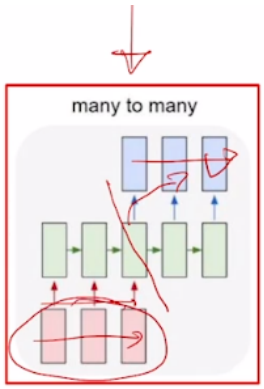
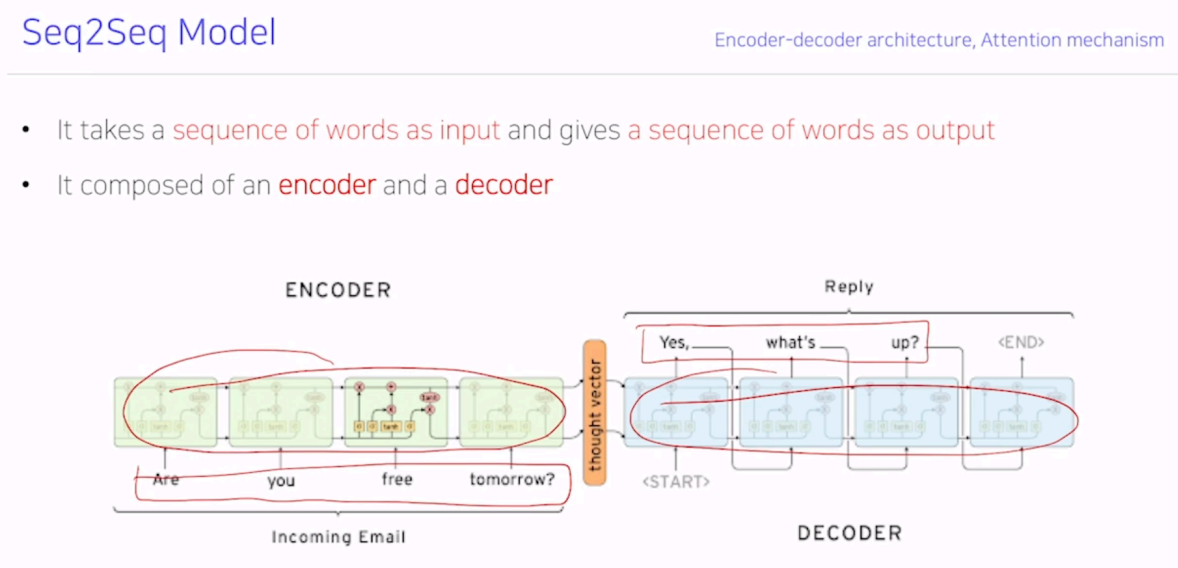
