



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0121568
(43) 공개일자 2022년09월01일

(51) 국제특허분류(Int. Cl.)
G05B 19/418 (2006.01) G05B 23/02 (2006.01)
G06N 20/00 (2019.01)
(52) CPC특허분류
G05B 19/41865 (2013.01)
G05B 19/4183 (2013.01)
(21) 출원번호 10-2021-0025807
(22) 출원일자 2021년02월25일
심사청구일자 2021년02월25일

(71) 출원인
창원대학교 산학협력단
경상남도 창원시 의창구 창원대로 20(퇴촌동)
(72) 발명자
김병욱
경상남도 창원시 성산구 토월로27번길 18, 105동 1404호 (상남동, 꿈에그린)
김성준
부산광역시 기장군 정관읍 정관4로 24, 103동 203호 (정관동일스위트)
(74) 대리인
김종선, 이형석

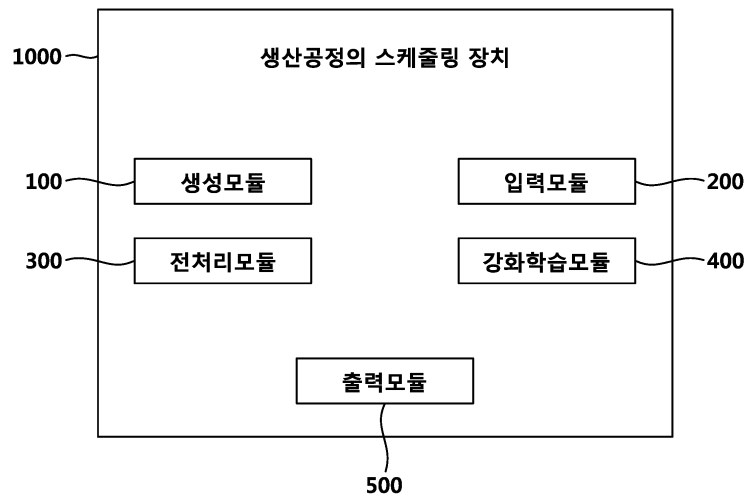
전체 청구항 수 : 총 12 항

(54) 발명의 명칭 강화학습을 이용한 생산공정의 스케줄링 장치 및 방법

(57) 요약

본 발명은 강화학습을 이용한 생산공정의 스케줄링 장치 및 방법에 관한 것이다. 본 발명에 따른 강화학습을 이용한 생산공정의 스케줄링 장치는 작업에 대한 일정을 추정함수로 생성하는 생성모듈, 상기 작업에 대한 파라미터를 입력하는 입력모듈, 상기 추정함수 및 파라미터를 전처리하는 전처리모듈, 상기 전처리된 추정함수 및 파라미터를 이용하여 추정함수를 학습하는 강화학습모듈 그리고 상기 강화학습된 추정함수를 통해 작업순열을 출력하는 출력모듈을 포함하여 구성된다.

대표도 - 도1



(52) CPC특허분류

G05B 19/41885 (2013.01)

G05B 23/02 (2013.01)

G06N 20/00 (2021.08)

이 발명을 지원한 국가연구개발사업

과제고유번호	1345323362
과제번호	LINCPLUS-2020-41
부처명	교육부
과제관리(전문)기관명	한국연구재단
연구사업명	산학연협력고도화지원(R&D)
연구과제명	사회맞춤형산학협력선도대학(LINC+)육성(0.5)
기 여 율	1/1
과제수행기관명	창원대학교 산학협력단
연구기간	2020.07.01 ~ 2020.12.31

명세서

청구범위

청구항 1

작업에 대한 일정을 추정함수로 생성하는 생성모듈;
 상기 작업에 대한 파라미터를 입력하는 입력모듈;
 상기 추정함수 및 파라미터를 전처리하는 전처리모듈;
 상기 전처리된 추정함수 및 파라미터를 이용하여 추정함수를 학습하는 강화학습모듈; 및
 상기 강화학습된 추정함수를 통해 작업순열을 출력하는 출력모듈을 포함하는 강화학습을 이용한 생산공정의 스케줄링 장치.

청구항 2

제1항에 있어서,
 상기 강화학습모듈은,
 상기 추정함수의 학습을 기설정된 횟수만큼 반복하는 반복부;
 상기 상기 추정함수의 반복 횟수여부를 탐색하는 탐색부;
 상기 탐색부의 탐색 결과에 따라 상기 추정함수의 입실론 계수와 무작위 난수를 비교하는 비교부;
 상기 비교결과에 따라 작업순열을 결정하는 결정부; 및
 상기 작업순열을 이용하여 상기 추정함수를 갱신하는 갱신부를 포함하는 강화학습을 이용한 생산공정의 스케줄링 장치.

