

Lenovo Schedules Laptop Manufacturing Using Deep Reinforcement Learning

2025. 08. 29

임제원



- ✓ Lenovo의 노트북 제조 시설인 LCFC 공장에서 수동 생산 스케줄링 시스템을 Deep Reinforcement Learning(DRL)으로 대체
(~2015: 전적으로 계획 팀의 담당자 판단에 의존)
(~2017: 수동 스케줄링 결과를 ERP 및 MES 시스템에 연결하여 모니터링 하고 효율 향상, 스케줄링 주기 최대 6시간)
(2018~: Advanced Planning and Scheduling System(DRL) 구축. OR 및 AI 기반으로 새로운 스케줄링 플랫폼 생성)
- ✓ 생산량, 전환 비용 및 주문 이행의 상대적 우선순위 균형을 맞추며,
4개의 제조 공장과 43개의 조립 제조 생산 라인, 2개 교대조(shift)가 있는 상황에서 스케줄링
- ✓ 생산 주문의 백로그 20% 감소, 이행률 23% 향상, 스케줄링 프로세스 6시간 → 30분으로 단축

- ✓ 20개 이상의 제품 시리즈와 550개의 제품 모델로 구성되어 있음
- ✓ 생산 시작 전에 주문이 Manufacturing Order(MO)로 나뉘며, 하나의 MO는 수천 대의 PC로 구성 될 수 있음
- ✓ 각 MO의 노트북은 동일한 모델이며 약속된 배송 날짜(Duedate)는 비슷함
- ✓ 노트북 생산 과정은 3단계로 나뉨
 - ✓ First Stage(표면 실장 기술)
: 작업장에서 마더보드 생산을 담당, 자동으로 실행되며 사람 개입이 거의 필요하지 않음
 - ✓ Second Stage(하위 조립 작업장)
: 생산의 두번째 단계로, 노트북 케이스를 디스플레이 및 키보드에 연결. 수동 개입 필요
 - ✓ **Third Stage**
: **처리할 반제품 및 예비부품이 MO를 기준으로 43개의 생산 라인에 할당됨. 가장 많은 시간이 소요**
- ✓ 시간당 생산량(unit-per-hour, UPH) 매트릭스를 이용하여 제품 및 생산 라인에 대한 행과 열로 효율성을 표현
- ✓ UPH는 직원 출석 변동, 생산 라인의 기계 상태, 도구 및 재료의 가용성에 따라 변경되기 쉬움

II Problem Definition

Figure 1. (Color online) Production Schedule Can Be Improved by Reallocating and Resequencing MOs



Notes. (Left) Example of improving the schedule by sequencing (i.e., reordering) MO #3 so that it is assembled before MO #2 online A and by scheduling MO #4 online A instead of online B. This reduces the total changeover time and thus reduces the production time for MO #4. The improved schedule also allows a new MO (#6) to be processed within the same amount of time.

✓ 생산 스케줄링(PLPP, Production Line Planning Problem)은 두 가지 문제를 해결해야 함

1. MO를 어떤 라인에서 실행(할당)하는지 결정하는 것
2. 모든 조립라인에서 각 MO의 실행 순서

✓ KPI

1. 특정 기간 동안 생산된 제품의 양을 나타내는, 생산량
2. 주문되었지만 아직 배송되지 않은 주문 수인, 백로그 주문량
3. 만료 예정 주문 이행률(Expiring Order Fulfillment Rate)(납기 준수)

- ✓ Lenovo의 생산 스케줄링은 기존 전문가 판단에 의해 몇 시간이 걸리는 상황
- ✓ 오늘날 현대 제조 기업들은 생산 자원 공급의 변동성으로 인해 엄청난 압력을 받고 있으므로, 아래와 같은 특성을 가진 생산 관리 시스템이 필요

1. 대규모 스케줄링 문제 해결 능력

: 하루에 최대 수만 건의 주문을 처리 가능해야 함

2. 빠른 답 능력

: 공급의 변동에 신속하게 대응해야 함

3. 더 나은 KPI 성능

: 이러한 개선은 회사의 이익 뿐 아니라 고객 만족도 향상을 통한 간접적 이익도 됨

4. 다중 기준 최적화 목표를 위한 유연한 구성 기능

: 담당자가 KPI 임계값을 구성하고, 우선 순위(가중치)를 설정하는 등 의사 결정 프로세스에 참여할 수 있어야 함

→ 그러나 기존의 **B&B** 및 **Cuttingplane** 방법은 **소규모 문제에만 국한되며**, **Tabu Search**와 같은 방식은 대규모 문제를 해결하기에는 **너무 느리다**는 단점이 있음

