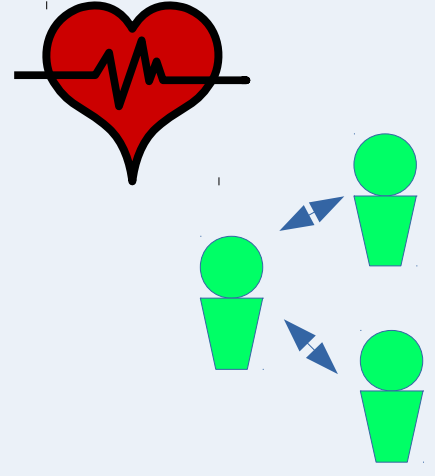


## 1. Motivation

Multivariate categorical distributions

- Medical diagnosis
- Social network analysis
- ...

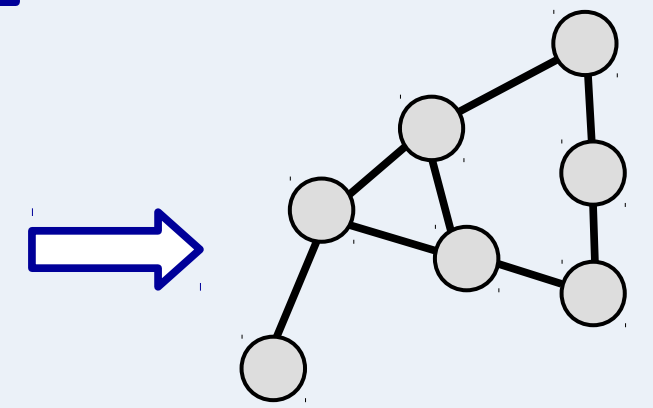
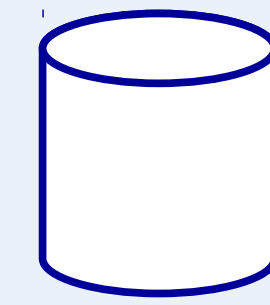
However, learning them is **hard**.  
Performing inference is **hard**.



## 2. Problem statement

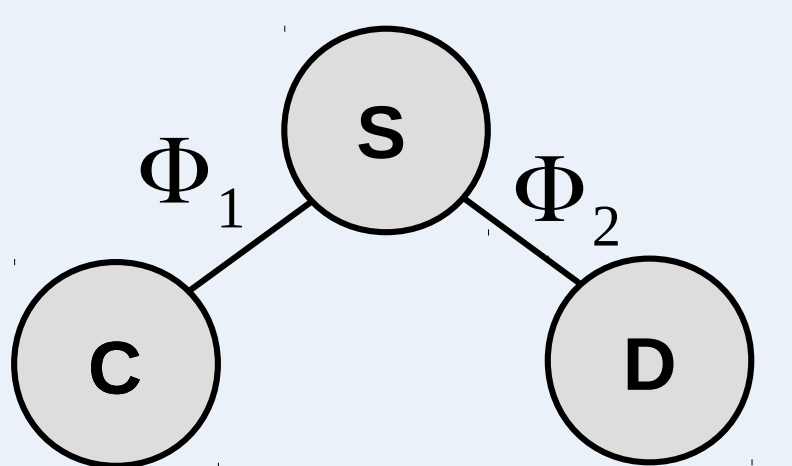
**Goal:** Learn tractable Markov networks from data using genetic algorithms.