청구항 3

제1항에 있어서,
 상기 강화학습모듈은,
 Double-Q learning 기법 또는 Dueling Architecture 기법을 활용하여 학습하는 것을 특징으로 하는 강화학습을 이용한 생산공정의 스케줄링 장치.

청구항 4

제2항에 있어서,
 상기 갱신부의 갱신시점을 판별하는 판별부를 더 포함하는 강화학습을 이용한 생산공정의 스케줄링 장치.

청구항 5

제2항에 있어서,
 상기 결정부는,
 상기 추정함수의 입실론 계수가 상기 무작위 난수보다 클 경우 무작위로 작업순열을 결정하고,

상기 추정함수의 입실론 계수가 무작위 난수보다 작거나 같을 경우 상기 추정함수에 기초하여 작업순열을 결정하는 것을 특징으로 하는 강화학습을 이용한 생산공정의 스케줄링 장치.

청구항 6

제1항에 있어서,

상기 출력모듈은,

상기 강화학습된 추정함수의 가중치 정보에 따라 상기 작업순열을 출력하는 것을 특징으로 하는 강화학습을 이용한 생산공정의 스케줄링 장치.

청구항 7

제1항에 있어서,

상기 가중치 정보는,

상태함수, 행동함수, 이점함수 간의 연산을 통한 가중치를 이용하는 것을 특징으로 하는 강화학습을 이용한 생산공정의 스케줄링 장치.

청구항 8

강화학습을 이용한 생산공정의 스케줄링 방법에 있어서,

작업에 대한 일정을 추정함수로 생성하는 생성단계;

상기 작업에 대한 파라미터를 입력하는 입력단계;

상기 추정함수 및 파라미터를 전처리하는 전처리단계;

상기 전처리된 추정함수 및 파라미터를 이용하여 추정함수를 학습하는 강화학습단계; 및

상기 강화학습된 추정함수를 통해 작업순열을 출력하는 출력단계를 포함하는 강화학습을 이용한 생산공정의 스케줄링 방법.

청구항 9

제8항에 있어서,

상기 강화학습단계는,

상기 추정함수의 학습을 기설정된 횟수만큼 반복하는 반복단계;

상기 상기 추정함수의 반복 횟수여부를 탐색하는 탐색단계;

상기 탐색 결과에 따라 상기 추정함수의 입실론 계수와 무작위 난수를 비교하는 비교단계;

상기 비교결과에 따라 작업순열을 결정하는 결정단계; 및

상기 작업순열을 이용하여 상기 추정함수를 갱신하는 갱신단계를 포함하는 강화학습을 이용한 생산공정의 스케줄링 방법.

청구항 10

제9항에 있어서,

상기 갱신단계의 갱신시점을 판별하는 판별단계를 더 포함하는 강화학습을 이용한 생산공정의 스케줄링 방법.

청구항 11

제8항에 있어서,

상기 결정단계는,

상기 추정함수의 입실론 계수가 상기 무작위 난수보다 클 경우 무작위로 작업순열을 결정하고,

상기 추정함수의 입실론 계수가 무작위 난수보다 작거나 같을 경우 상기 추정함수에 기초하여 작업순열을 결정하는 것을 특징으로 하는 강화학습을 이용한 생산공정의 스케줄링 방법.

청구항 12

제8항에 있어서,

상기 출력단계는,

상태함수, 행동함수, 이점함수를 연산한 가중치를 기초로 한 추정함수의 가중치 정보에 따라 상기 작업순열을 출력하는 것을 특징으로 하는 강화학습을 이용한 생산공정의 스케줄링 방법.

발명의 설명

기술 분야

[0001] 본 발명은 강화학습을 이용한 생산공정의 스케줄링 장치 및 방법에 관한 것으로, 보다 상세하게는 서로 다른 작업 시간을 갖는 기계와 다수의 작업에 대해 최종 공정시간을 최소화하도록 강화학습을 이용하여 생산공정의 작업순열을 스케줄링하는 장치 및 방법에 관한 것이다.

배경 기술

[0002] 산업 발전에 따라 소품종 대량생산 설비, 다품종 소량생산 설비 등의 제조공정에서 설비는 성과목표가 향상된 기술, 생산의 효율성이 증대된 기술을 요구하고 있다. 이러한 요구에 따라 흐름생산 공정 또는 개별 작업 생산공정의 작업 일정을 최적화하는 생산일정 연구가 활발히 진행되고 있다.

