

데이터 사이언스 (Data Science)



원지혜
(jhwon@pknu.ac.kr)

Chapter 5

패턴 분석

- 패턴 분석이란 무엇인지
- 관찰자의 시각에 따라: 순차형, 선택형, 반복형
- 연관규칙
- 반복되는 패턴 분석

1. 순차형 구조

■ 순차형 구조란?

- 어떤 업무를 수행할 때 분기되거나 반복되는 내용 없이 **단순히 한 번 처리하고 끝나는 형태의 구조**를 의미
 - 예를 들어 은행에서 업무를 보려면 먼저 번호표를 뽑고 기다리다 순서가 되면 업무를 마치고 나오는 형태를 순차형 구조
- 이와 같은 구조를 순차형 패턴 흐름이라고 함

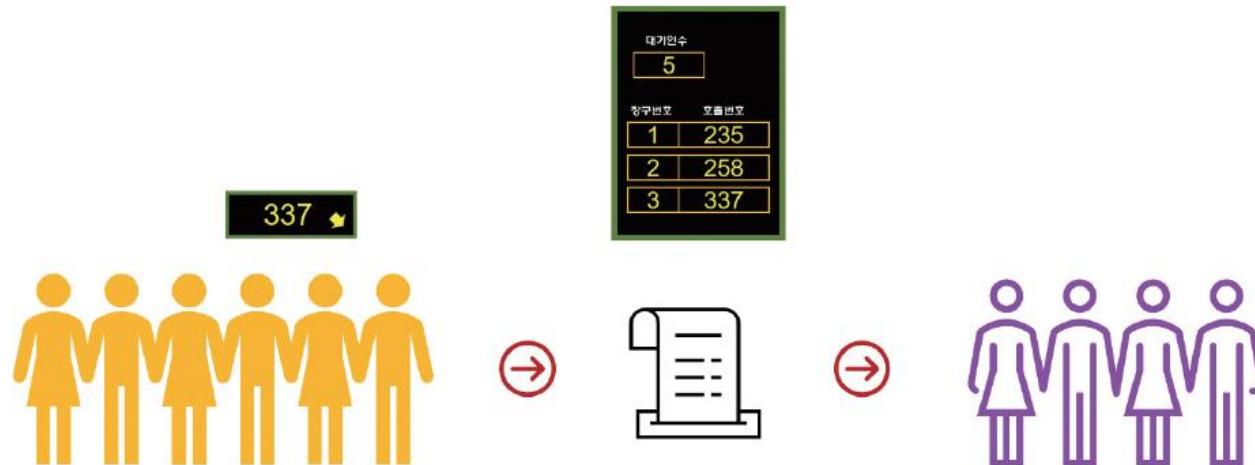


그림 5-1 순차형 패턴 흐름

2. 선택형 구조

■ 선택형 구조란?

- 여러 가지 선택 항목 중 자신의 기준에 가장 적합한 항목을 선택하는 구조를 의미
 - 예를 들어 대학교에 입학하고자 할 때 어떤 학과를 선택해서 지원해야 할 지에 대한 구조를 선택적 구조라고 함
- 선택형 구조에서는 선택하기 위한 기준이 되는 특정 조건이 제시되며 주어진 조건의 판별에 따라 참과 거짓으로 구분되어 항목을 선택하게 됨
- 이와 같은 구조를 선택형 패턴 흐름이라고 함

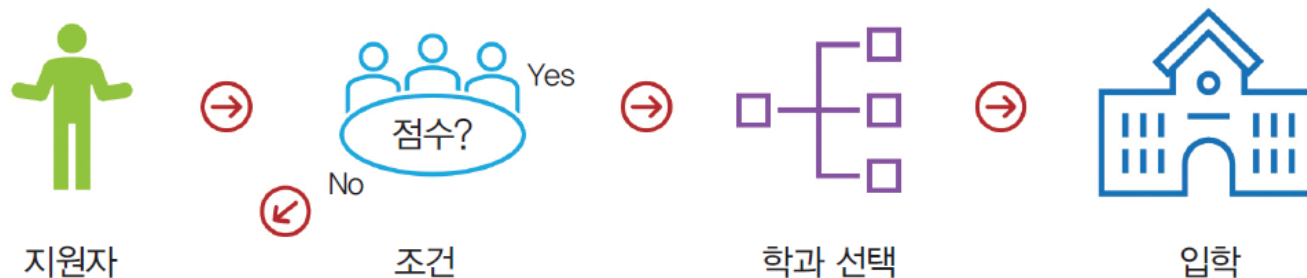


그림 5-2 선택형 패턴 흐름

3. 반복형 구조

■ 반복형 구조란?