❖ Encoder

- ✓ 단순한 네트워크 구조로는 전환비용 계산, 생산 라인 선택과 같은 비선형적 구조를 제대로 모델링 할 수 없음
- ✓ 따라서 기존의 encoder를 2-layer nonlinear Convolutional Neural Network(CNN) 기반으로 업그레이드하여 사용

❖ Decoder

- ✓ 생산 시스템의 가장 중요한 네 가지 제약
 1. 생산 시간: 각 주문의 생산시간, 휴식시간과 교대 인수인계 시간을 포함하는 구성된 총 시간
 2. 생산 수량: 제품 모델에 특수 장비가 필요한 경우, 특정 기간 내의 총 생산 수량이 제한될 수 있음
ex) 2시간 마다 최대 200개
 3. 할당 가능한 생산 라인: 각 주문은 모델을 처리할 수 있는 생산 라인에만 할당 가능
 4. 연관성: 일부 주문은 특정 플래그에 지정되어(특정 라인에 할당되어) 완료되어야 함
- ✓ 이러한 제약 조건을 만족하기 위해 새로운 마스크 메커니즘 도입
(각 요소는 특정 위치의 특정 라인에 주문을 배치하는 것이 가능한지 여부를 제어)

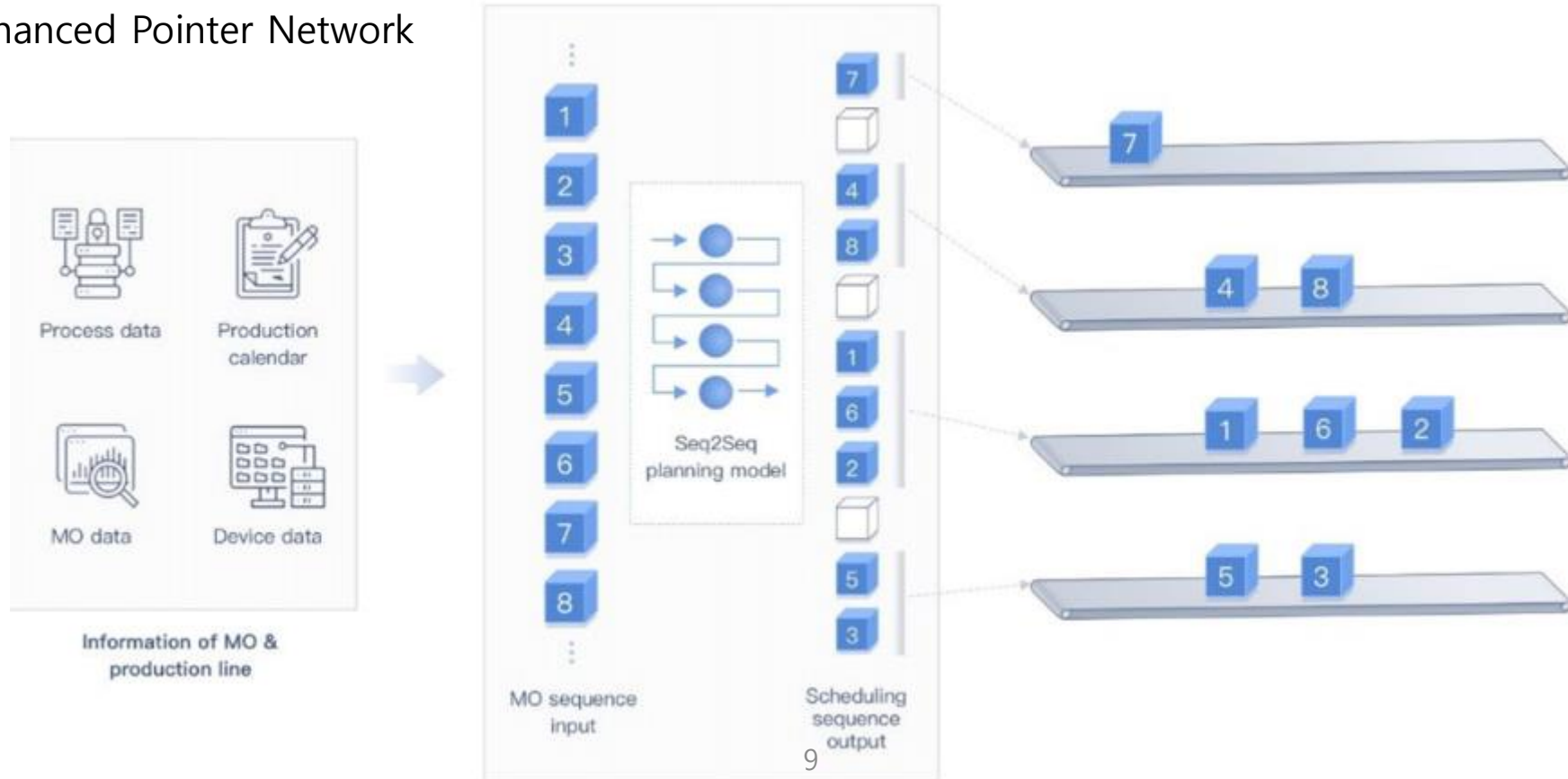
- ✓ 이러한 문제 해결을 위해, PLPP를 MDP로 공식화 진행
- ✓ 임의의 초기 시퀀스(주문시간 기준)를 받은 뒤에, **Sequence-to-Sequence 기반 Encoder-Decoder 구조로** 문제를 해결
- ✓ Encoder는 MO의 초기 시퀀스를 취하고, Decoder는 최적화된 MO 시퀀스를 생성
- ✓ 이 때, **출력 값으로는 MO Index와 분리 flag**가 나옴 (나윤, 민주 논문 방식과 매우 유사)
ex) MO1, MO6, MO2, flag, MO5, MO4, MO3
- ✓ 이러한 방식은 문제 크기가 증가해도 계산 시간에서 이점이 생기며,
Pointer Network를 도입하여 주문 개수가 늘어나도 처리가 가능함