- Learn:
  - Structure
  - Parameters
- Tractable
- Genetic algorithms



## 3. Background

**Markov Networks:** Compact representation of distribution!



**C:** eating chicken soup  
**S:** being sick  
**D:** visiting the doctor

$\Phi_i$	$W_i$
$\Phi_1: \neg C \wedge \neg S$	1
$\Phi_2: S \Rightarrow D$	3

$$P(X) = \frac{1}{Z} \prod_{\Phi_i, w_i} \exp(\Phi_i(X) w_i)$$

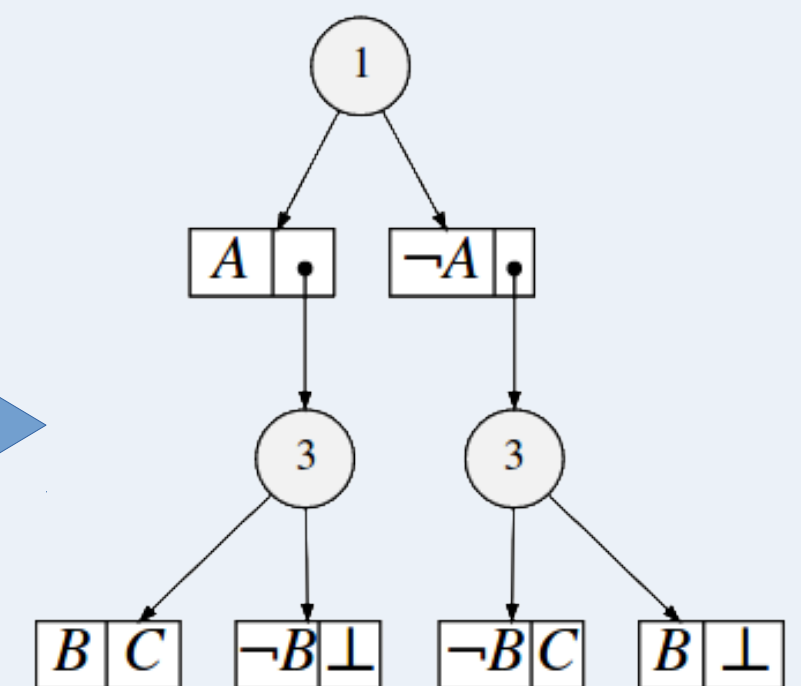
$$Z = \sum_X \prod_{\Phi_i, w_i} \exp(\Phi_i(X) w_i) \quad !! \quad \text{Sad face icon}$$

Encoding

Theory

$$(A \Leftrightarrow B) \wedge C$$

Knowledge compilation



Can compute Z in  $O(|SDD|)$



## 4. Approach

### 1) Fitness

What to consider:

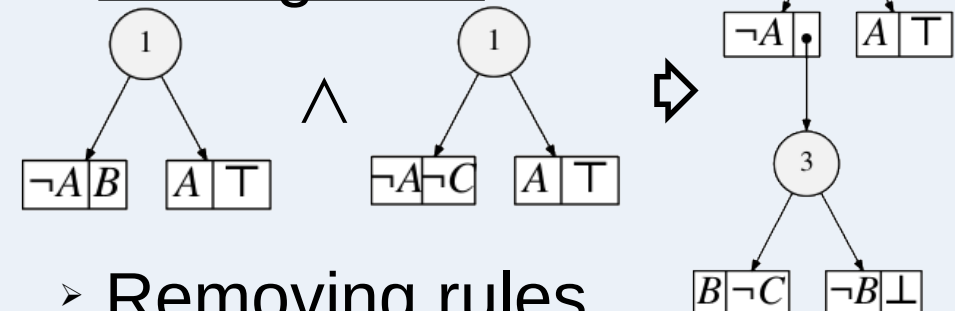
- Data fit
- Size of SDD

$$f(SDD) = LL(SDD) - \alpha \log(|SDD|)$$

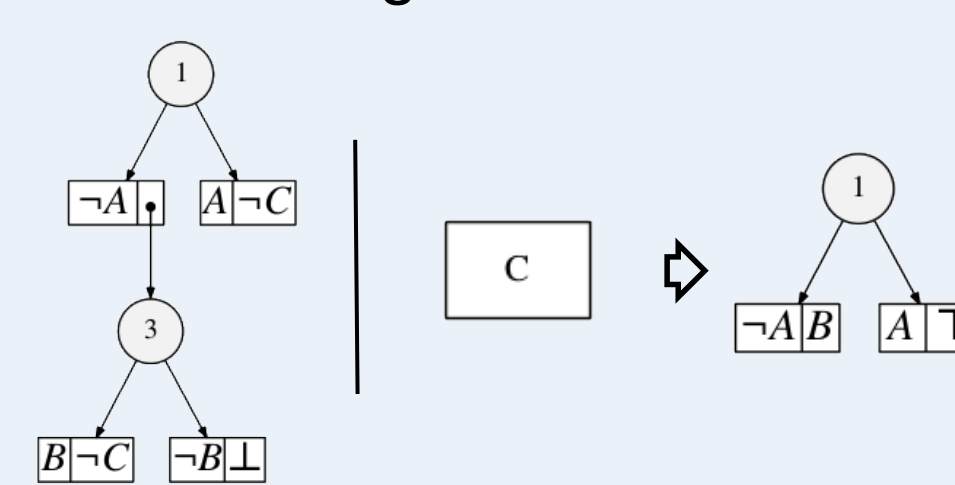
- $LL(SDD)$ : Log-likelihood
- $|SDD|$ : Size of the SDD
- $\alpha$ : Trade-off parameter

