる大幅な遅延の発生を回避できるからであると考えられる。

[0076] 状態情報取得部82aが基板処理装置10の制御部70から取得する状態情報は、第1処理ユニット20および第2処理ユニット30での表面処理（研磨処理）のレシピ情報をさらに含んでもよい。本件発明者が鋭意検討を重ねた結果、第1処理ユニット20および第2処理ユニット30での処理時間（たとえば終点検出により決定される研磨時間）は、第1処理ユニット

20および第2処理ユニット30での表面処理（研磨処理）のレシピ情報と相関関係があることが見出された。したがって、後述する予測モデル85に入力される状態情報が、第1処理ユニット20および第2処理ユニット30での表面処理（研磨処理）のレシピ情報を含んでいる場合には、予測モデル85による予測精度を向上させることができる。

[0077] 行動選択部82bは、ある状態 s_t において、新たな基板Wをカセット12から取り出すか否か、および、取り出す場合には第1処理ユニット20および第2処理ユニット30のどちらに搬送するかの行動を行うことに対する価値（Q学習におけるQ値）を予測する予測モデル85（図6参照）を有している。

[0078] 図6は、予測モデル85の構成の一例を説明するための模式図である。図6に示す例では、予測モデル85は、ニューラルネットワークシステムであり、入力層と、入力層に接続された1または2以上の中間層と、中間層に接続され出力層とを有する階層型のニューラルネットワークまたは量子ニューラルネットワーク（QNN）を含んでいる。図6では、階層型のニューラルネットワークとして、フィードフォワードニューラルネットワークが図示されているが、畳み込みニューラルネットワーク（CNN）やリカレントニューラルネットワーク（RNN）など、様々なタイプのニューラルネットワークが使用され得る。予測モデル85は、中間層が2層以上に多層化されたニューラルネットワーク、すなわちディープラーニング（深層学習）を含んでもよい。

[0079] 図6に示すように、予測モデル85は、状態情報取得部82aにより取得された状態情報が入力層に入力されると、新たな基板Wをカセット12から取り出すか否かおよび取り出す場合には第1処理ユニット20および第2処理ユニット30のどちらに搬送するかの行動を行うことに対する価値（Q学習におけるQ値）を予測して出力層から出力する。

[0080] 行動選択部82bは、複数の予測モデル85を有し、当該複数の予測モデル85による予測結果の組み合わせ（すなわちアンサンブル学習）に基づい

て、各行動の価値（Q値）を推定して出力してもよい。

- [0081] 行動選択部82bは、状態情報取得部82aにより取得された状態情報を入力として予測モデル85に基づいて1つの行動（すなわち、新たな基板Wをカセット12から取り出して第1処理ユニット20に搬送する行動と、新たな基板Wをカセット12から取り出して第2処理ユニット20に搬送する行動と、新たな基板Wをカセット12から取り出さない行動のうちのいずれか）を選択する。選択方法としては、たとえば、行動選択部82bは、予測モデル85により予測された各行動の価値（Q値）を比較して、最も価値（Q値）が高い行動を選択してもよい（greedy法）、予め定められた確率 ε 以下でランダムに行動を選択し、それ以外では最も価値（Q値）が高い行動を選択してもよい（ ε -greedy法）。
- [0082] 指示信号送信部82cは、行動選択部82bにより選択された行動を行うように基板処理装置10の制御部70に指示信号を送信する。基板処理装置10の制御部70が指示信号送信部82cから受信した指示信号に従って行動することにより、基板処理装置10内の状態 s_t は、次の状態 s_{t+1} に遷移する。
- [0083] 予測モデル更新部82eは、遷移後の状態 s_{t+1} が終端状態（予め定められた枚数の基板処理が終了した状態）ではなかった場合には、状態情報取得部82aにより取得される遷移後の状態 s_{t+1} の状態情報を予測モデル85の入力層に入力した場合に出力層から出力される各行動の価値のうちの最大の価値（Q値）に基づいて予測モデル85を更新（たとえば、ニューラルネットワークにおける各ノードのパラメータ（重みや閾値など）を更新）してもよい。
- [0084] 動作結果取得部82dは、予め定められた枚数の基板処理終了後（すなわち遷移後の状態 s_{t+1} が終端状態である場合）、単位時間あたりの処理枚数と、表面処理後の基板が洗浄ユニット40にて洗浄開始となるまでに待たされた待ち時間とを含む動作結果を、基板処理装置10の制御部70から取得する。ここで「待ち時間」は、処理された複数枚の基板の各々の待ち時間のう