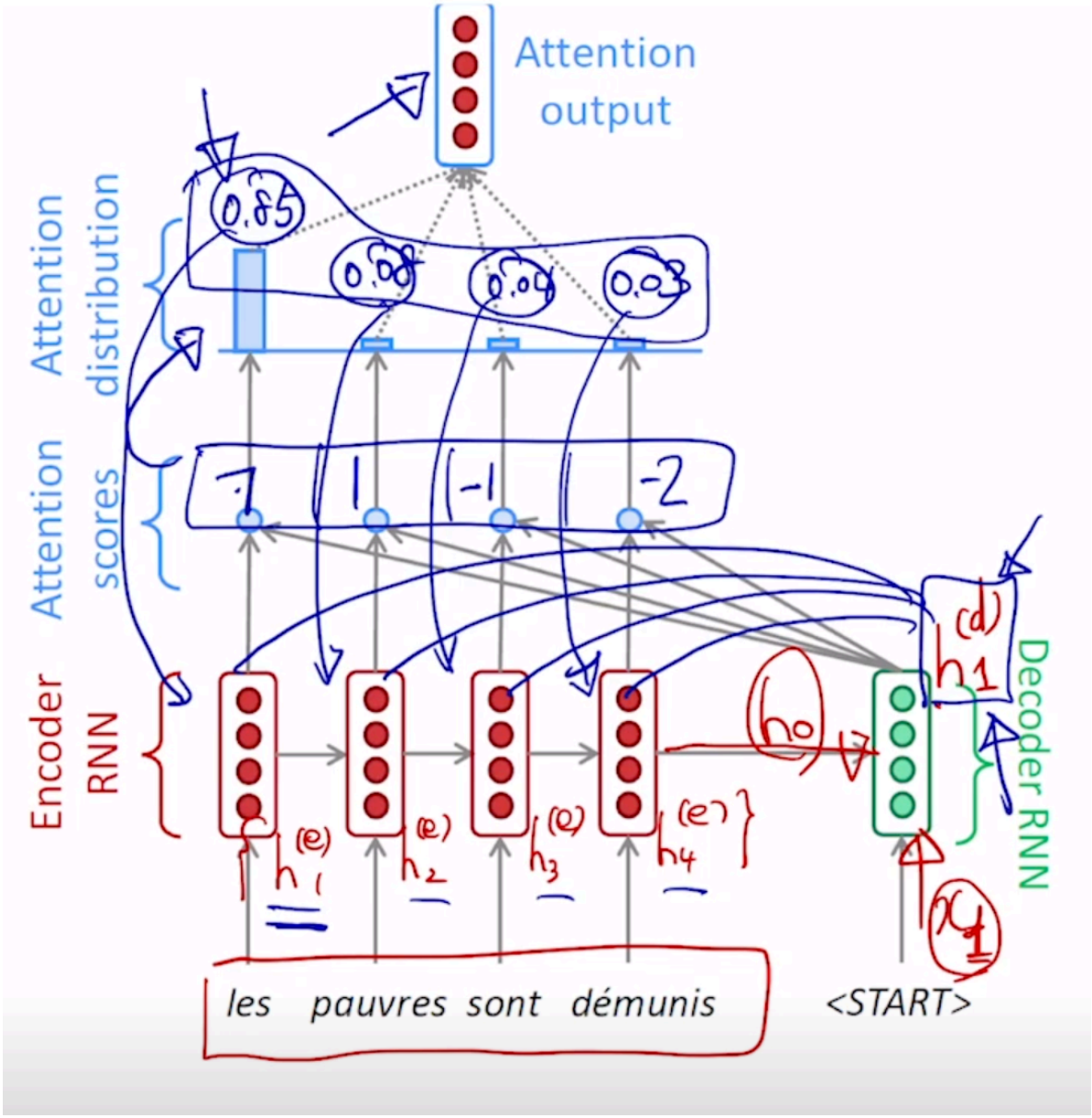
Week4-3