### 2) Mutation

➢ Adding rules



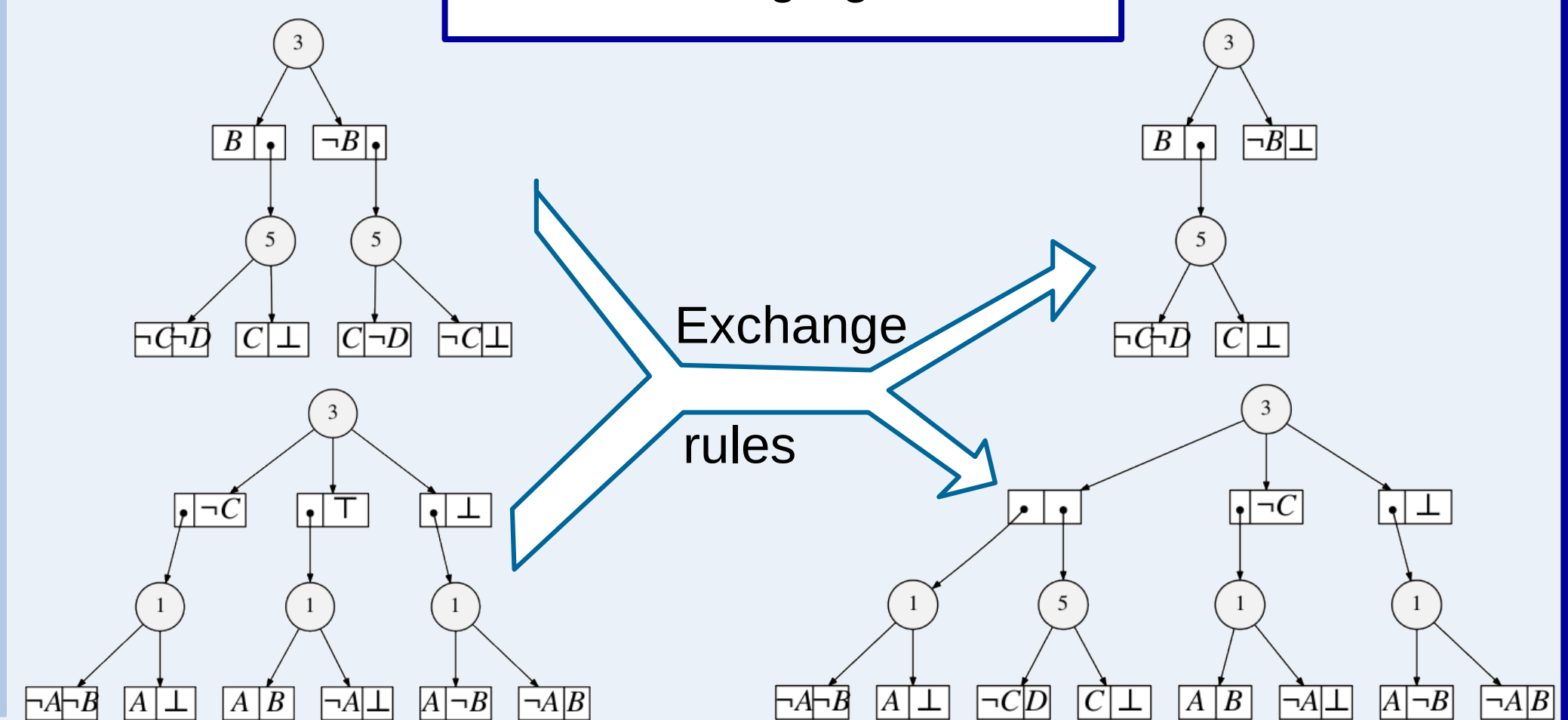
➢ Removing rules



➢ Altering rules: Add/remove

### 3) Crossover

Idea: exchange good rules!