ちの最大値であってもよいし、平均値であってもよい。

- [0085] 予測モデル更新部 8 2 e は、予め定められた枚数の基板処理終了後（すなわち遷移後の状態 s_{t+1} が終端状態である場合）、処理枚数が多くかつ待ち時間が短いほど報酬が大きくなるように、動作結果取得部 8 2 d により取得された動作結果に基づいて報酬を計算し、当該報酬に基づいて予測モデル 8 5 を更新（たとえば、ニューラルネットワークにおける各ノードのパラメータ（重みや閾値など）を更新）する。
- [0086] 次に、このような構成からなる機械学習装置 8 0 による機械学習方法の一例について説明する。図 7 は、機械学習方法の一例を示すフローチャートである。
- [0087] 図 7 に示すように、まず、基板処理装置 1 0 にて 1 サイクルの処理（すなわち、あらかじめ定められた枚数ないしロットの処理）が開始されると、機械学習装置 8 0 の制御部 8 2 が、基板処理装置 1 0 の制御部 7 0 から処理開始通知を受信する（ステップ S 1 0）。
- [0088] そして、状態情報取得部 8 2 a が、基板処理装置 1 0 内における基板 W の位置および各ユニット 2 0、3 0、4 0 内に位置する基板 W の当該ユニット内での経過時間を含む状態情報を、基板処理装置 1 0 の制御部 7 0 から取得する（ステップ S 1 1）。
- [0089] 次に、行動選択部 8 2 b が、状態情報取得部 8 2 a により取得された状態情報を入力として、予測モデル 8 5 に基づいて、1 つの行動（すなわち、新たな基板 W をカセット 1 2 から取り出して第 1 処理ユニット 2 0 に搬送する行動と、新たな基板 W をカセット 1 2 から取り出して第 2 処理ユニット 2 0 に搬送する行動と、新たな基板 W をカセット 1 2 から取り出さない行動のうちのいずれか）を選択する（ステップ S 1 2）。
- [0090] そして、指示信号送信部 8 2 c が、行動選択部 8 2 b により選択された行動を行うように基板処理装置 1 0 の制御部 7 0 に指示信号を送信する（ステップ S 1 3）。基板処理装置 1 0 の制御部 7 0 が指示信号送信部 8 2 c から受信した指示信号に従って行動することにより、基板処理装置 1 0 内の状態

s_t は、次の状態 s_{t+1} に遷移する。

- [0091] 遷移後の状態 s_{t+1} が終端状態（予め定められた枚数の基板処理が終了した状態）ではなかった場合には（ステップ S 14 : NO）、ステップ S 11 から処理を繰り返す。この場合、予測モデル更新部 82 e は、状態情報取得部 82 a により取得される遷移後の状態 s_{t+1} の状態情報を予測モデル 85 の入力層に入力した場合に出力層から出力される各行動の価値のうちの最大の価値（Q 値）に基づいて予測モデル 85 を更新（たとえば、ニューラルネットワークにおける各ノードのパラメータ（重みや閾値など）を更新）してもよい。
- [0092] 予め定められた枚数の基板処理終了後（すなわち、遷移後の状態 s_{t+1} が終端状態である場合）には（ステップ S 14 : YES）、動作結果取得部 82 d が、単位時間あたりの処理枚数と、表面処理後の基板 W が洗浄ユニット 40 にて洗浄開始となるまでに待たされた待ち時間とを含む動作結果を、基板処理装置 10 の制御部 70 から取得する（ステップ S 15）。
- [0093] 次に、予測モデル更新部 82 e は、予め定められた枚数の基板処理終了後（すなわち遷移後の状態 s_{t+1} が終端状態である場合）、処理枚数が多くかつ待ち時間が短いほど報酬が大きくなるように、動作結果取得部 82 d により取得された動作結果に基づいて報酬を計算する（ステップ S 16）。
- [0094] そして、予測モデル更新部 82 e は、計算された報酬に基づいて予測モデル 85 を更新（たとえば、ニューラルネットワークにおける各ノードのパラメータ（重みや閾値など）を更新）する（ステップ S 17）。
- [0095] 機械学習装置 80 の制御部 82 は、あらかじめ定められた学習回数（たとえば 10000 回）に到達したか否かを判断し、当該学習回数に到達していない場合には（ステップ S 18 : NO）、ステップ S 10 から処理を繰り返す。他方、あらかじめ定められた学習回数に到達した場合には（ステップ S 18 : YES）、処理を終了する。これにより、学習済みの予測モデル 85（たとえば、チューニングされたニューラルネットワークシステム）が得られる。

[0096] 機械学習装置80により生成された学習済みの予測モデル85（たとえば、チューニングされたニューラルネットワークシステム）は、基板処理装置10の制御部70にインストールされて利用され得る。学習済みの予測モデル85がインストールされた基板処理装置10の制御部70は、基板処理装置10内における基板Wの位置および各ユニット20、30、40内に位置する基板の当該ユニット内での経過時間を含む状態情報を入力として、学習済みの予測モデル85に基づいて、新たな基板Wをカセット12から取り出すか否かおよび取り出す場合には第1処理ユニット20および第2処理ユニット30のどちらに搬送するかの行動を選択し、選択した行動を行うように搬送部50の動作を制御する。

[0097] 以上のような第1の実施形態によれば、機械学習装置80は、基板処理装置10内におけるその時その時の基板Wの位置および各ユニット20、30、40内に位置する基板Wの当該ユニット内での経過時間を含む状態情報に応じて、予測モデル85に基づいて、新たな基板Wをカセットから取り出すか否かおよび取り出す場合には第1処理ユニット20および第2処理ユニット30のどちらに搬送するかの行動を選択することを試行錯誤し、あらかじめ定められた枚数の基板処理終了後、単位時間あたりの処理枚数が多くかつ表面処理後の基板が洗浄開始となるまでに待たされた待ち時間が短くなるほど大きな報酬を獲得し、当該報酬に基づいて予測モデルを更新することを繰り返すことにより、予測モデル85の機械学習（強化学習）を行っている。そのため、このような機械学習装置80により生成された学習済みの予測モデル85を利用することにより、基板Wの搬送開始のタイミングおよびその搬送ルート、基板処理装置10内におけるその時その時の状態に応じて、（単位時間あたりの処理枚数が多くかつ待ち時間が短くなるように）適切に決定することが可能になる。