1. 학습정리
    1. seq2seq with attention

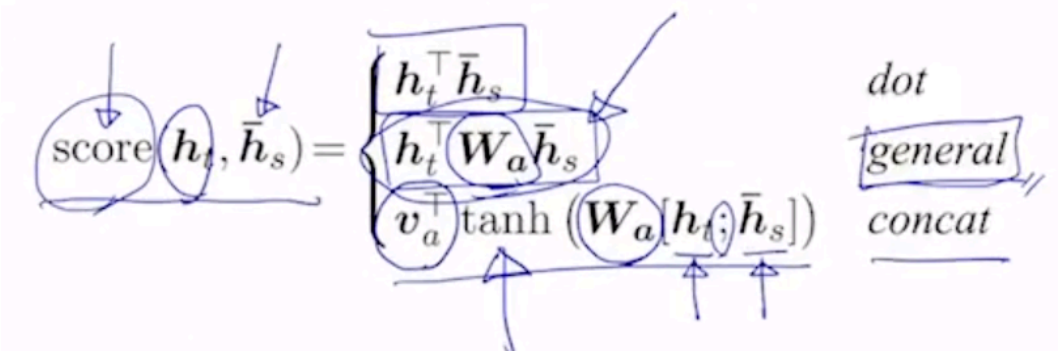


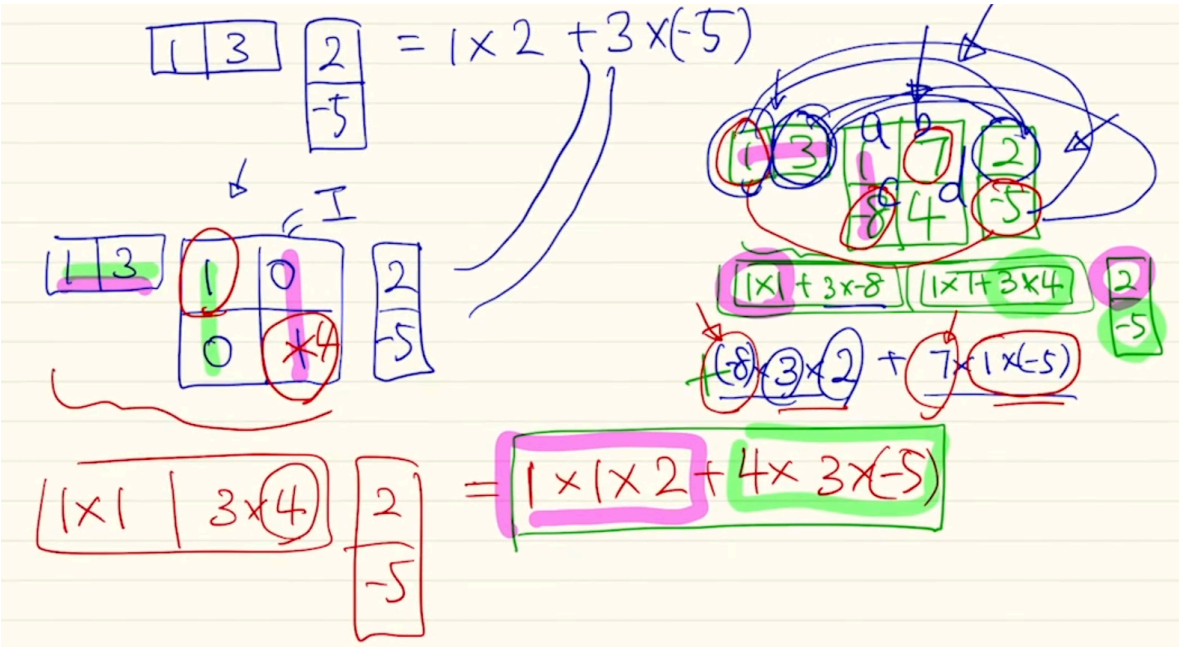




1. many to many 구조의 RNN에 해당
2. 인코더의 마지막 히든 스테이트가 디코더의 초기 히든 스테이트로 들어감
3. 어텐션 모듈
    1. 디코더에서 그때그때 필요한 인코더의 히든 스테이트를 가져옴
    2. 어텐션 스코어
        1. 디코더 히든스테이트 벡터와 인코더 히든스테이트 벡터의 각 스텝을 내적한 것
    3. 각각의 스코어를 소프트 맥스를 통과시키면 각각의 히든 인코더 스테이트에 대응하는 확률 값으로 계산됨
    4. 위 확률값을 인코더 히든스테이트에 적용되는 가중치로 사용해 가중평균을 해서 나오는 하나의 인코딩 벡터를 디코더에서 사용
    5. 입력
        1. 디코더 히든 스테이트 벡터
        2. 인코더 히든 스테이트 벡터 전부
    6. 출력
        1. 인코더 히든스테이트 벡터의 가중평균된 벡터 하나
4. 아웃풋