- 주어진 조건이 참일 경우 계속해서 반복하고 거짓이 될 경우 반복을 멈추는 구조를 의미
 - 예를 들어 자동차 면허를 취득하기 위해 학과 시험에 응시하여 불합격하였다면 학과 시험을 합격할 때까지 계속해서 도전하는 구조
- 이와 같은 구조를 반복형 흐름 패턴이라고 함



그림 5-3 반복형 패턴 흐름

1. 연관 규칙

■ 연관 규칙이란?

- 명시된 한계에 따라서 빈번하게 나오는 항목값의 조합을 나타내는 공통된 약속을 의미
- 연관 규칙은 특정 사건이 발생했을 때 빈번하게 발생하는 또 다른 규칙을 발견하는 규칙 기반 학습임
- 연관 규칙은 항목 순서 또는 인과 관계를 고려하지 않고 서로 연관성만을 가지는 의미로 해석해야 함



그림 5-4 키 성장과 운동의 연관성 (출처 : 네이버)

2. 패턴 찾기

■ 패턴 찾기란?

- 패턴을 찾는다는 것은
일정한 변화의 연관 규칙을 찾는 행위로 수집된 데이터 또는
표본 집합에 나타나는 일정한 경향이나 규칙적인 패턴을 찾는 과정
- 패턴을 찾는다: 일정한 변화의 연관 규칙을 찾는다는 것을 의미함

2. 패턴 찾기

실습 05 - 01

주어진 도형의 배치 패턴을 보고 연관 규칙 찾아내기

- **Step 01** | 3색 도형에서 한 가지 색상에 집중하여 움직이는 방향을 주시하게 되면 마지막에 위치할 도형의 모양을 유추할 수 있습니다.



그림 5-5 주어진 도형의 배치 현황

- **Step 02** | 화살표 방향으로 진행하는 도형의 모양을 살펴보면 녹색 도형이 시계방향으로 회전하는 패턴을 찾을 수 있습니다.

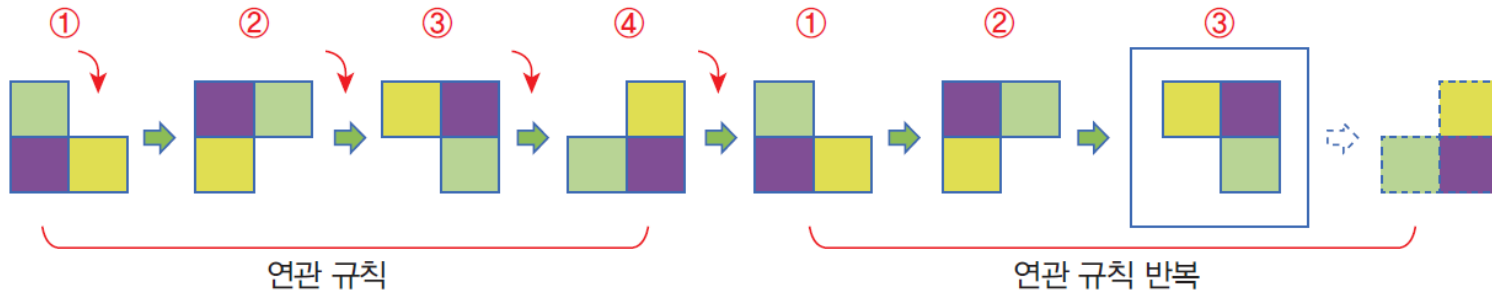


그림 5-6 연관 규칙을 통한 도형 패턴 찾기 성공

2. 패턴 찾기

도전 05 - 01

다음 지시사항을 수행하시오.