[0098] なお、上述した第1の実施形態に係る機械学習装置80は、基板処理装置10の実機に対して機械学習を行ったが、これに限定されず、基板処理装置10のシミュレータに対して機械学習を行ってもよいし、機械学習の初期に

は基板処理装置 10 のシミュレータに対して機械学習を行い、ある程度学習が進んだ後で、基板処理装置 10 の実機に対して機械学習を行ってもよい。

[0099] (第 2 の実施形態)

次に、第 2 の実施形態について説明する。基板の搬送、処理（研磨）および洗浄の工程を予め定められたタイムチャートに従って管理するスケジューラを使用した従来の制御方法では、研磨ユニットでの研磨時間が終点検出により決定されることで研磨時間にばらつきが存在することなどを理由として、平均研磨時間、平均搬送時間および平均洗浄時間に基づいて計算した時刻どおりに（許容時間なしで）制御を行うと、確実に遅れが生じてスループットが悪化する。そのため、装置内にて基板が多少滞留してしまうことを許容し、少し早めに目的箇所に着するように制御することで、遅れが生じないようにすることが行われる。この許容時間は、従来は人間が経験で調整しており、装置内におけるその時その時の状態に関わらず一律に決められていた。

[0100] 第 2 の実施形態に係る機械学習装置 180 は、基板処理装置 10 の制御部 70 が、カセット 12 から取り出される基板 W の順番と第 1 処理ユニット 20 および第 2 処理ユニット 30 のどちらに搬送するかとの対応関係が規定された搬送ルールに従って、第 1 処理ユニット 20 および第 2 処理ユニット 30 と洗浄ユニット 40 と搬送部 50 の動作を制御する場合（すなわち、カセット 12 から新たに取り出す基板 W を第 1 処理ユニット 20 および第 2 処理ユニット 30 のどちらに搬送するか搬送ルートが予め決められている場合）に、基板 W の搬送開始のタイミングを、基板処理装置 10 内におけるその時その時の状態に応じて、（単位時間あたりの処理枚数が多くなるように）適切に決定することを可能にできるものである。

[0101] 図 8 は、第 2 の実施形態に係る機械学習装置 180 の構成を示すブロック図である。機械学習装置 180 の少なくとも一部は、1 つのコンピュータまたは量子コンピューティングシステム、もしくは互いにネットワークを介して接続された複数のコンピュータまたは量子コンピューティングシステムに

よって構成されている。

- [0102] 図8に示すように、機械学習装置180は、通信部181と、制御部182と、記憶部183とを有している。各部181～183は、バスやネットワークを介して通信可能に接続されている。
- [0103] このうち通信部181は、基板処理装置10の制御部70に対する通信インターフェースである。通信部181は、基板処理装置10の制御部70に有線で接続されていてもよいし、無線で接続されていてもよい。
- [0104] 記憶部183は、たとえばフラッシュメモリなどの不揮発性データストレージである。記憶部183には、制御部182が取り扱う各種データが記憶される。
- [0105] 図8に示すように、制御部182は、状態情報取得部182aと、行動選択部182bと、指示信号送信部182cと、動作結果取得部182dと、予測モデル更新部182eとを有している。これらの各部は、機械学習装置180内のプロセッサが所定のプログラムを実行することにより実現されてもよいし、ハードウェアで実装されてもよい。
- [0106] 本実施の形態において、制御部182は、単位時間あたりの処理枚数が多く、かつ、表面処理後の基板が洗浄ユニット40にて洗浄開始となるまでに待たされる待ち時間が短くなるような基板の搬送開始のタイミングおよびその搬送ルートを、基板処理装置10内におけるその時その時の状態に応じた試行錯誤を繰り返すことで、強化学習するものである。強化学習のアルゴリズムは、特に限定されるものではないが、たとえばQ学習、SARSA法、方策勾配法、Actor-Critic法などが用いられ得る。
- [0107] 状態情報取得部182aは、基板処理装置10内における基板Wの位置および各ユニット20、30、40内に位置する基板Wの当該ユニット内での経過時間を含む状態情報を、基板処理装置10の制御部70から所定の時間間隔（たとえば0.1sごと）で繰り返し取得する。
- [0108] 状態情報取得部182aが基板処理装置10の制御部70から取得する状態情報は、第1処理ユニット20および第2処理ユニット30にて使用され