[0003] 생산일정 연구에서 NP-hard의 문제는 일련의 m개의 기계에서 n개의 작업이 처리되도록 흐름생산 공정을 적용할 수 있다. 예를 들어, 흐름생산 공정은 전자, 종이 및 섬유 산업, 콘크리트 생산 산업, 사진 필름 제조 산업과 같은 일반적인 제조 분야 또는 인터넷 서비스 시스템, 컨테이너 처리 시스템, 토목 시스템과 같은 비 제조 분야에서 활용할 수 있다.

[0004] 종래의 흐름생산 공정 기술의 문제점을 해결하기 위해 단일 또는 복수의 기계에서 최적의 작업 시간을 도출하는 분기 한정법 알고리즘, NEH 휴리스틱 알고리즘, α -탐욕 선택에 기반으로 한 휴리스틱 알고리즘, 확률 분포를 통한 선택 알고리즘이 개발되었다. 이와 같은 알고리즘은 소규모 흐름생산에 대한 최적의 생산 일정을 도출할 수 있다. 또한 Simulated Annealing(SA), Tabu Search(TS), Genetic Algorithm(GA)와 같은 메타 휴리스틱 접근 알고리즘, 흐름 상점 스케줄링을 위한 하이브리드 유전 알고리즘(hybrid genetic algorithm, HGA), 전체 공정 시간을 최소화하는 변형 SA 알고리즘은 초기의 작업 일정 결과를 기초하여 작업순열을 직접적으로 탐색할 수 있다.

[0005] 다만, 상기한 알고리즘은 계산 복잡도 및 계산 비용이 매우 높아 비효율적이고 산업분야에 적용하기 어렵다. 또한 복잡한 품종 생산 작업에서는 최적의 작업 순열을 도출하는 방향의 알고리즘을 수립시키기 어렵다는 문제점이 있다.

[0006] 이러한 문제점을 해결하기 위해 생산일정 기술은 유한 시간 내에 최적의 작업 일정을 도출하여 기계의 유휴 시간을 최소화하도록 일정계획 정책 및 추정함수의 정책을 탐색하는 강화학습 기반의 스케줄링 기술이 요구된다.

발명의 내용

해결하려는 과제

- [0007] 본 발명은 전술한 종래 기술의 문제점을 해결하기 위한 것으로서, 본 발명은 작업일정에 대한 추정함수를 학습 및 탐색하여 최적의 작업일정을 출력하는 생산공정의 스케줄링 장치 및 방법을 제공하는 것을 목적으로 한다.
- [0008] 본 발명은 Double-Q learning 기법으로 작업순열을 추정함수에 갱신하여 추정함수를 학습 및 탐색하는 생산공정의 스케줄링 장치 및 방법을 제공하는 것을 목적으로 한다.
- [0009] 본 발명의 목적은 이상에서 언급한 것으로 제한되지 않으며, 언급되지 않은 또 다른 목적들은 아래의 기재로부터 본 발명이 속하는 기술 분야의 통상의 지식을 가진 자에게 명확히 이해될 수 있을 것이다.

