

# 확률적 프로그래밍 추론 엔진의 성능을 향상시키는 연구

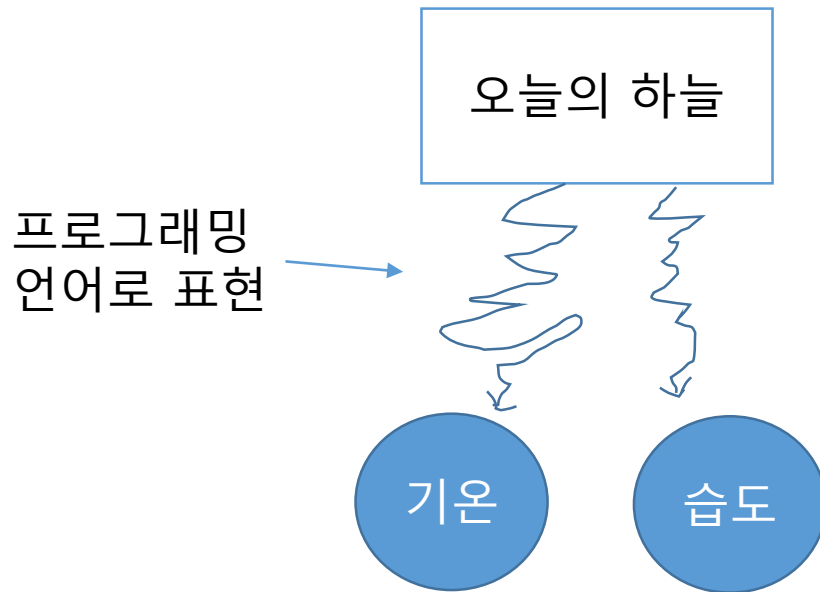
- 프로그램 변환을 통한 최적화 -  
서울대학교 소프트웨어 원리 연구실  
김윤승

# 확률적 프로그래밍으로 풀려는 문제

- 자연 현상, 사회 현상에서 관찰된 정보로부터 몰랐던 정보 찾기

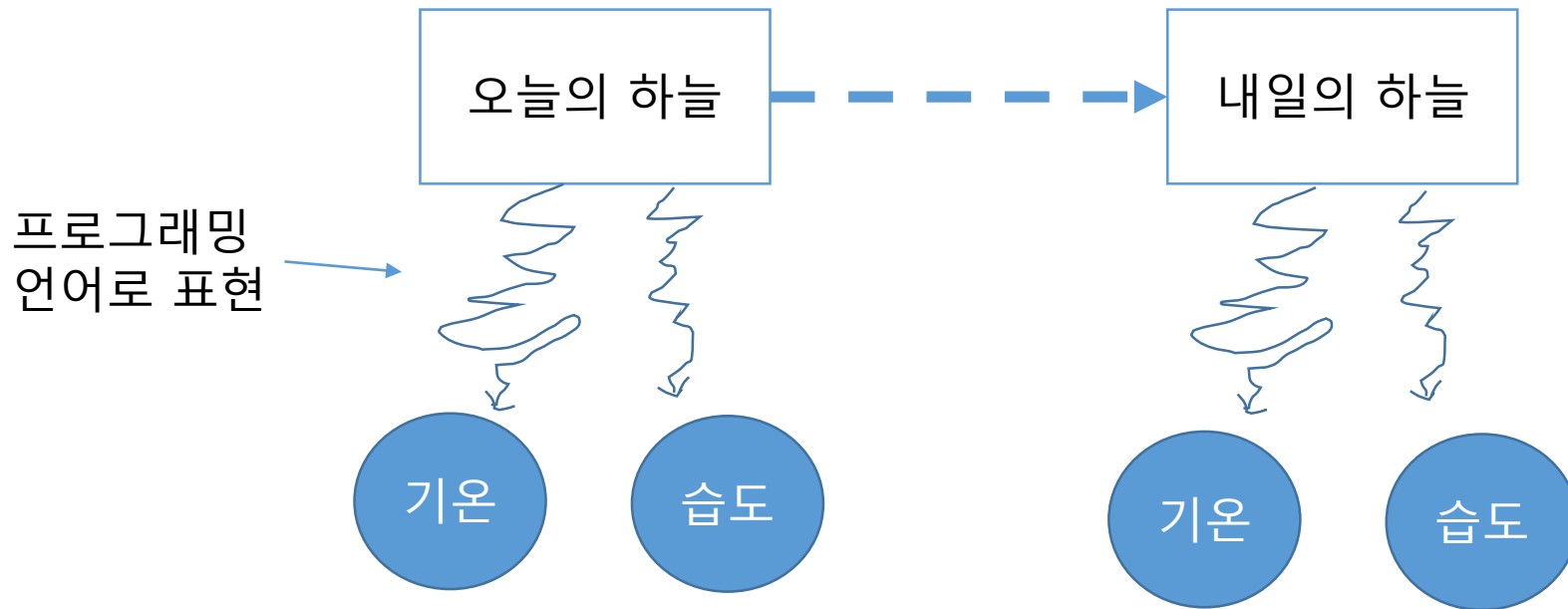
# 확률적 프로그래밍으로 풀려는 문제

- 자연 현상, 사회 현상에서 관찰된 정보로부터 몰랐던 정보 찾기
- Ex) 기상 예보



# 확률적 프로그래밍으로 풀려는 문제

- 자연 현상, 사회 현상에서 관찰된 정보로부터 몰랐던 정보 찾기
- Ex) 기상 예보



# 확률적 프로그래밍이란..

- 프로그램으로 확률 분포 나타내기
  - 사용자의 역할
  
- 그 프로그램에서 분포 알아내기
  - 언어에 내장된 추론 엔진이 자동으로

# 확률적 프로그래밍이란..

- 프로그램으로 확률 분포 나타내기
- 그 프로그램에서 분포 알아내기
- Ex.