る消耗部材の使用時間をさらに含んでもよい。本件発明者が鋭意検討を重ねた結果、第1処理ユニット20および第2処理ユニット30での処理時間（たとえば終点検出により決定される研磨時間）は、第1処理ユニット20および第2処理ユニット30にて使用される消耗部材の使用時間と相関関係があることが見出された。したがって、後述する予測モデル185に入力される状態情報が、第1処理ユニット20および第2処理ユニット30にて使用される消耗部材の使用時間を含んでいる場合には、予測モデル185による予測精度をさらに向上させることができる。消耗部材は、たとえば、回転テーブル22b、24b、32b、34bに取り付けられた研磨パッド、トップリング22a、24a、32a、34aに取り付けられて基板Wの外周を支持するリテーナリング、トップリング22a、24a、32a、34aに取り付けられて基板Wの裏面を支持する弾性膜のうちの1つまたは2つ以上であってもよい。

[0109] 状態情報取得部182aが基板処理装置10の制御部70から取得する状態情報は、カセット12内に收容された基板Wに予め施されている処理のレシピ情報（たとえば図1Bに示す基板W表面の銅膜7の成膜条件）をさらに含んでもよい。本件発明者が鋭意検討を重ねた結果、第1処理ユニット20および第2処理ユニット30での処理時間（たとえば終点検出により決定される研磨時間）は、カセット12内に收容された基板Wに予め施されている処理のレシピ情報と相関関係があることが見出された。したがって、後述する予測モデル185に入力される状態情報が、カセット12内に收容された基板Wに予め施されている処理のレシピ情報を含んでいる場合には、予測モデル185による予測精度を向上させることができる。

[0110] 状態情報取得部182aが基板処理装置10の制御部70から取得する状態情報は、第1処理ユニット20および第2処理ユニット30の連続運転時間をさらに含んでもよい。本件発明者が鋭意検討を重ねた結果、第1処理ユニット20および第2処理ユニット30にて運転間隔が空くと水が滞留したりして一回洗い直すことによりコンデションが大きく変わることから、

第1処理ユニット20および第2処理ユニット30での処理時間（たとえば終点検出により決定される研磨時間）は、第1処理ユニット20および第2処理ユニット30の連続運転時間と相関関係があることが見出された。したがって、後述する予測モデル85に入力される状態情報が、第1処理ユニット20および第2処理ユニット30の連続運転時間を含んでいる場合には、予測モデル85による予測精度を向上させることができる。

[0111] 状態情報取得部182aが基板処理装置10の制御部70から取得する状態情報は、第1処理ユニット20および第2処理ユニット30での表面処理（研磨処理）のレシピ情報をさらに含んでもよい。本件発明者が鋭意検討を重ねた結果、第1処理ユニット20および第2処理ユニット30での処理時間（たとえば終点検出により決定される研磨時間）は、第1処理ユニット20および第2処理ユニット30での表面処理（研磨処理）のレシピ情報と相関関係があることが見出された。したがって、後述する予測モデル185に入力される状態情報が、第1処理ユニット20および第2処理ユニット30での表面処理（研磨処理）のレシピ情報を含んでいる場合には、予測モデル185による予測精度を向上させることができる。

[0112] 行動選択部182bは、ある状態 s_t において、新たな基板Wをカセット12から取り出すか否かの行動を行うことに対する価値（Q学習におけるQ値）を予測する予測モデル185（図9参照）を有している。

[0113] 図9は、予測モデル185の構成の一例を説明するための模式図である。図9に示す例では、予測モデル185は、ニューラルネットワークシステムであり、入力層と、入力層に接続された1または2以上の中間層と、中間層に接続され出力層とを有する階層型のニューラルネットワークまたは量子ニューラルネットワーク（QNN）を含んでいる。図9では、階層型のニューラルネットワークとして、フィードフォワードニューラルネットワークが図示されているが、畳み込みニューラルネットワーク（CNN）やリカレントニューラルネットワーク（RNN）など、様々なタイプのニューラルネットワークが使用され得る。予測モデル185は、中間層が2層以上に多層化さ

れたニューラルネットワーク、すなわちディープラーニング（深層学習）を含んでいてもよい。

[0114] 図9に示すように、予測モデル185は、状態情報取得部182aにより取得された状態情報が入力層に入力されると、新たな基板Wをカセット12から取り出すか否かおよび取り出す場合には第1処理ユニット20および第2処理ユニット30のどちらに搬送するかの行動を行うことに対する価値（Q学習におけるQ値）を予測して出力層から出力する。

[0115] 行動選択部182bは、複数の予測モデル185を有し、当該複数の予測モデル185による予測結果の組み合わせ（すなわちアンサンブル学習）に基づいて、各行動の価値（Q値）を推定して出力してもよい。

[0116] 行動選択部182bは、状態情報取得部182aにより取得された状態情報を入力として予測モデル185に基づいて1つの行動（すなわち、新たな基板Wをカセット12から取り出す行動と、新たな基板Wをカセット12から取り出さない行動のいずれか）を選択する。選択方法としては、たとえば、行動選択部182bは、予測モデル185により予測された各行動の価値（Q値）を比較して、最も価値（Q値）が高い行動を選択してもよいし（greedy法）、予め定められた確率 ε 以下でランダムに行動を選択し、それ以外では最も価値（Q値）が高い行動を選択してもよい（ ε -greedy法）。