과제의 해결 수단

- [0010] 이와 같은 목적을 달성하기 위한 본 발명은, 작업에 대한 일정을 추정함수로 생성하는 생성모듈, 상기 작업에 대한 파라미터를 입력하는 입력모듈, 상기 추정함수 및 파라미터를 전처리하는 전처리모듈, 상기 전처리된 추정함수 및 파라미터를 이용하여 추정함수를 학습하는 강화학습모듈 그리고 상기 강화학습된 추정함수를 통해 작업순열을 출력하는 출력모듈을 포함하는 강화학습을 이용한 생산공정의 스케줄링 장치를 제공한다.
- [0011] 상기 강화학습모듈은, 상기 추정함수의 학습을 기설정된 횟수만큼 반복하는 반복부, 상기 추정함수의 반복 횟수 여부를 탐색하는 탐색부, 상기 탐색부의 탐색 결과에 따라 상기 추정함수의 입실론 계수와 무작위 난수를 비교하는 비교부, 상기 비교결과에 따라 작업순열을 결정하는 결정부 및 상기 작업순열을 상기 반복부의 추정함수에 갱신하는 갱신부를 포함한다.
- [0012] 상기 강화학습모듈은, Double-Q learning 기법 또는 Dueling Architecture 기법을 활용하여 학습하는 것을 특징으로 한다.
- [0013] 상기 갱신부의 갱신시점을 판별하는 판별부를 더 포함한다.
- [0014] 상기 결정부는, 상기 추정함수의 입실론 계수가 상기 무작위 난수보다 클 경우 무작위로 작업순열을 결정하고, 상기 추정함수의 입실론 계수가 무작위 난수보다 작거나 같을 경우 상기 추정함수에 기초하여 작업순열을 결정하는 것을 특징으로 한다.
- [0015] 상기 출력모듈은, 상기 강화학습된 추정함수의 가중치 정보에 따라 상기 작업순열을 출력하는 것을 특징으로 한다.
- [0016] 상기 가중치 정보는, 상태함수, 행동함수, 이점함수에 대해 연산한 가중치를 이용하는 것을 특징으로 한다.
- [0017] 본 발명의 다른 특징에 따르면, 강화학습을 이용한 생산공정의 스케줄링 방법에 있어서, 작업에 대한 일정을 추정함수로 생성하는 생성단계, 상기 작업에 대한 파라미터를 입력하는 입력단계, 상기 추정함수 및 파라미터를 전처리하는 전처리단계, 상기 전처리된 추정함수 및 파라미터를 이용하여 추정함수를 학습하는 강화학습단계, 그리고 상기 강화학습된 추정함수를 통해 작업순열을 출력하는 출력단계를 포함하는 강화학습을 이용한 생산공정의 스케줄링 방법을 제공한다.
- [0018] 상기 강화학습단계는, 상기 추정함수의 학습을 기설정된 횟수만큼 반복하는 반복단계, 상기 추정함수의 반복 횟수 여부를 탐색하는 탐색단계, 상기 추정함수의 입실론 계수와 무작위 난수를 비교하는 비교단계, 상기 비교결과에 따라 작업순열을 결정하는 결정단계 및 상기 작업순열을 상기 반복부의 추정함수에 갱신하는 갱신단계를 포함한다.
- [0019] 상기 갱신단계의 갱신시점을 판별하는 판별단계를 더 포함한다.
- [0020] 상기 결정단계는, 상기 추정함수의 입실론 계수가 상기 무작위 난수보다 클 경우 무작위로 작업순열을 결정하고, 상기 추정함수의 입실론 계수가 무작위 난수보다 작거나 같을 경우 상기 추정함수에 기초하여 작업순열을 결정하는 것을 특징으로 한다.
- [0021] 상기 출력단계는, 상태함수, 행동함수, 이점함수를 연산한 가중치를 기초로 한 추정함수의 가중치 정보에 따라 상기 작업순열을 출력하는 것을 특징으로 한다.

발명의 효과

- [0022] 이상과 같은 본 발명의 강화학습을 이용한 생산공정의 스케줄링 장치 및 방법에 따르면, 서로 다른 처리시간을 갖는 작업일정에 대한 추정함수를 학습 및 탐색하여 최적의 작업일정을 출력함으로써, 기계의 유휴시간을 최소화하여 기계시스템의 생산 효율성을 높일 수 있다.
- [0023] 본 발명은 Double-Q learning 기법으로 작업순열을 추정함수에 갱신하여 추정함수의 학습 및 탐색을 수행함으로써, 특정 상태에서 기계의 행동이 지나치게 과대평가되는 과적합 문제를 방지할 수 있는 기대효과를 가진다.
- [0024] 또한, 본 발명은 추정함수의 가중치 정보에 기초하여 강화 학습함으로써, 분산을 낮추고 수렴속도를 증가시킬 수 있다. 또한, 실제 산업현장에서 유동적으로 변경되는 작업순열 및 비용을 실시간으로 도출할 수 있어 실제 산업현장에서 적용할 수 있다.
- [0025] 본 발명에서 얻을 수 있는 효과는 이상에서 언급한 효과들로 제한되지 않으며, 언급하지 않은 또 다른 효과들은 아래의 기재로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

- [0026] 도 1은 본 발명에 따른 강화학습을 이용한 생산공정의 스케줄링 장치를 설명하기 위한 블록도
- 도 2는 도 1의 강화학습모듈을 설명하기 위한 블록도
- 도 3은 본 발명에 따른 강화학습을 이용한 생산공정의 스케줄링 방법을 설명하기 위한 흐름도
- 도 4는 도 3의 강화학습 단계를 설명하기 위한 흐름도