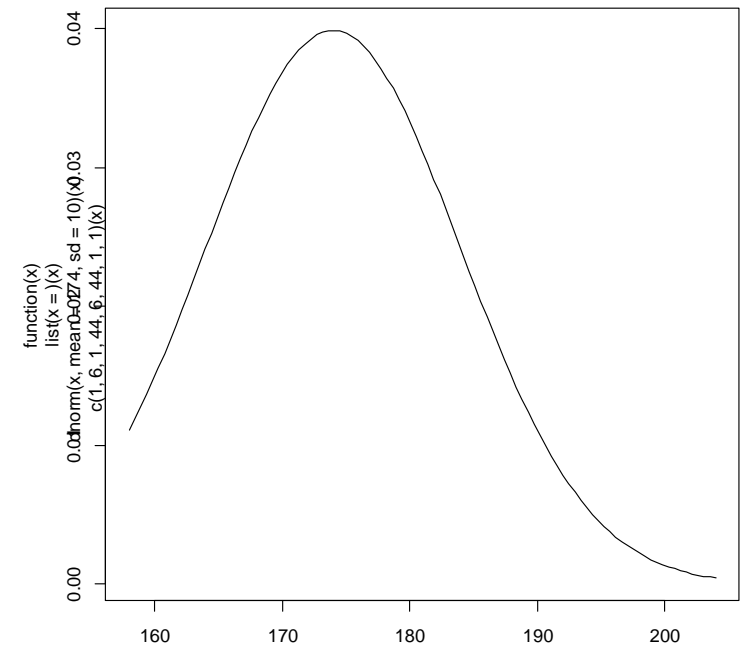


---

체중과 관계없이 신장 146 ~ 158cm	4급	공익근무대상
체중과 관계없이 신장 204cm 이상	4급	공익근무대상

---

```
x ~ Gaussian(174, 10)
observe(x > 158)
observe(x < 204)
return x
```

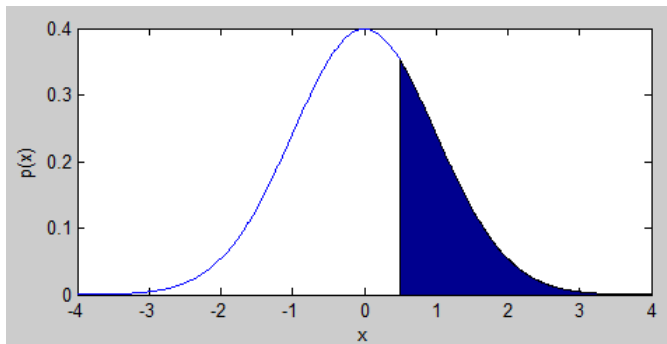


# 확률적 프로그래밍이 마주친 문제

- 프로그램에서 어떻게 확률 공간을 정의하지?
- 어떻게 해야 질 좋은 샘플을 빨리 얻을까?
- 샘플을 얻을 공간의 차원이 일반적으로 너무 넓다. 줄일 수 없을까?

# 간단한 차원 없애기

- 
- 
- $x \sim \text{Gaussian}(\theta, 1)$
- observe( $x > 1$ )
- 
- 



from Infer.NET

- 
- 
- flip( $1 - \Phi_{0,1}(1)$ )
- 
- 

쓸모가 있을까?



# 정규 분포 압축하기

$s1 \sim \text{Gaussian}(0, 10)$

$s2 \sim \text{Gaussian}(0, 10)$

$p1 \sim \text{Gaussian}(s1, 5)$

$p2 \sim \text{Gaussian}(s2, 5)$

observe( $p2 - p1 > 1.0$ )

return ( $s1, s2$ )

$s1 \sim \text{Gaussian}(0, 10)$

$s2 \sim \text{Gaussian}(0, 10)$

$p \sim \text{Gaussian}(s2 - s1, 5\sqrt{2})$

observe( $p > 1.0$ )

return ( $s1, s2$ )

# 정규 분포 압축 + 차원 낮추기

`s1 ~ Gaussian(0,10)`

`s2 ~ Gaussian(0,10)`

`p ~ Gaussian(s2-s1, 5 $\sqrt{2}$ )`

`observe(p > 1.0)`

`return (s1,s2)`

`s1 ~ Gaussian(0,10)`

`s2 ~ Gaussian(0,10)`

`flip(1- $\Phi_{s2-s1,5\sqrt{2}}$ (1))`

`return (s1,s2)`

# 하고 있는 일

- 이 최적화를 일반화하기 - loop, branch가 존재하는 프로그램 상에서
- Mixture Model 위에서의 최적화
- 질 좋은 샘플을 얻기 위한 연구

# 감사합니다

- 질문, 의견 부탁드립니다