[0117] 指示信号送信部182cは、行動選択部182bにより選択された行動を行うように基板処理装置10の制御部70に指示信号を送信する。基板処理装置10の制御部70が指示信号送信部182cから受信した指示信号に従って行動することにより、基板処理装置10内の状態 s_t は、次の状態 s_{t+1} に遷移する。

[0118] 予測モデル更新部182eは、遷移後の状態 s_{t+1} が終端状態（予め定められた枚数の基板処理が終了した状態）ではなかった場合には、状態情報取得部182aにより取得される遷移後の状態 s_{t+1} の状態情報を予測モデル185の入力層に入力した場合に出力層から出力される各行動の価値のうちの最

大の価値（Q値）に基づいて予測モデル185を更新（たとえば、ニューラルネットワークにおける各ノードのパラメータ（重みや閾値など）を更新）してもよい。

- [0119] 動作結果取得部182dは、予め定められた枚数の基板処理終了後（すなわち遷移後の状態 s_{t+1} が終端状態である場合）、単位時間あたりの処理枚数を含む動作結果を、基板処理装置10の制御部70から取得する。
- [0120] 予測モデル更新部182eは、予め定められた枚数の基板処理終了後（すなわち遷移後の状態 s_{t+1} が終端状態である場合）、処理枚数が多いほど報酬が大きくなるように、動作結果取得部182dにより取得された動作結果に基づいて報酬を計算し、当該報酬に基づいて予測モデル185を更新（たとえば、ニューラルネットワークにおける各ノードのパラメータ（重みや閾値など）を更新）する。
- [0121] 次に、このような構成からなる機械学習装置180による機械学習方法の一例について説明する。図10は、機械学習方法の一例を示すフローチャートである。
- [0122] 図10に示すように、まず、基板処理装置10にて1サイクルの処理（すなわち、あらかじめ定められた枚数ないしロットの処理）が開始されると、機械学習装置180の制御部182が、基板処理装置10の制御部70から処理開始通知を受信する（ステップS110）。
- [0123] そして、状態情報取得部182aが、基板処理装置10内における基板Wの位置および各ユニット20、30、40内に位置する基板Wの当該ユニット内での経過時間を含む状態情報を、基板処理装置10の制御部70から取得する（ステップS111）。
- [0124] 次に、行動選択部182bが、状態情報取得部182aにより取得された状態情報を入力として、予測モデル185に基づいて、1つの行動（すなわち、新たな基板Wをカセット12から取り出す行動と、新たな基板Wをカセット12から取り出さない行動のいずれか）を選択する（ステップS112）。

- [0125] そして、指示信号送信部 182c が、行動選択部 182b により選択された行動を行うように基板処理装置 10 の制御部 70 に指示信号を送信する（ステップ S 113）。基板処理装置 10 の制御部 70 が指示信号送信部 82c から受信した指示信号に従って行動することにより、基板処理装置 10 内の状態 s_t は、次の状態 s_{t+1} に遷移する。
- [0126] 遷移後の状態 s_{t+1} が終端状態（予め定められた枚数の基板処理が終了した状態）ではなかった場合には（ステップ S 114：NO）、ステップ S 111 から処理を繰り返す。この場合、予測モデル更新部 182e は、状態情報取得部 182a により取得される遷移後の状態 s_{t+1} の状態情報を予測モデル 185 の入力層に入力した場合に出力層から出力される各行動の価値のうちの最大の価値（Q 値）に基づいて予測モデル 185 を更新（たとえば、ニューラルネットワークにおける各ノードのパラメータ（重みや閾値など）を更新）してもよい。
- [0127] 予め定められた枚数の基板処理終了後（すなわち、遷移後の状態 s_{t+1} が終端状態である場合）には（ステップ S 114：YES）、動作結果取得部 182d が、単位時間あたりの処理枚数を含む動作結果を、基板処理装置 10 の制御部 70 から取得する（ステップ S 115）。
- [0128] 次に、予測モデル更新部 182e は、予め定められた枚数の基板処理終了後（すなわち遷移後の状態 s_{t+1} が終端状態である場合）、処理枚数が多くなるように、動作結果取得部 182d により取得された動作結果に基づいて報酬を計算する（ステップ S 116）。
- [0129] そして、予測モデル更新部 182e は、計算された報酬に基づいて予測モデル 185 を更新（たとえば、ニューラルネットワークにおける各ノードのパラメータ（重みや閾値など）を更新）する（ステップ S 117）。
- [0130] その後、機械学習装置 180 の制御部 182 は、あらかじめ定められた学習回数（たとえば 10000 回）に到達したか否かを判断し、当該学習回数に到達していない場合には（ステップ S 118：NO）、ステップ S 110 から処理を繰り返す。他方、あらかじめ定められた学習回数に到達した場合

には（ステップS 1 1 8：Y E S）、処理を終了する。これにより、学習済みの予測モデル1 8 5（たとえば、チューニングされたニューラルネットワークシステム）が得られる。