발명을 실시하기 위한 구체적인 내용

- [0027] 본 발명의 목적 및 효과, 그리고 그것들을 달성하기 위한 기술적 구성들은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시 예들을 참조하면 명확해질 것이다. 본 발명을 설명함에 있어서 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략할 것이다.
- [0028] 그리고 후술되는 용어들은 본 발명에서의 기능을 고려하여 정의된 용어들로서 이는 사용자, 운용자의 의도 또는 관례 등에 따라 달라질 수 있다.
- [0029] 그러나 본 발명은 이하에서 개시되는 실시 예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있다. 단지 본 실시 예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 그러므로 그 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다.
- [0030] 이와 같은 본 발명은 비록 한정된 실시 예와 도면에 의해 설명되나, 본 발명은 이것에 의해 한정되지 않으며, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 본 발명의 기술 사상과 아래에 기재될 청구범위의 균등 범위 내에서 다양한 수정 및 변형이 가능함은 물론이다.
- [0031] 이하에서는 도면에 도시한 실시 예에 기초하면서 본 발명에 대하여 더욱 상세하게 설명하기로 한다. 그러나, 본원이 이러한 실시 예와 도면에 제한되는 것은 아니다.
- [0032] 이하, 첨부된 도면을 참조하여 본 발명의 구성에 대하여 상세히 설명한다.
- [0033] 도 1은 본 발명에 따른 강화학습을 이용한 생산공정의 스케줄링 장치(1000)를 설명하기 위한 블록도이다.
- [0034] 도 1에 도시된 바와 같이, 본 발명에 따른 강화학습을 이용한 생산공정의 스케줄링 장치(1000)는 작업에 대한 일정을 추정함수로 생성하는 생성모듈(100), 상기 작업에 대한 파라미터를 입력하는 입력모듈(200), 상기 추정함수 및 파라미터를 전처리하는 전처리모듈(300), 상기 전처리된 추정함수 및 파라미터를 이용하여 추정함수를 학습하는 강화학습모듈(400) 및 상기 강화학습된 추정함수를 통해 작업순열을 출력하는 출력모듈(500)을 포함한다.
- [0035] 생성모듈(100)은 기계 작업에 대한 일정을 추정함수로 생성한다. 여기서 추정함수는 가중치 정보를 포함하고 있는 상태함수, 행동함수, 이점함수일 수 있으며, 추정함수의 가중치 정보는 각 작업들의 가치에 대해 상대적으로

부여되는 중요도이며 상태함수, 행동함수, 이점함수 간의 연산을 통해 획득한 값일 수 있다.

[0036] 또한, 추정함수는 듀얼 구조를 통해 현재상태에 따라 어떤 행동을 수행하는지를 판단하는 이점함수(advantage function)일 수 도 있다.

[0037] 상기 이점함수는 하기 (수학식 1)과 같으며, 상기 (수학식 1)의 이점함수를 이용하여 하기 (수학식 2)와 같은 추정함수를 생성할 수 있다.

수학식 1

$$A^{\pi}(s,a) = Q^{\pi}(s,a) - V^{\pi}(s)$$

[0038]

[0039] 여기서, s는 현재 상태의 작업 또는 작업들, a는 다음 선택할 작업 또는 행동을 나타내며, S함수는 현재 상태의 작업(s)에 대해 다음 선택할 작업(a)들의 가중치 정보 함수이며, V함수는 작업들(s)에 대한 가중치 정보 함수를 의미한다.

수학식 2

$$Q(s,a) = V(s) + (A(s,a) - 1/|A|\sum A(s,a'))$$

[0040]

[0041] 여기서, s는 현재 상태의 작업 또는 작업들, a는 다음 선택할 작업 또는 행동을 나타내며, V함수는 전체 작업들(s)에 대한 가중치 정보 함수, A함수는 전체 작업들에서 다음 작업을 선택할 때의 가중치 정보 함수, $1/|A|\sum A(s,a')$ 함수는 어떤 작업(s)에서의 선택할 수 있는 다음 작업(a')들에 대한 가중치 합의 평균을 구하는 함수를 의미한다.

[0042] 본 발명의 일 실시예에 따르면, 생성모듈(100)은 학습환경의 과적합 문제를 해결하기 위해 두 개의 추정함수 또는 복수 개의 추정함수를 활용할 수 있다. 이는 특정 상태에서 선택한 행동에 대해 지나치게 과대평가된 가치를 학습하여 노이즈로 인해 학습환경에서 발생하는 과적합 문제를 방지하기 위함이다.

[0043] 입력모듈(200)은 각각의 기계 및 작업의 파라미터가 입력되는 구성이다. 입력모듈(200)의 파라미터는 학습의 횟수, 속도, 작업의 소요시간, 무작위 난수 설정값을 포함한다.