1. 도형의 연관 규칙에 따라 마지막 도형 그리기
2. 찾은 패턴을 적용하여 네모 칸에 숫자로 채우기
3. 기타 사항은 주어진 도형 그림을 참조

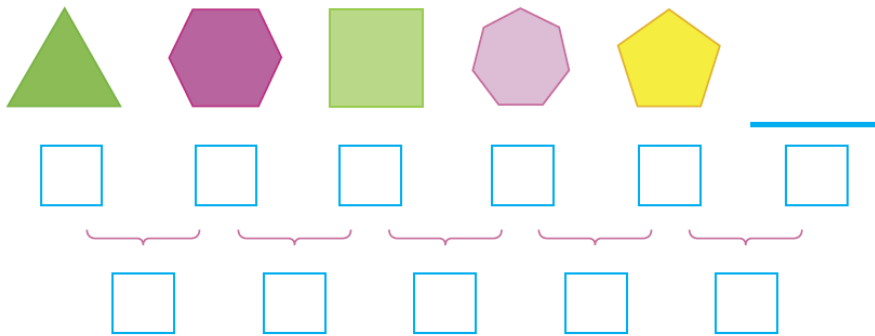
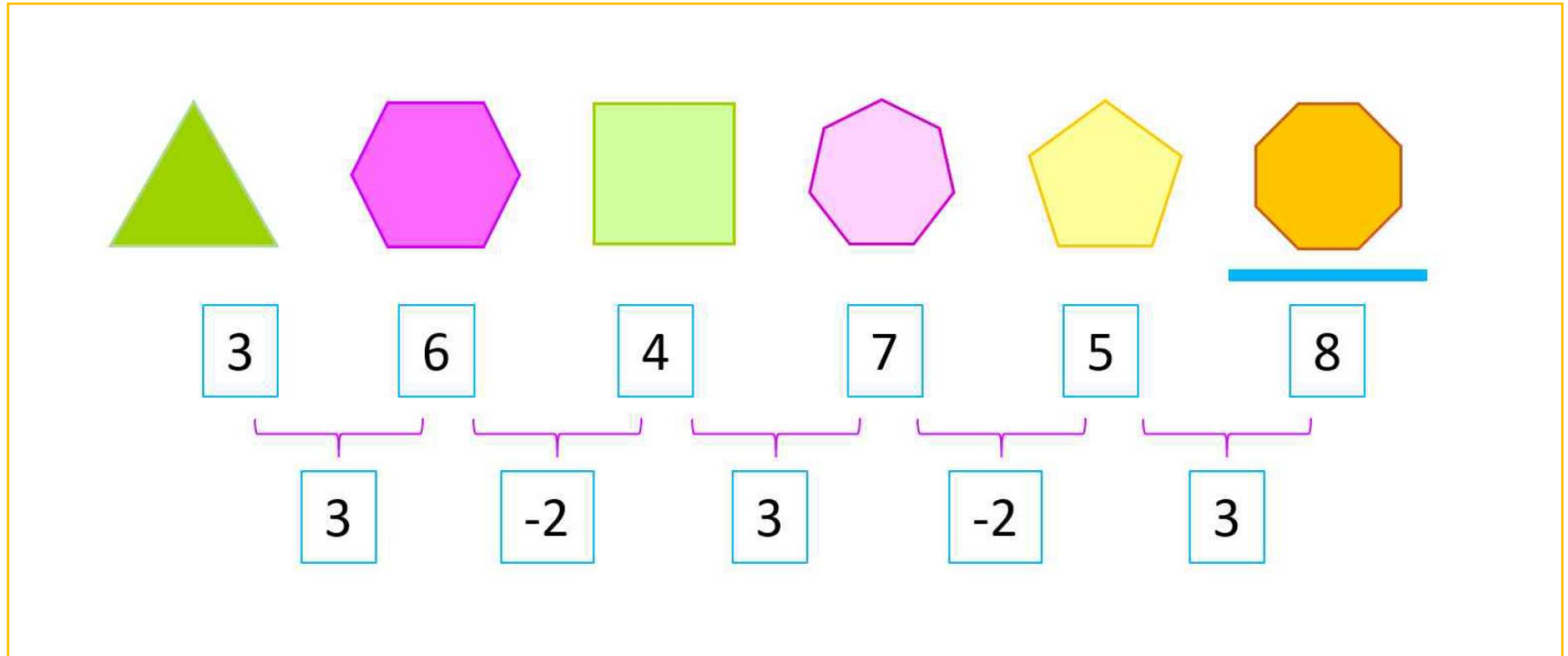