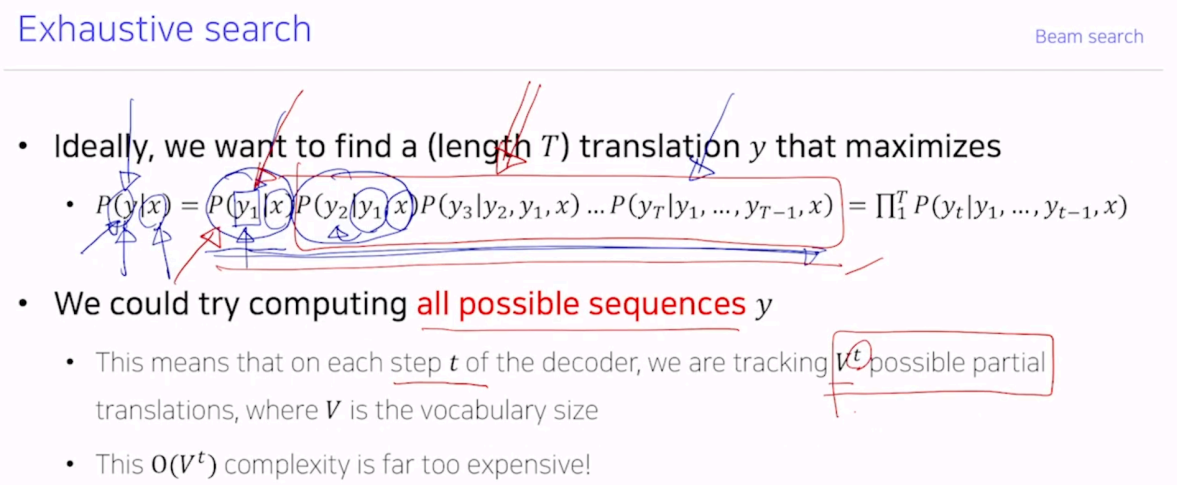
    1. 디코더의 히든 스테이트와 어텐션 모듈의 아웃풋 벡터가 컨캣되어서 아웃풋 레이어의 입력으로 들어가서 다음 나올 단어 예측
5. 디코더의 인풋은 teacher forcing 방식으로 입력
    1. 학습이 빠르고 용이
    2. 테스트 케이스에서는 잘 안될 수 있음
    3. 학습의 초반부에만 티쳐포싱을 사용하고 후반부에는 제대로 하는 방식 존재