[0044] 전처리모듈(300)은 추정함수의 학습에 용이하도록 추정함수 및 파라미터를 전처리하고 추정함수 및 파라미터를 기 설정된 초기값으로 설정한다. 본 발명의 일 실시예에 따르면, 전처리모듈(300)은 Minmax Scaler 기법 또는 Standard Scaler 기법을 활용하여 파라미터를 전처리할 수 있다.

[0045] 강화학습모듈(400)은 추정함수 및 파라미터를 이용하여 추정함수를 학습한다. 본 발명의 일 실시예에 따르면, 강화학습모듈(400)은 다중 목적함수 기반 강화학습 흐름 생산일정 및 목적함수의 파라미터에 대한 성능향상을 위해 Double-Q learning 기법 또는 Dueling Architecture 기법을 활용하여 학습할 수 있다.

[0046] 도 2에 도시된 바와 같이, 강화학습모듈(400)은 반복부(410), 탐색부(420), 비교부(430), 결정부(440), 갱신부(450), 판별부(460)을 포함하여 구성된다.

[0047] 반복부(410)는 추정함수의 학습을 기설정된 횟수만큼 반복하고, 탐색부(420)는 추정함수를 탐색하여 반복된 추정함수의 반복횟수를 탐색하며, 비교부(430)는 추정함수의 입실론 계수와 무작위 난수를 비교한다.

[0048] 여기서, 입실론 계수는 추정함수의 탐색의 비율을 결정하는 계수이며 입실론 계수는 초기값인 '1'에서 점차 '0'으로 수렴할 수 있다. 그리고, 무작위 난수는 사용자에게 의해 입력모듈(200)에 입력된 설정값에 해당된다.

[0049] 결정부(440)는 비교부(430)의 비교결과에 따라 무작위 작업순열 또는 추정함수를 통한 작업순열을 결정하는 역할이다. 상기 무작위 작업순열은 입력된 작업의 파라미터가 무작위로 섞인 작업 배열을 의미한다.

[0050] 갱신부(450)는 결정된 무작위 작업순열 또는 추정함수를 통한 작업순열을 이용하여 추정함수를 갱신하는 역할이며, 작업순열에 따라 다음작업의 가중치 정보를 기반으로 갱신할 수 있다.