그림 5-7 도형 배열의 연관 규칙 찾기

2. 패턴 찾기



2. 패턴 찾기

■ 런 길이 부호화

- 데이터에서 연속된 문자와 개수만으로 표현하여 간단하게 비손실 압축을 수행하는 방법을 런 길이 부호화라고 함
- 비손실 압축은 압축 과정에서 아무런 손실 없이 압축을 수행하는 방법을 의미

실습 05 - 02

주어진 데이터를 보고 런 길이 부호화 방법으로 패턴 찾기

- **Step01** | 주어진 문자열 데이터의 연관 규칙을 살펴봅니다.



그림 5-8 문자열 데이터

2. 패턴 찾기

- **Step 02** | 주어진 문자열 데이터를 살펴보면 알파벳 A와 B가 연속해서 일정한 수만큼 나열되어 있다는 것을 알 수 있습니다. 여기에서 찾은 패턴은 다음과 같습니다.

- 패턴 분석 : 알파벳 A와 B가 연속해서 일정한 수만큼 나열됨

- **Step 03** | 패턴을 찾았으므로 런 길이 부호화를 적용하기 위한 일반화 과정을 수행합니다.

- 일반화 : $nAmB \rightarrow n$ 과 m 은 연속된 데이터의 수를 의미(단, 1은 표시하지 않음)

- **Step 04** | 주어진 문자열 데이터에서 찾은 패턴을 런 길이 부호화 방법으로 다음과 같이 정형화를 표기할 수 있습니다.

- 정형화 : 3AB5A3BA

2. 패턴 찾기

도전 05 – 02

다음 지시사항을 수행하시오.

1. 주어진 문자열 데이터의 패턴 분석 제시하기
2. 찾은 연관 규칙을 적용한 런 길이 부호화를 적용하여 정형화 제시하기
3. 기타 사항은 주어진 도형 그림을 참조

[문] 주어진 데이터



[답1] 주어진 데이터에 대한 패턴 분석 : 알파벳 AB가 연속해서 일정한 수 만큼 나열됨

[답2] 연관 규칙을 적용한 런 길이 부호화 : 10(AB)A

1. 순차형 패턴 분석

■ 패턴 분석이란?

- 변화의 유형에 대한 추세선이 나타날 수 있는 변동 모형을 미리 정형화 하여 실제로 발생할 수 있는 움직임과 비교
- 앞으로 발생할 수 있는 추이를 예측하는 기법을 의미

■ 순차형 패턴 분석

- 주어진 문제를 해결해야 하는 과정에서 당면 과제를 작은 모듈 단위로 분해 모듈을 차례대로 진행하며 문제를 해결하는 과정에서의 패턴을 분석하는 방법
 - 은행에서 업무를 처리하기 위해 기다리는 순차형 패턴을 분석하여 자연어 알고리즘과 순서도 알고리즘을 작성하기 위해 다음 예제를 수행

1. 순차형 패턴 분석

예제 05 - 01

은행에서 차례를 기다리는 순차형 구조의 패턴 분석하기

- **Step 01** | 다음과 같이 주어진 스케줄을 보고 순차형 구조로 패턴 분석을 수행합니다.

은행에 입금하기 위해 신분증을 지참하고 승용차로 가까운 은행에 방문하였습니다. 은행에는 이미 많은 사람들이 대기하고 있는 상태입니다. 은행에서 업무를 마치면 귀가하는 길에 제과점에 들러 빵과 케익을 사서 들고 오기 위해 외출합니다.

- 패턴 분석 : 도착한 순서대로 은행 업무를 수행하기 위한 번호표 발급받기

1. 순차형 패턴 분석

- **Step 02** | 은행에서 번호표를 발급받아 순서를 기다리는 순차형 구조의 패턴을 분석하여 다음과 같이 자연어 알고리즘을 설계합니다.