**4) Weight learning:** Maximum Likelihood

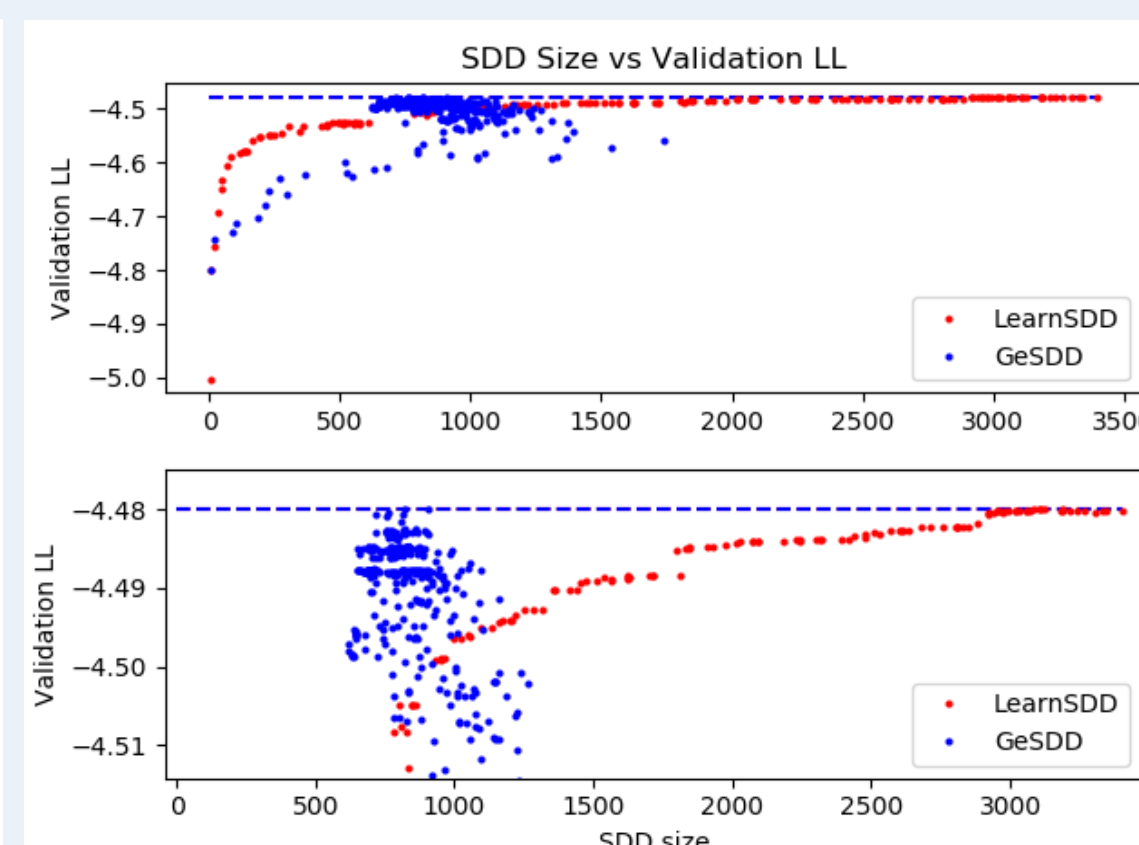
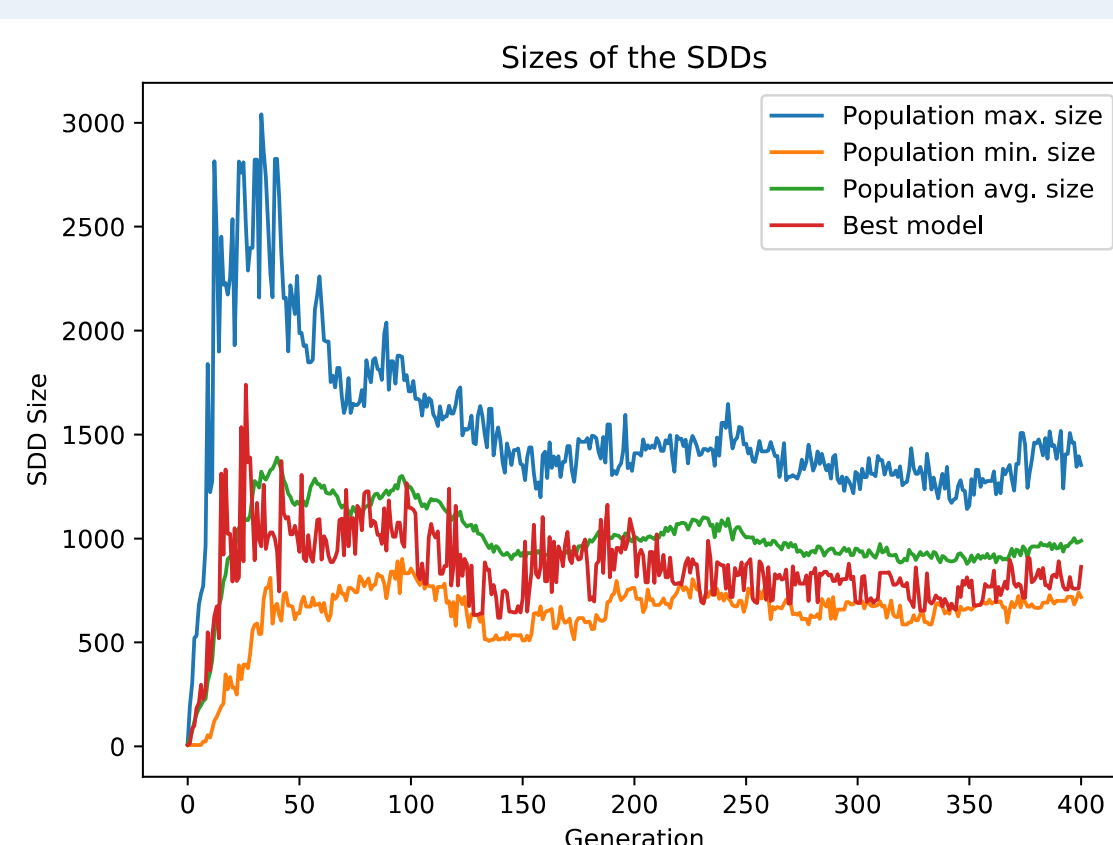
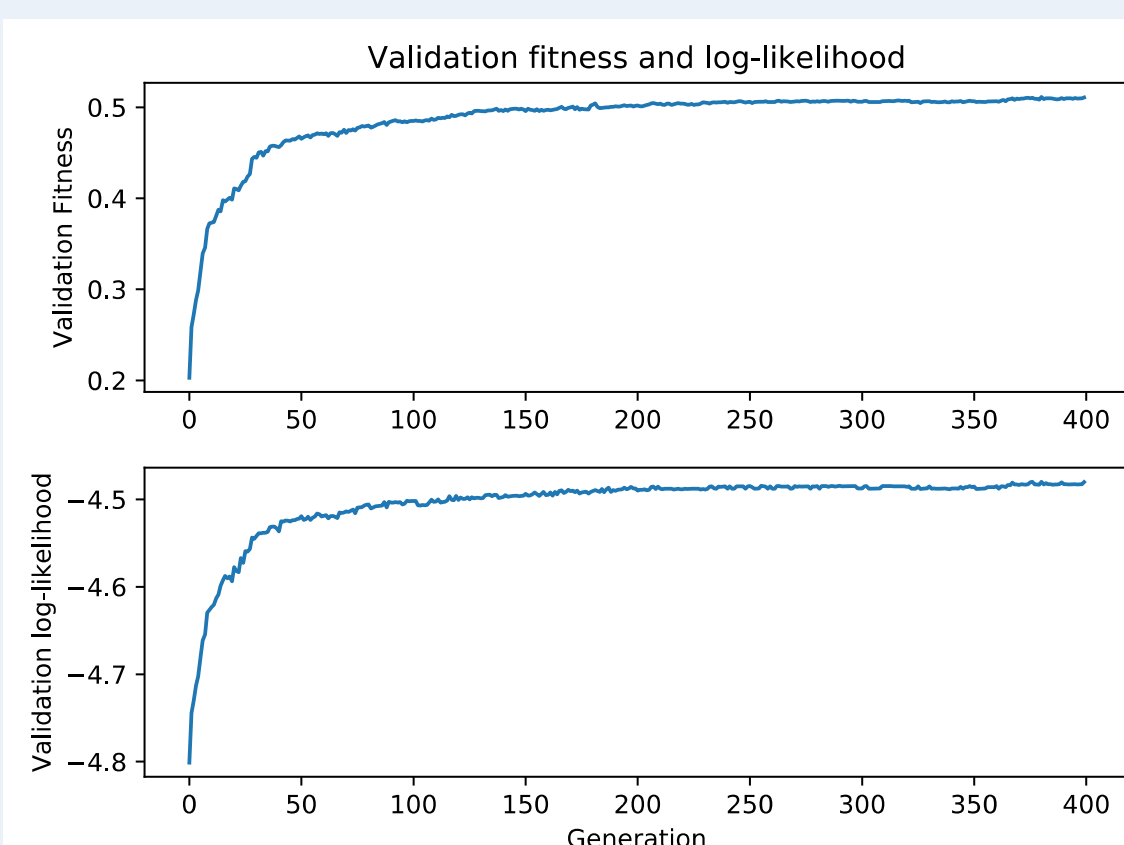
$$\arg \max_w \sum_{\Phi_i, w_i} \text{count}(\Phi_i) * w_i - n \log(Z)$$



**No closed form:** BFGS based method

## 5. Results

**Comparison:** LearnSDD, greedy learning.



**Main takeaways:**

**Results:**

- Comparable LL
- Better Size

**Trade-off:**

- Offline learning
- Online using