- [0051] 판별부(460)는 강화학습모듈(400)에서 잘못된 행동에 대한 가중치가 과평가되어 잘못된 방향으로 학습이 진행되는 것을 방지하기 위해 반복부(410)의 기설정된 반복횟수만큼 학습을 수행했는지, 탐색부(420)의 탐색을 수행했는지, 갱신부(450)의 갱신시점, 출력모듈(500)의 작업순열 결과를 판별한다.
- [0052] 출력모듈(500)은 강화학습된 추정함수를 통해 작업순열을 출력한다. 본 발명에 따른 추정함수는 가중치 정보를 포함하고 있어, 출력모듈(500)은 우수한 작업순열을 도출한 추정함수의 가중치 정보를 기반으로 최적의 작업순열을 출력하게 된다.
- [0053] 도 3은 본 발명에 따른 강화학습을 이용한 생산공정의 스케줄링 방법을 설명하기 위한 흐름도이다,
- [0054] 먼저, 생성모듈(100)은 작업에 대한 일정을 추정함수로 생성(S100)하고, 입력모듈(200)은 작업에 대한 파라미터를 입력(S200)한다. 여기서, S100단계는 가중치 정보를 포함하는 추정함수를 생성할 수 있으며, S200단계는 사용자에게 의해 각각의 기계 및 작업에 따른 소요시간에 대한 정보를 입력하고, 학습 시 반복횟수 설정값, 학습 시 진행속도값, 난수 설정값 등을 입력할 수 있다.
- [0055] 그리고 나서, 전처리모듈(300)은 학습에 용이하도록 전처리(S300)하고, 추정함수와 파라미터를 기 설정된 초기값으로 설정한다.
- [0056] 강화학습모듈(400)은 전처리된 추정함수 및 파라미터를 이용하여 추정함수를 학습한다(S400). 추정함수 학습방법은 도 4를 참조한다. 도 4에 도시된 바와 같이, 강화학습모듈(400)에서 반복부(410)는 추정함수를 사용자에게 의해 기 설정된 학습의 횟수만큼 반복한다(S410).
- [0057] 반복부(410)에서 기설정된 횟수만큼 반복 시, S500 단계로 이어져 출력모듈(500)은 작업 순열의 결과를 출력할 수 있다. 그러나, 기설정된 횟수만큼 반복하지 못한 경우, S420 단계로 이어져 탐색부(420)는 추정함수의 재 학습횟수를 결정하기 위해 추정함수를 탐색한다(S420).
- [0058] 탐색부(420)는 추정함수를 탐색하여 추정함수의 재 학습횟수가 결정된 경우 S410 단계로 진행되어 작업순열을 출력하고, 추정함수를 탐색하지 못한 경우 S430 단계로 진행되어 비교부(430)가 추정함수의 입실론 계수와 기 설정된 무작위 난수를 비교한다(S430).
- [0059] 입실론 계수와 무작위 난수의 비교 결과에 따라, 결정부(440)는 작업순열을 결정한다(S440). 결정부(440)는 추정함수의 입실론 계수가 무작위 난수보다 클 경우 무작위로 작업순열을 결정하고, 추정함수의 입실론 계수가 무작위 난수보다 작거나 같을 경우, 상기 추정함수에 기초하여 작업순열을 결정한다.
- [0060] 본 발명의 일 실시예에 따르면, 결정부(440)는 작업과 기계간의 가중치를 포함하는 추정함수를 기반으로 작업순열을 결정하는 것으로, 현재 진행 중인 작업에서 다음에 수행할 작업 중 가장 낮은 가중치를 갖는 작업을 선택하여 작업순열을 결정하게 된다.
- [0061] 갱신부(450)는 결정된 작업순열을 이용하여 탐색단계(S420)의 추정함수를 갱신(S450)하고, 갱신된 추정함수를 이용하여 탐색부(420)는 재 탐색을 수행한다.
- [0062] 상기 S410단계 내지 S450단계는 강화학습모듈(400)이 현재 작업상태에서 다음 작업 상태를 선택하기 위해 추정함수의 가중치를 학습하는 구성이다.
- [0063] 그리고 나서, 출력모듈(500)은 추정함수의 가중치 정보를 기반으로 하여 작업순열을 출력한다(S500). 즉, 출력모듈(500)은 작업의 가중치에 따라 작업을 스케줄링할 수 있다.
- [0064] 본 발명의 일 실시예에 따르면, 출력모듈(500)은 상태-행동함수와 상태함수의 차인 이점함수를 통해 작업순열의 결과를 출력할 수 있다. 예를 들어, 출력모듈(500)은 각 작업들의 가중치 정보를 포함하는 상태함수를 이용하여 가장 낮은 가중치를 갖는 작업을 첫 번째 작업으로 선정하고, 현재 상태의 작업에서 다음 작업들에 대한 가중치 정보를 포함하는 상태-행동함수를 이용하여 현재 상태의 작업에서 가장 낮은 가중치를 갖는 작업을 출력할 수 있다. 또, 현재 상태에 대한 이점함수의 평균치를 뺀 값과 상태함수의 가중치를 합한 상태-행동함수를 이용하여 다음 작업순열을 결정하여 출력할 수 있다.
- [0065] 이와 같이 본 발명에 따르면, 강화학습을 이용한 생산공정의 스케줄링 장치(1000) 및 방법은 추정함수를 Double-Q learning 기법으로 강화학습하여 최적의 작업일정을 출력할 수 있으므로, 기계의 유휴시간을 최소화하여 기계시스템의 생산 효율성을 높일 수 있고, 기계의 과적합 문제를 방지할 수 있다.
- [0066] 이상과 같이 본 발명의 도시된 실시 예를 참고하여 설명하고 있으나, 이는 예시적인 것들에 불과하며, 본 발명