표 5-1 순차형 패턴 분석 알고리즘

Algorithm Name : 은행업무 순차형 프로세스 알고리즘

Begin :

1. 입금하기 위해 은행을 방문하여 기다리는 순서를 정한다.
 - 1.1 번호표 발급기에서 번호표를 발급받는다.
 - 1.2 대기 의자에 앉아서 순서를 기다린다.
 - 1.3 본인 확인절차를 거치기 위해 신분증을 미리 준비한다.
 - 1.4 기다리는 순서가 카운트되는 스크린을 주시한다.
2. 본인의 번호가 호출되면 해당 창구로 이동한다.
 - 2.1 은행 직원의 요청에 따라 업무를 추진한다.
 - 2.2 신분증과 서명 등에 대한 요청이 있을 경우 제시한다.
 - 2.3 은행업무 중 추가로 궁금한 사항에 대해서는 은행직원에게 문의한다.
 - 2.4 은행업무를 마쳤으면 신분증 등 소지품을 잘 챙긴다.
3. 은행업무를 모두 수행하였다면 귀가한다.
 - 3.1 은행 문을 나선다.
 - 3.2 주차되어 있는 장소로 이동한다.
 - 3.3 귀갓길에 제과점에 들려 빵과 케익을 사서 들고간다.

End :

1. 순차형 패턴 분석

- **Step 03** | 순차형 패턴 분석 알고리즘을 적용하여 순서도를 그립니다.

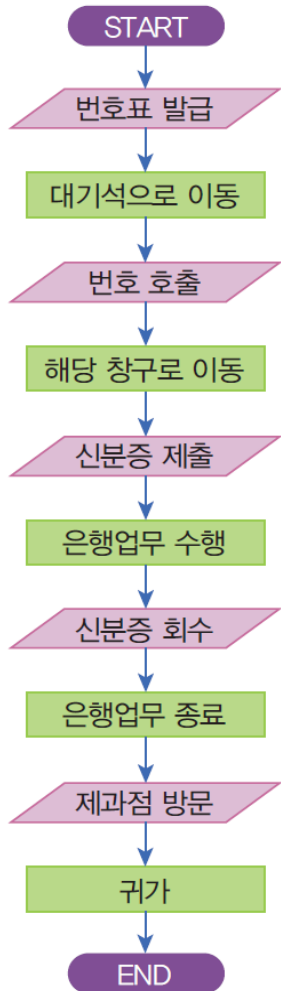


그림 5-9 순차형 순서도 작성

2. 선택형 패턴 분석

■ 선택형 패턴 분석이란?

- 여러 가지 선택 항목이 있을 경우
자신에게 가장 잘 맞는 항목을 선택하는 문제를
해결하는 과정에서 패턴을 분석하는 방법
 - 데이터사이언스학과가 개설되어 있는 대학교를 찾아서
입학지원서를 접수하기 위한 선택형 패턴 분석을 수행하여
자연어 알고리즘과 순서도 알고리즘을 작성하기 위해 다음 예제를 수행

2. 선택형 패턴 분석

예제 05 - 02

3개 대학교 중 지원할 1개 대학교를 선택하는 패턴 분석하기

- **Step 01** | 다음과 같이 주어진 대학교 지원계획을 보고 선택형 구조로 패턴 분석을 수행합니다.

전국에 있는 대학교를 알아본 결과 데이터사이언스학과에 지원할 수 있는 3개(A, B, C 대학교)의 대학교를 찾았습니다. 각 대학교에 지원할 수 있는 점수는 1000점 만점 중 A 대학교는 850점 이상, B 대학교는 800점 이상, C 대학교는 750점 이상이어야만 지원할 수 있다는 것을 확인하였습니다. 각 대학교의 모집 요강에 모두 참석하고 입학지원서를 확보하려고 합니다. 입학지원서 접수는 기간 내 해당 대학교에 직접 방문하여 접수할 계획입니다.

- 패턴 분석 : 지원학과가 있는 대학교를 방문하여 입학지원서 접수하기

2. 선택형 패턴 분석

- **Step 02** | 본인이 지원하고 싶은 대학교를 선택하기 위한 선택형 구조의 패턴을 분석하여 다음과 같이 자연어 알고리즘을 설계합니다.