[0131] 機械学習装置1 8 0により生成された学習済みの予測モデル1 8 5（たとえば、チューニングされたニューラルネットワークシステム）は、基板処理装置1 0の制御部7 0にインストールされて利用され得る。学習済みの予測モデル1 8 5がインストールされた基板処理装置1 0の制御部7 0は、カセット1 2から取り出される基板Wの順番と第1 処理ユニット2 0および第2 処理ユニット3 0のどちらに搬送するかとの対応関係が規定された搬送ルールに従って、第1 処理ユニット2 0および第2 処理ユニット3 0と洗浄ユニット4 0と搬送部5 0の動作を制御するものであって、基板処理装置1 0内における基板Wの位置および各ユニット2 0、3 0、4 0内に位置する基板の当該ユニット内での経過時間を含む状態情報を入力として、学習済みの予測モデル1 8 5に基づいて、新たな基板Wをカセット1 2から取り出すか否かの行動を選択し、選択した行動を行うように搬送部5 0の動作を制御する。

[0132] 以上のような第2 の実施形態によれば、機械学習装置1 8 0は、基板処理装置1 0内におけるその時その時の基板Wの位置および各ユニット2 0、3 0、4 0内に位置する基板Wの当該ユニット内での経過時間を含む状態情報に応じて、予測モデル1 8 5に基づいて、新たな基板Wをカセットから取り出すか否かの行動を選択することを試行錯誤し、あらかじめ定められた枚数の基板処理終了後、単位時間あたりの処理枚数が多くなるほど大きな報酬を獲得し、当該報酬に基づいて予測モデルを更新することを繰り返すことにより、予測モデル1 8 5の機械学習（強化学習）を行っている。そのため、このような機械学習装置1 8 0により生成された学習済みの予測モデル1 8 5を利用することにより、基板Wの搬送開始のタイミングを、基板処理装置1 0内におけるその時その時の状態に応じて、（単位時間あたりの処理枚数が多くなるように）適切に決定することが可能になる。

[0133] なお、上述した第2の実施形態に係る機械学習装置180は、基板処理装置10の実機に対して機械学習を行ったが、これに限定されず、基板処理装置10のシミュレータに対して機械学習を行ってもよいし、機械学習の初期には基板処理装置10のシミュレータに対して機械学習を行い、ある程度学習が進んだ後で、基板処理装置10の実機に対して機械学習を行ってもよい。

[0134] (第3の実施形態)

次に、第3の実施形態について説明する。基板の搬送、処理（研磨）および洗浄の工程を予め定められたタイムチャートに従って管理するスケジューラを使用した従来の制御方法では、同じレシピであっても研磨時間と消耗部材の使用時間との間に相関があることなどを理由として、平均研磨時間、平均搬送時間および平均洗浄時間に基づいて計算された時刻どおりに制御を行うと、遅れが生じてスループットが悪化することがある。

[0135] 第3の実施形態に係る機械学習装置280は、基板処理装置10の制御部70が、カセット12から取り出される基板Wの順番と第1処理ユニット20および第2処理ユニット30のどちらに搬送するか、およびその搬送開始時刻との対応関係が規定された搬送ルールに従って、第1処理ユニット20および第2処理ユニット30と洗浄ユニット40と搬送部50の動作を制御する場合（すなわち、カセット12から新たに基板Wを取り出すタイミングと、取り出した基板Wを第1処理ユニット20および第2処理ユニット30のどちらに搬送するかの搬送ルートとが予め決められている場合）に、処理ユニットでの表面処理（研磨）のレシピ情報と、基板情報だけでなく、処理ユニット内にて使用される消耗部材の使用時間と、処理ユニットの連続運転時間をも考慮して、処理ユニットにおける表面処理時間を精度よく予測することを可能にでき、これにより、タイムチャート（搬送ルール）の作成時に、当該予測された表面処理時間に基づいて、基板の搬送開始のタイミングを精度よく決定することを可能にできるものである。

[0136] 図11は、第3の実施形態に係る機械学習装置280の構成を示すブロッ

ク図である。機械学習装置280の少なくとも一部は、1つのコンピュータまたは量子コンピューティングシステム、もしくは互いにネットワークを介して接続された複数のコンピュータまたは量子コンピューティングシステムによって構成されている。

[0137] 図11に示すように、機械学習装置280は、通信部281と、制御部282と、記憶部283とを有している。各部281～283は、バスやネットワークを介して通信可能に接続されている。

[0138] このうち通信部281は、基板処理装置10の制御部70に対する通信インターフェースである。通信部281は、基板処理装置10の制御部70に有線で接続されていてもよいし、無線で接続されていてもよい。

[0139] 記憶部283は、たとえばフラッシュメモリなどの不揮発性データストレージである。記憶部283には、制御部282が取り扱う各種データが記憶される。

[0140] 図11に示すように、制御部282は、入力情報取得部282aと、予測部282bと、実表面時間取得部282cと、予測モデル更新部282dとを有している。これらの各部は、機械学習装置280内のプロセッサが所定のプログラムを実行することにより実現されてもよいし、ハードウェアで実装されてもよい。

[0141] 本実施の形態において、制御部282は、基板Wを表面処理する第1処理ユニット20（または第2処理ユニット30）における表面処理のレシピ情報と、基板情報と、第1処理ユニット20（または第2処理ユニット30）内にて使用される消耗部材の使用時間と、第1処理ユニット20（または第2処理ユニット30）の連続運転時間と、第1処理ユニット20（または第2処理ユニット30）における実際の表面処理時間との関係性を機械学習（教師あり学習）するものである。