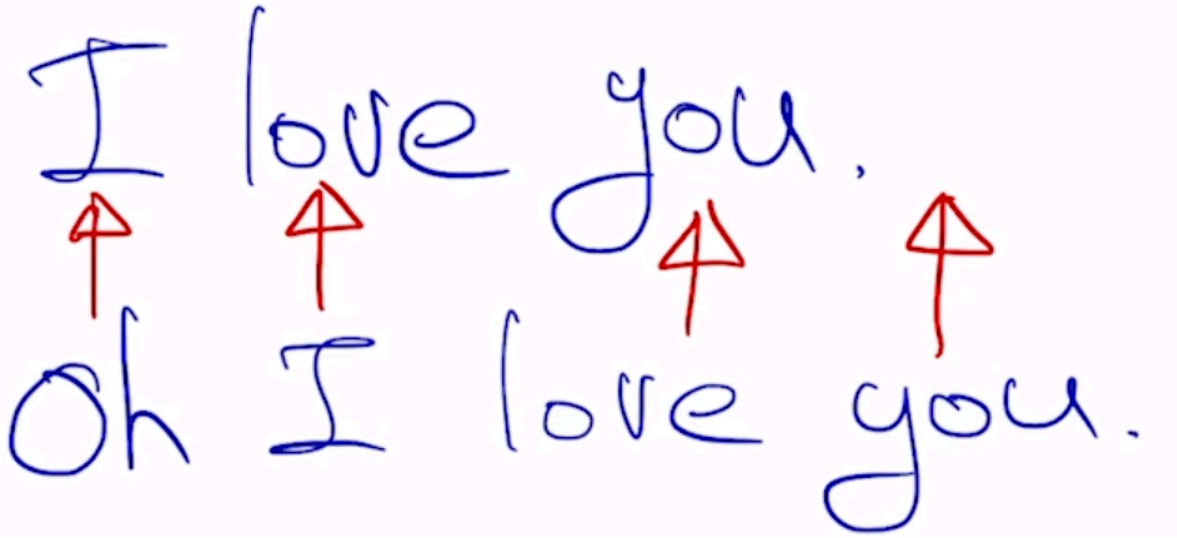


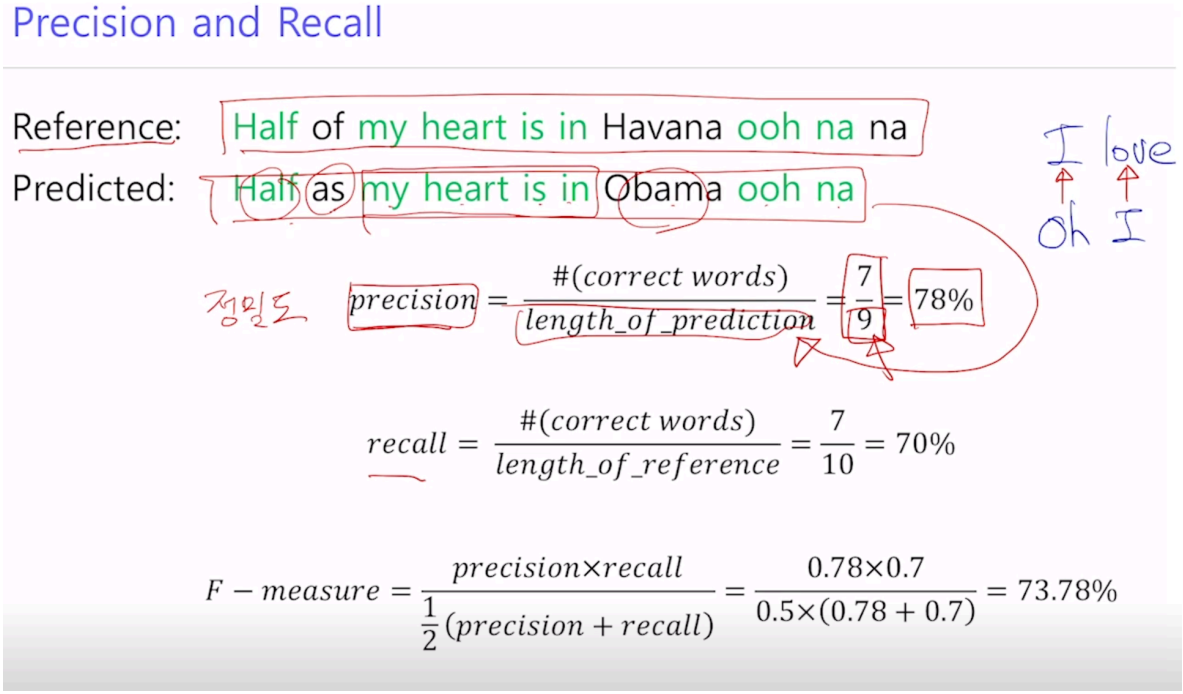
6. 어텐션 모듈에서 스코어를 구할 때 내적연산의 확장
    1. 제너럴 어텐션
        1. 중간에 학습 가능한 파라미터로써의 행렬을 만듦
        2. 내적 연산을 수행할 때 서로다른 디멘전 간의 곱셈 간에 부여되는 가중치로 활용될 수 있음
    2. 컨캣 기반의 어텐션
        1. 유사도를 구해야 하는 두 벡터를 컨캣해서 멀티레이어 뉴럴 네트워크의 입력으로 주는 방식
7. 어텐션의 장점
    1. 기계 번역의 성능 향상
    2. 인코더 디코더간의 병목 문제 해결
        1. 긴 문장에서의 번역의 어려움 문제
    3. 그래디언트 배니싱 문제 해결
    4. 모델의 번역과정을 해석가능
        1. 번역할 때 인코더의 어떤 단어에 집중했는지 볼 수 있음