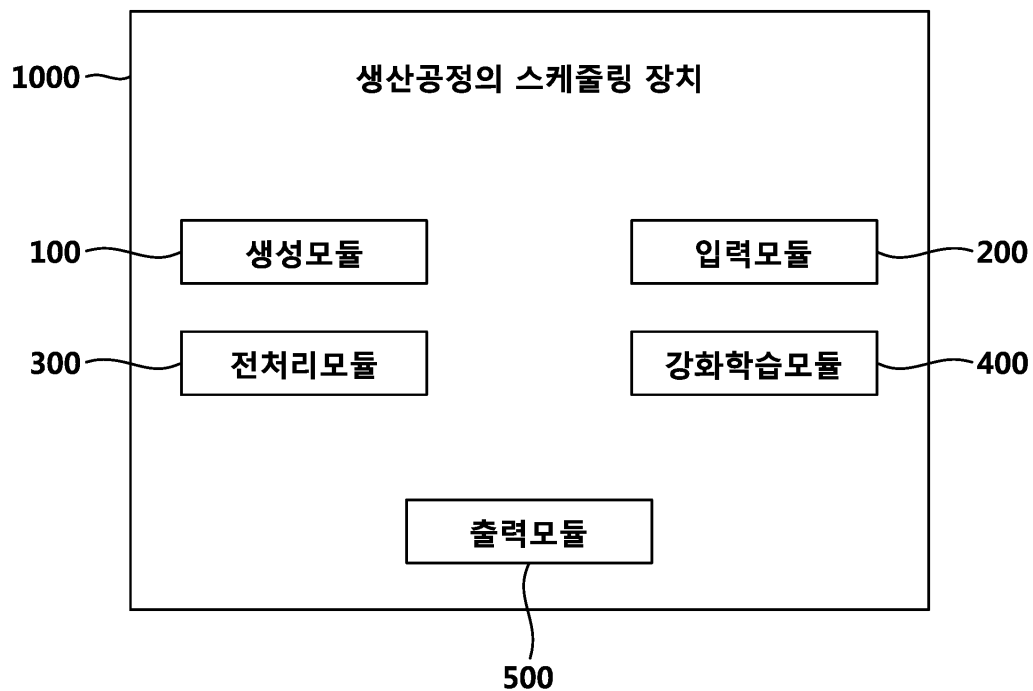
이 속하는 기술 분야의 통상의 지식을 가진 자라면 본 발명의 요지 및 범위에 벗어나지 않으면서도 다양한 변형, 변경 및 균등한 타 실시 예들이 가능하다는 것을 명백하게 알 수 있을 것이다. 따라서 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적인 사상에 의해 정해져야 할 것이다.

부호의 설명

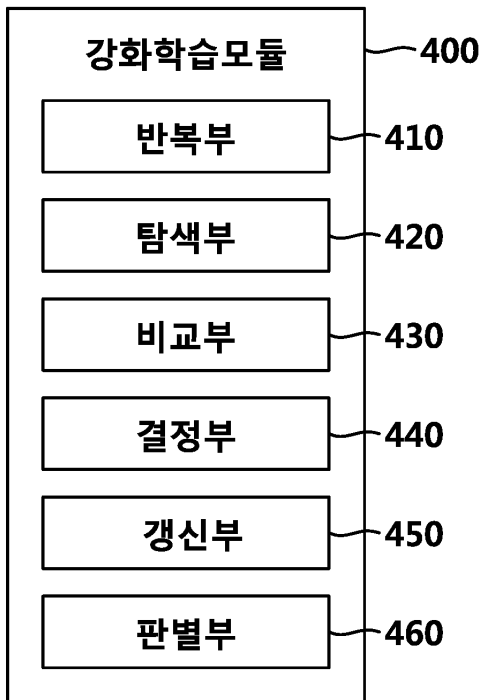
1000: 생산공정 스케줄링 장치
100: 생성모듈
200: 입력모듈
300: 전처리모듈
400: 강화학습모듈
410: 반복부
420: 탐색부
430: 비교부
440: 결정부
450: 갱신부
460: 판별부
500: 출력모듈

도면

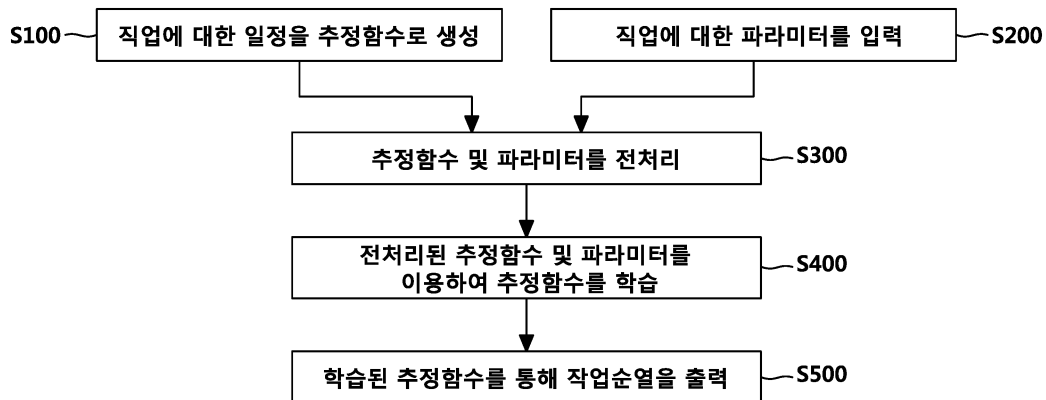
도면1



도면2



도면3



도면4