[0142] 入力情報取得部282aは、第1処理ユニット20（または第2処理ユニット30）での表面処理のレシピ情報と、基板情報（たとえば図1Bに示す基板W表面の銅膜7の成膜条件）と、第1処理ユニット20（または第2処

理ユニット 30) 内にて使用される消耗部材の使用時間と、第 1 処理ユニット 20 (または第 2 処理ユニット 30) の連続運転時間とを、基板処理装置 10 の制御部 70 から入力情報として取得する。消耗部材は、たとえば、回転テーブル 22 b、24 b、32 b、34 b に取り付けられた研磨パッド、トップリング 22 a、24 a、32 a、34 a に取り付けられて基板 W の外周を支持するリテーナリング、トップリング 22 a、24 a、32 a、34 a に取り付けられて基板 W の裏面を支持する弾性膜のうちの 1 つまたは 2 つ以上であってもよい。

[0143] 本件発明者が鋭意検討を重ねた結果、第 1 処理ユニット 20 (または第 2 処理ユニット 30) での処理時間 (たとえば終点検出により決定される研磨時間) は、第 1 処理ユニット 20 (または第 2 処理ユニット 30) にて使用される消耗部材の使用時間と相関関係があることが見出された。また、本件発明者が鋭意検討を重ねた結果、第 1 処理ユニット 20 (または第 2 処理ユニット 30) にて運転間隔が空くと水が滞留したりして一回洗い直すことによりコンデションが大きく変わることから、第 1 処理ユニット 20 (または第 2 処理ユニット 30) での処理時間 (たとえば終点検出により決定される研磨時間) は、第 1 処理ユニット 20 (または第 2 処理ユニット 30) の連続運転時間と相関関係があることが見出された。したがって、後述する予測モデル 285 に入力される入力情報が、消耗部材の使用時間と当該処理ユニットの連続運転時間とを含んでいることにより、予測モデル 285 による予測精度を顕著に向上させることが可能である。

[0144] 予測部 282 b は、第 1 処理ユニット 20 (または第 2 処理ユニット 30) での表面処理のレシピ情報と、基板情報と、第 1 処理ユニット 20 (または第 2 処理ユニット 30) 内にて使用される消耗部材の使用時間と、第 1 処理ユニット 20 (または第 2 処理ユニット 30) の連続運転時間とに基づいて、第 1 処理ユニット 20 (または第 2 処理ユニット 30) における表面処理時間を予測する予測モデル 285 (図 12 参照) を有している。

[0145] 図 12 は、予測モデル 285 の構成の一例を説明するための模式図である

。図 1 2 に示す例では、予測モデル 2 8 5 は、ニューラルネットワークシステムであり、入力層と、入力層に接続された 1 または 2 以上の中間層と、中間層に接続され出力層とを有する階層型のニューラルネットワークまたは量子ニューラルネットワーク (QNN) を含んでいる。図 1 2 では、階層型のニューラルネットワークとして、フィードフォワードニューラルネットワークが図示されているが、畳み込みニューラルネットワーク (CNN) やリカレントニューラルネットワーク (RNN) など、様々なタイプのニューラルネットワークが使用され得る。予測モデル 2 8 5 は、中間層が 2 層以上に多層化されたニューラルネットワーク、すなわちディープラーニング (深層学習) を含んでいてもよい。

[0146] 図 1 2 に示すように、予測モデル 2 8 5 は、入力情報取得部 2 8 2 a により取得された入力情報 (すなわち、第 1 処理ユニット 2 0 (または第 2 処理ユニット 3 0) での表面処理のレシピ情報と、基板情報と、第 1 処理ユニット 2 0 (または第 2 処理ユニット 3 0) 内にて使用される消耗部材の使用時間と、第 1 処理ユニット 2 0 (または第 2 処理ユニット 3 0) の連続運転時間) とが入力層に入力されると、第 1 処理ユニット 2 0 (または第 2 処理ユニット 3 0) における表面処理時間を予測して出力層から出力する。

[0147] 実表面処理時間取得部 2 8 2 c は、第 1 処理ユニット 2 0 (または第 2 処理ユニット 3 0) における実際の表面処理時間を、基板処理装置 1 0 の制御部 7 0 から取得する。

[0148] 予測モデル更新部 2 8 2 d は、実表面処理時間取得部 2 8 2 c により取得された実際の表面処理時間と、予測部 2 9 2 b により予測された表面処理時間とを比較し、その誤差に応じて予測モデル 2 8 5 を更新 (たとえば、ニューラルネットワークにおける各ノードのパラメータ (重みや閾値など) を更新) する。

[0149] 次に、このような構成からなる機械学習装置 2 8 0 による機械学習方法の一例について説明する。図 1 3 は、機械学習方法の一例を示すフローチャートである。