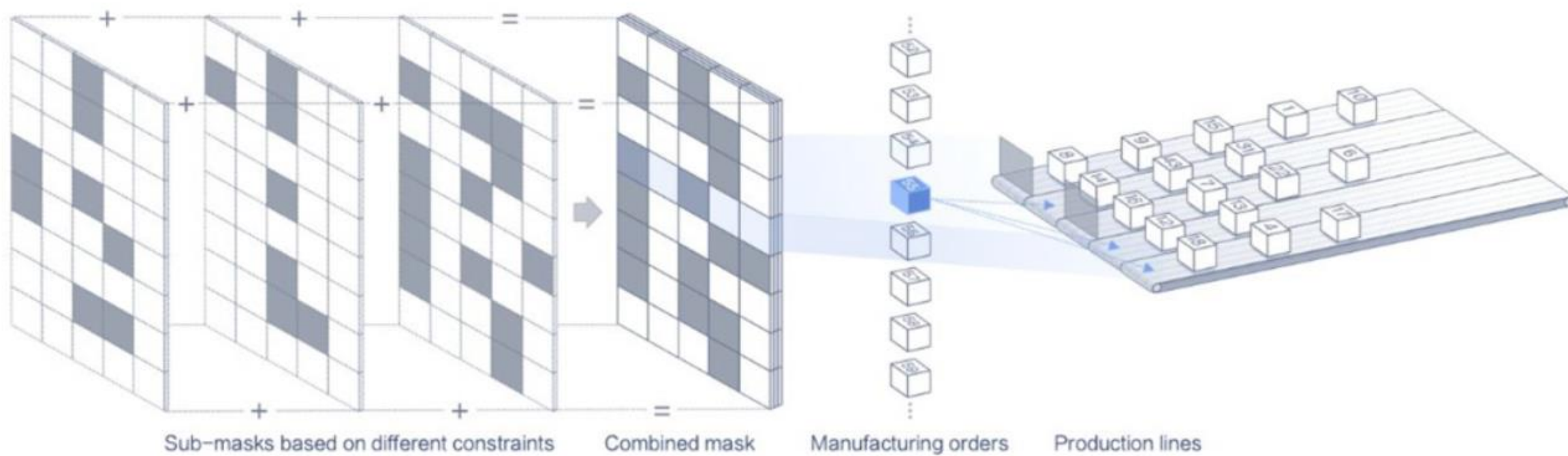
- ✓ 입력값: 주문측 정보와 공장 측 정보가 포함됨

주문 측 정보: 각 주문의 요청된 제품 수량(Quantity), 제품 시리즈 및 제품 ID(Type)