표 5-2 선택형 패턴 분석 알고리즘

Algorithm Name : 대학교 입학지원서 선택형 프로세스 알고리즘

Begin :

1. 대학교 입학지원서를 접수하기 위해 다음과 같은 절차를 수행한다.
 - 1.1 지원할 대학교의 모집요강을 모두 참석한다.
 - 1.2 지원하고 싶은 대학교의 입학지원서를 확보한다.
 - 1.3 본인의 성적을 확인한다.
 - 1.4 합격할 가능성이 있는 대학교를 찾는다.
 - 1.5 입학지원서를 접수하고 귀가한다.
2. 지원하고자 하는 대학교의 커트라인 점수는 다음과 같다.
 - 2.1 A 대학교의 커트라인 점수는 850점이다.
 - 2.2 B 대학교의 커트라인 점수는 800점이다.
 - 2.3 C 대학교의 커트라인 점수는 750점이다.
3. 입학지원서 접수는 다음과 같이 진행한다.
 - 3.1 합격할 수 있는 대학교를 선택한다.
 - 3.2 입학지원서는 기간내 접수한다.
 - 3.3 입학을 지원할 대학교에 직접 방문하여 지원서를 접수한다.

End :

2. 선택형 패턴 분석

- **Step 03** | 선택형 패턴 분석 알고리즘을 적용하여 순서도를 그립니다.