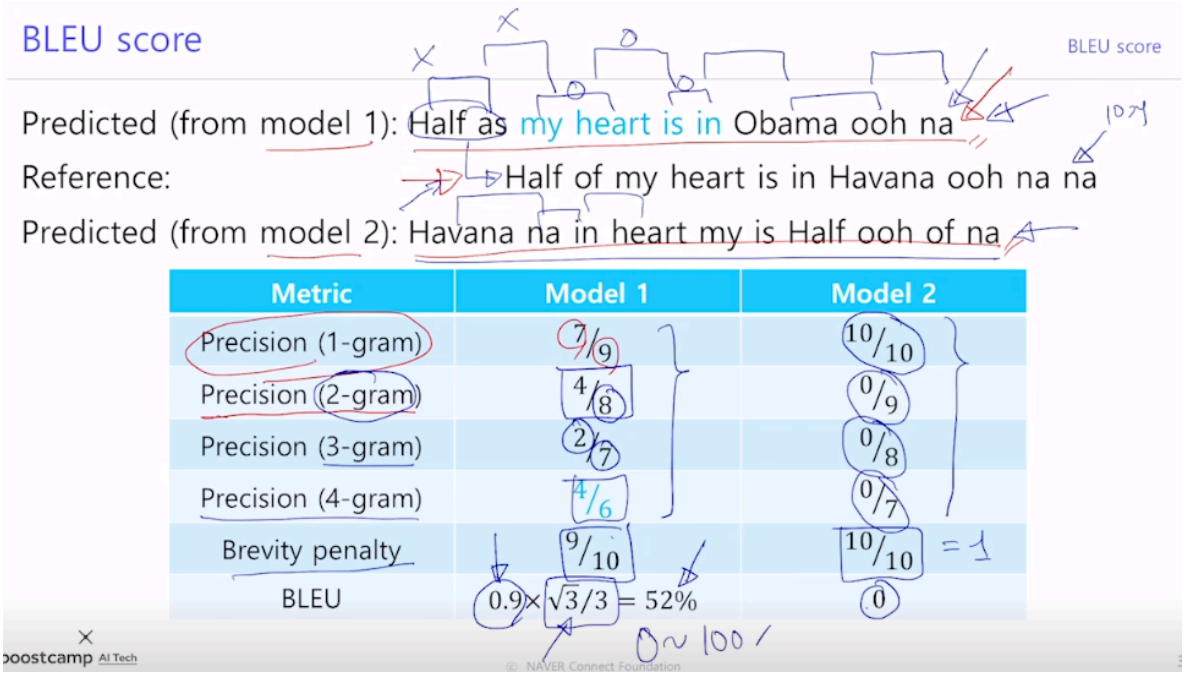
2. Beam Search
    1. 자연어 생성 모델이 테스트 타임에서 더 좋은 생성결과를 얻을 수 있게 하는 기법
    2. 지금까지 배운 건 그리디 디코딩
        1. 디코더의 아웃풋으로 가장 확률이 높은 단어 하나만을 선택
        2. 단어를 잘 못 생성했다는 걸 뒤늦게 깨달아도 뒤로 돌아갈 수 없음
        3. -> 최적의 예측값을 예상하지 못함
    3. 가능한 모든 경우를 다 따지는 디코딩



        1. 모든 가능성을 다 계산하는 것은 복잡도가 너무 심함
    4. 빔 서치는 위 두 방식의 중간
    5. 핵심 아이디어
        1. 정해놓은 k(빔 사이즈) 개의 가지수만 고려하는 것
        2. 최종적으로 k 개의 후보 중 하나 고르는 것
    6. 종료조건
        1. 두 후보가 모두 종료 토큰을 찾았을 때
    7. 아웃풋
        1. 두 후보의 길이가 다르면 긴 쪽의 예측값이 상대적으로 낮게 나옴
        2. 길이로 노말라이즈 해주어서 해결
4. BLEU 스코어
    1. 자연어 생성 모델의 정확도를 평가하는 척도







2. 매 타임스텝에서의 결과값과 예측값을 비교하는 식으로 평가하면 위의 예시는 정확도 0%가 됨
3. precision (정밀도)
    1. 맞춣 단어 수 / 예측값
4. recall (재현율)
    1. 맞춘 단어 수 / 레퍼런스
    2. 원래 나왔어야 하는 단어 중 몇 개를 놓쳤는지
5. F-measure
    1. precision 과 recall 의 조화평균
    2. 둘 중 작은 값에 치중된 계산을 하겠다는 의도
6. 블루스코어
    1. 얼마나 겹치는가를 계산 뿐만 아니라 연속된 N 개의 단어가 얼마나

            겹치는 가를 계산 (N-gram)
        2. precision 만을 고려
    7. brevity penalty
        1. 예측 길이 /레퍼런스 길이
        2. 길이가 일치하지 않을 때 받는 패널티
    8. 스코어 결과
        1. 1-gram 에서 4-gram 까지를 기하평균 한 후 브레비티 패널티를 곱한 값

2. 피어세션
    1. 수업 리뷰
    2. 과제에 대한 조언
    3. 알고리즘 문제 풀이
        1. https://leetcode.com/problems/132-pattern/