공장 측 정보: 생산 라인 수, 각 생산 라인의 각 모델에 대한 생산 효율성, 각 생산 모델 쌍간의 전환 비용 및 제조 규칙

- ✓ 기계 가용성 상태와 UPH 및 생산 일정을 포함한 주문 정보와 생산 정보는 MO에 포함됨
- ✓ Encoder-enhanced Pointer Network





- ✓ EEPN 검증을 위해 Heuristic, Meta-heuristic, RL과 비교를 진행
 Heuristic: CW(Clarke-Wright heuristics), EDF(Earliest deadline first heuristics)
 Meta-heuristic: GA(Genetic Algorithm), PSO(Particle swarm optimization)
 RL: Nazari et al.(VRP 문제를 RNN Decoder + Attention으로 해결한 2018 NeurIPS 논문)
- ✓ 검증은 15개의 생산라인, 교대조 당 2,000~9,000 오더, 30개의 교대조(shift)를 스케줄링
 (학습 사이즈에 대한 언급이 없는 것으로 보아, 해당 사이즈로 학습한 것으로 추측)
- ✓ 테스트 샘플에는 8개의 제품 시리즈와 92개의 laptop 모델이 포함되어 있음
- ✓ 목적함수는 비즈니스 제약 조건을 위반하지 않는 선에서 제품 수를 최대화 하는 것

Table 1. EEPN Yields Higher Production Volumes Than Conventional Heuristic Methods and a Previous Reinforcement Learning Method and Exhibits Faster or Similar Computation Speed

Algorithm	EEPN	Nazari18	PSO	GA	EDF	RANDOM	CW
Production volume	17,673.5	14,836.8	16,597.6	16,596.9	14,625.5	13,118.4	15,696.3
Runtime (s)	2.71	2.48	1,814.90	1,810.86	1.16	1.06	48.12

Note. Narari18 refers to the reinforcement learning method described in Nazari et al. (2018); RANDOM refers to random scheduling.

- ✓ 다중 목표 최적화 문제를 해결하기 위해서는 비지배 해 집합을 생성해야 함
- ✓ 즉, 이 집합의 모든 해는 모든 목적함수에 대해 집합의 다른 해보다 나쁘지 않으며 각 해는 특정 목표 조합에 대해 최적

ex) 우측의 표를 볼 경우, 해 D는 해 A에 비해 비용도 크고

지연도 길게 발생하므로 A에 지배당함

해 집합	비용	지연
A	100	5
B	120	1
C	150	0
D	120	10

- ✓ 계획 담당자는 생성된 해 집합에서 여러 기준의 기대를 가장 잘 충족하는 하나의 해를 선택해야 함
- ✓ 계획 담당자가 여러 가중치를 조절해본 뒤 최적의 솔루션을 시각적으로 선택할 수 있도록 함
- ✓ 이를 위해 다양한 목적함수 가중치를 추가 입력 데이터로 활용하여 여러 가중치를 단일 모델로 학습
(각 가중치 별 학습을 하는 것은 현실적으로 불가능하기에)

- ✓ 이후 해당 모델을 LCFC에 적용하여 2019년에 19억1천만 달러, 20년에 26억 9천만 달러의 수익 증가를 가져옴

❖ Contribution

- ✓ 다중 목표 최적화를 위해 여러 가중치를 입력으로 받아, 계획자가 조정 가능하도록 설계한 모델
- ✓ Flag(Line Token)을 이용하여 작업이나 머신 수에 영향 받지 않는 모델을 설계
- ✓ 다양한 제약 조건을 Masking Mechanism을 이용하여 표현하고 해결

❖ Disappointing Point

- ✓ 2-layer nonlinear Convolutional Neural Network에 대한 설명이 없음
- ✓ 학습된 사이즈를 언급하지 않았으며, 모델에 대한 전반적인 설명만 있는 상태
(기존보다 작은 사이즈로 학습하였다고만 기재되어 있음)
- ✓ 실제 공정에 사용한 모델은 다른 비교군과 실험한 결과가 기재되어 있지 않으며, 기존 전문가 대비 향상된 성능만 언급