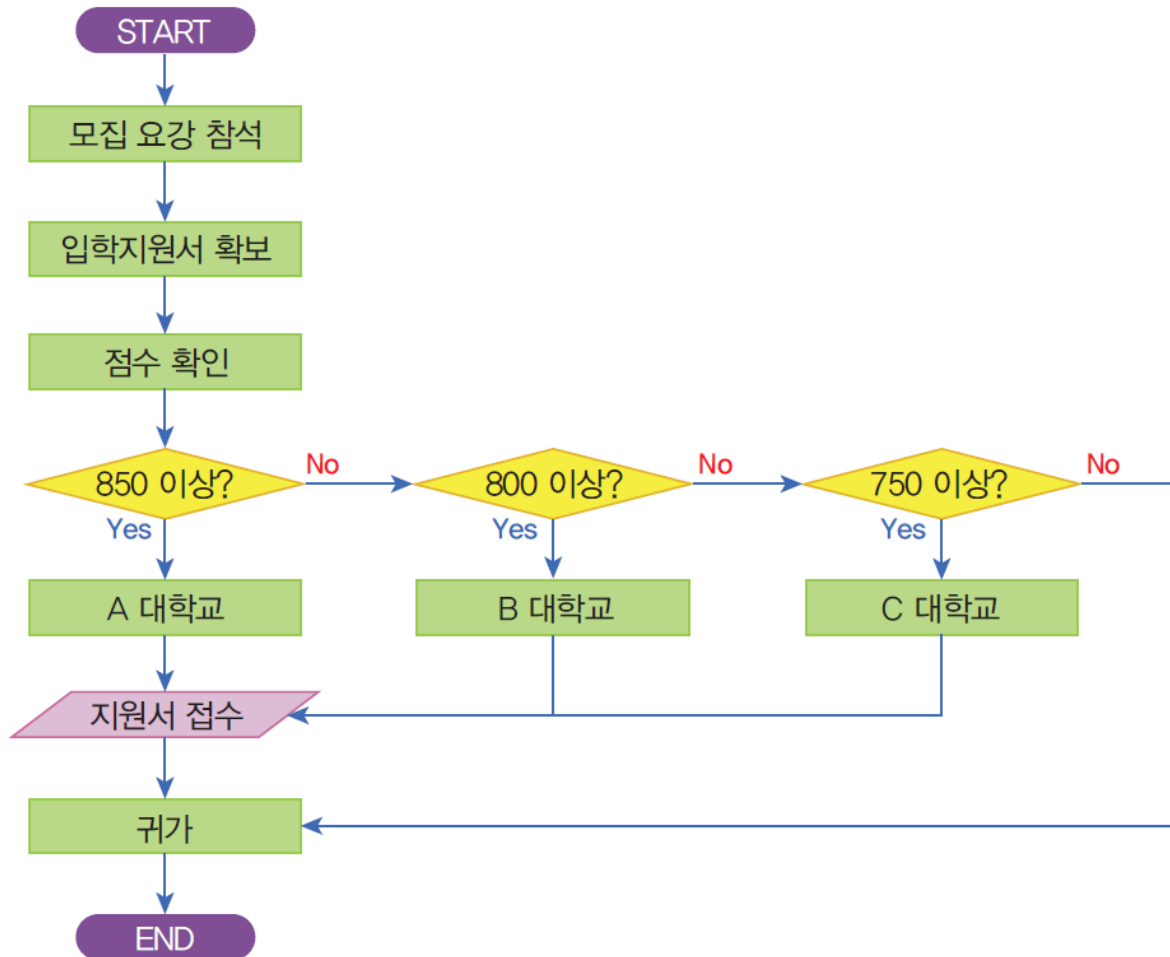


그림 5-10 선택형 순서도 작성

3. 반복형 패턴 분석

■ 반복형 패턴 분석이란?

- 주어진 조건이 참일 경우 계속해서 반복 수행하고 거짓이 될 경우 반복 수행을 멈추는 유형으로 문제를 해결하기 위한 패턴 분석
 - 운전면허를 취득하기 위한
[학과]→[코스]→[주행] 시험 중 학과 시험이 불합격일 경우 계속 도전하고 합격하게 되면 도전을 멈추는
반복형 패턴 분석 방법을 살펴보기 위해 다음 예제를 수행

3. 반복형 패턴 분석

예제 05 - 03

학과 시험을 합격할 때까지 도전하는 반복형 구조의 패턴 분석하기

- **Step 01** | 다음과 같이 운전면허 취득을 위한 [학과] → [코스] → [주행] 시험 중 학과 시험을 합격할 때까지 도전하는 패턴을 반복형 구조로 분석을 수행합니다.

운전면허가 필요할 것 같아서 알아보던 중 1종 보통과 2종 보통을 가장 많이 신청하는 것으로 들었습니다. 1종 보통 면허는 정원 15명 이하의 승합자동차를 운전할 수 있고 2종 보통 면허는 일반 승용차와 승차정원 10인 이하의 승합차를 운전할 수 있다고 합니다. 학과 시험의 커트라인 점수는 1종 70점, 2종 60점이며 둘 중 어느 면허를 신청할지 고민하고 있습니다. 학과 시험에 불합격하게 되면 합격할 때까지 계속해서 원서를 접수할 계획입니다. 그리고 학과 시험에 합격하게 되면 코스 시험에 도전할 것입니다.

- 패턴 분석 : 운전면허를 취득하기 위해 학과 시험 합격할 때까지 응시하기

3. 반복형 패턴 분석

- **Step 02** | 운전면허를 취득하기 위한 학과 시험에 응시하여 불합격하게 되면 합격할 때까지 계속해서 도전하는 반복형 구조의 패턴을 분석하여 다음과 같이 자연어 알고리즘을 설계합니다.

표 5-3 반복형 패턴 분석 알고리즘

Algorithm Name : 운전면허 학과 시험 반복형 프로세스 알고리즘

Begin :

1. 운전면허증을 취득하기 위해 다음 절차를 수행한다.
 - 1.1 취득하려는 운전면허의 종류에 대해 살펴본다.
 - 1.2 1종 보통 면허는 승차정원 15명 이하의 승합자동차를 운전할 수 있다.
 - 1.3 2종 보통 면허는 일반 승용차와 승차정원 10인 이하의 승합차를 운전할 수 있다.
 - 1.4 본인이 운전하려는 차량에 적합한 면허가 어떤 것인지를 검토한다.
 - 1.5 학과 시험에 합격하게 되면 코스 시험에 응시한다.
2. 운전면허 학과 시험의 커트라인 점수는 다음과 같다.
 - 2.1 1종 보통 면허의 커트라인 점수는 70점이다.
 - 2.2 2종 보통 면허의 커트라인 점수는 60점이다.
 - 2.3 학과 시험에 불합격하면 합격할 때까지 계속해서 도전한다.
3. 운전면허 학과 시험은 다음과 같은 절차로 지원한다.
 - 3.1 주변에 가까운 운전면허 시험장을 찾는다.
 - 3.2 시험일정에 맞추어 원서를 접수한다.
 - 3.3 시험당일 시험장에 방문하여 시험에 응시한다.

End :

Q&A