- [0150] 図13に示すように、まず、入力情報取得部282aが、第1処理ユニット20（または第2処理ユニット30）での表面処理のレシピ情報と、基板情報（たとえば図1Bに示す基板W表面の銅膜7の成膜条件）と、第1処理ユニット20（または第2処理ユニット30）内にて使用される消耗部材の使用時間と、第1処理ユニット20（または第2処理ユニット30）の連続運転時間とを、基板処理装置10の制御部70から入力情報として取得する（ステップS211）。
- [0151] 次に、予測部282bが、入力情報取得部282aにより取得された入力情報（すなわち、第1処理ユニット20（または第2処理ユニット30）での表面処理のレシピ情報と、基板情報と、第1処理ユニット20（または第2処理ユニット30）内にて使用される消耗部材の使用時間と、第1処理ユニット20（または第2処理ユニット30）の連続運転時間）を入力として、予測モデル285に基づいて、第1処理ユニット20（または第2処理ユニット30）における表面処理時間を予測して出力する（ステップS212）。
- [0152] 次いで、実表面処理時間取得部282cが、第1処理ユニット20（または第2処理ユニット30）における実際の表面処理時間を、基板処理装置10の制御部70から取得する（ステップS213）。
- [0153] そして、予測モデル更新部282dが、実表面処理時間取得部282cにより取得された実際の表面処理時間と、予測部292bにより予測された表面処理時間とを比較し、その誤差に応じて予測モデル285を更新（たとえば、ニューラルネットワークにおける各ノードのパラメータ（重みや閾値など）を更新）する（ステップS214）。
- [0154] その後、機械学習装置280の制御部282は、あらかじめ定められた学習回数（たとえば10000回）に到達したか否かを判断し、当該学習回数に到達していない場合には（ステップS215：NO）、ステップS211から処理を繰り返す。他方、あらかじめ定められた学習回数に到達した場合には（ステップS215：YES）、処理を終了する。これにより、学習済

みの予測モデル285（たとえば、チューニングされたニューラルネットワークシステム）が得られる。

[0155] 機械学習装置280により生成された学習済みの予測モデル285（たとえば、チューニングされたニューラルネットワークシステム）は、基板処理装置10の制御部70にインストールされて利用され得る。学習済みの予測モデル285がインストールされた基板処理装置10の制御部70は、カセット12から取り出される基板Wの順番と第1処理ユニット20および第2処理ユニット30のどちらに搬送するか、およびその搬送開始時刻との対応関係が規定された搬送ルールに従って、第1処理ユニット20および第2処理ユニット30と洗浄ユニット40と搬送部50の動作を制御するものであって、第1処理ユニット20（または第2処理ユニット30）での表面処理のレシピ情報と、基板情報（たとえば図1Bに示す基板W表面の銅膜7の成膜条件）と、第1処理ユニット20（または第2処理ユニット30）内にて使用される消耗部材の使用時間と、第1処理ユニット20（または第2処理ユニット30）の連続運転時間とを入力として、学習済みの予測モデル285に基づいて、第1処理ユニット20（または第2処理ユニット30）における表面処理時間を予測し、タイムチャート（搬送ルール）の作成時に、当該予測された表面処理時間に基づいて、基板の搬送開始のタイミングを決定する。なお、タイムチャート作成時に、予測表面処理時間に基づいて基板の搬送開始のタイミングを決定する具体的な手法としては、たとえば、特許第5023146号公報にて提案された手法を利用することができる。

[0156] 以上のような第3の実施の形態によれば、機械学習装置280は、第1処理ユニット20（または第2処理ユニット30）での表面処理のレシピ情報と、基板情報と、第1処理ユニット20（または第2処理ユニット30）内にて使用される消耗部材の使用時間と、第1処理ユニット20（または第2処理ユニット30）の連続運転時間と、第1処理ユニット20（または第2処理ユニット30）における実際の表面処理時間との対応関係を教師データとして、予測モデル285の機械学習（教師あり学習）を行っている。その

ため、このような機械学習装置 280 により生成された学習済みの予測モデル 285 を利用することにより、第 1 処理ユニット 20（または第 2 処理ユニット 30）での表面処理のレシピ情報と、基板情報だけでなく、第 1 処理ユニット 20（または第 2 処理ユニット 30）内にて使用される消耗部材の使用時間と、第 1 処理ユニット 20（または第 2 処理ユニット 30）の連続運転時間をも考慮して、第 1 処理ユニット 20（または第 2 処理ユニット 30）における表面処理時間を精度よく予測することが可能となり、これにより、タイムチャートの作成時に、当該予測された表面処理時間に基づいて、基板の搬送開始のタイミングを精度よく決定することが可能になる。

[0157] なお、上述した実施の形態に係る機械学習装置 80、180、280 は、1 つのコンピュータまたは量子コンピューティングシステム、もしくは互いにネットワークを介して接続された複数のコンピュータまたは量子コンピューティングシステムによって構成され得るが、1 または複数のコンピュータまたは量子コンピューティングシステムに機械学習装置 80、180、280 を実現させるためのプログラム及び当該プログラムを非一時的（*non-transitory*）に記録したコンピュータ読取可能な記録媒体も、本件の保護対象である。

[0158] 以上、実施の形態および変形例を例示により説明したが、本技術の範囲はこれらに限定されるものではなく、請求項に記載された範囲内において目的に応じて変更・変形することが可能である。また、各実施の形態および変形例は、処理内容を矛盾させない範囲で適宜組み合わせることが可能である。