Inferência em Redes de Crenças

- Sejam Z_i pais de Y_i e z_i, um conj de valores para Z_i.

  $P(E_X^- \mid X) = \prod_i P(E_{Y_i|X} \mid X)$

  $P(E_X^- \mid X) = \prod_i \sum_{y_i} \sum_{z_i} P(E_{Y_i|X} \mid X, y_i, z_i) P(y_i, z_i \mid X)$

- Decompondo $E_{Y_i|X}$ em dois componentes independentes $E_{Y_i}^+$ e $E_{Y_i|X}^-$

  $P(E_X^- \mid X) = \prod_i \sum_{y_i} \sum_{z_i} P(E_{Y_i}^- \mid X, y_i, z_i) P(E_{Y_i|X}^+ \mid X, y_i, z_i) P(y_i, z_i \mid X)$

  $P(E_X^- \mid X) = \prod_i \sum_{y_i} P(E_{Y_i}^- \mid y_i) \sum_{z_i} P(E_{Y_i|X}^+ \mid z_i) P(y_i, z_i \mid X)$

Inferência em Redes de Crenças

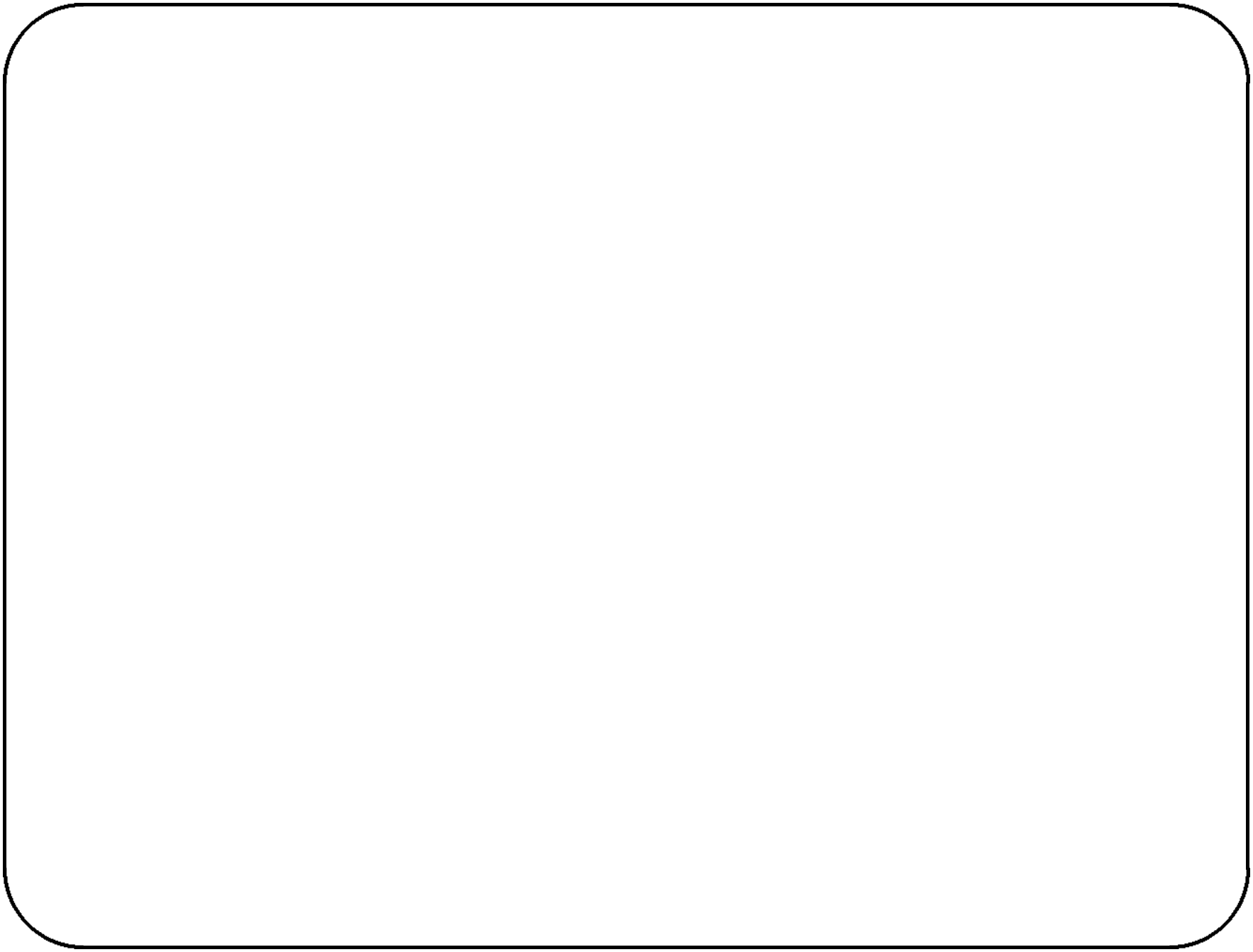
- Aplicando Bayes a $P(E^{+}_{Y_i|X} \mid z_i)$:

  $P(X \mid E^{-}_{X}) = \prod_i \sum_{y_i} P(E^{-}_{Y_i} \mid y_i) \sum_{z_i} \frac{P(z_i|E^{+}_{Y_i|X})P(E^{+}_{Y_i|X})}{P(z_i)} P(y_i, z_i \mid X)$

  ...

  $P(X \mid E^{-}_{X}) = \beta \prod_i \sum_{y_i} P(E^{-}_{Y_i} \mid y_i) \sum_{z_i} P(y_i \mid X, z_i) \prod_{z_i} P(z_{ij} \mid E_{Z_{ij}|Y_i})$

- $P(E^{-}_{Y_i} \mid y_i)$ é uma instância recursiva de $P(E^{-}_{X} \mid X)$.
- $P(y_i \mid X, z_i)$ é tirada diretamente da TPC de Y_i.
- $P(z_{ij} \mid E_{Z_{ij}|Y_i})$ é uma instância recursiva de $P(X \mid E)$.



function BELIEF-NET-ASK(X) **returns** a probability distribution over the values of X
inputs: X, a random variable

SUPPORT-EXCEPT(X, *null*)

---

function SUPPORT-EXCEPT(X, V) **returns** $\mathbf{P}(X|E_{X\setminus V})$

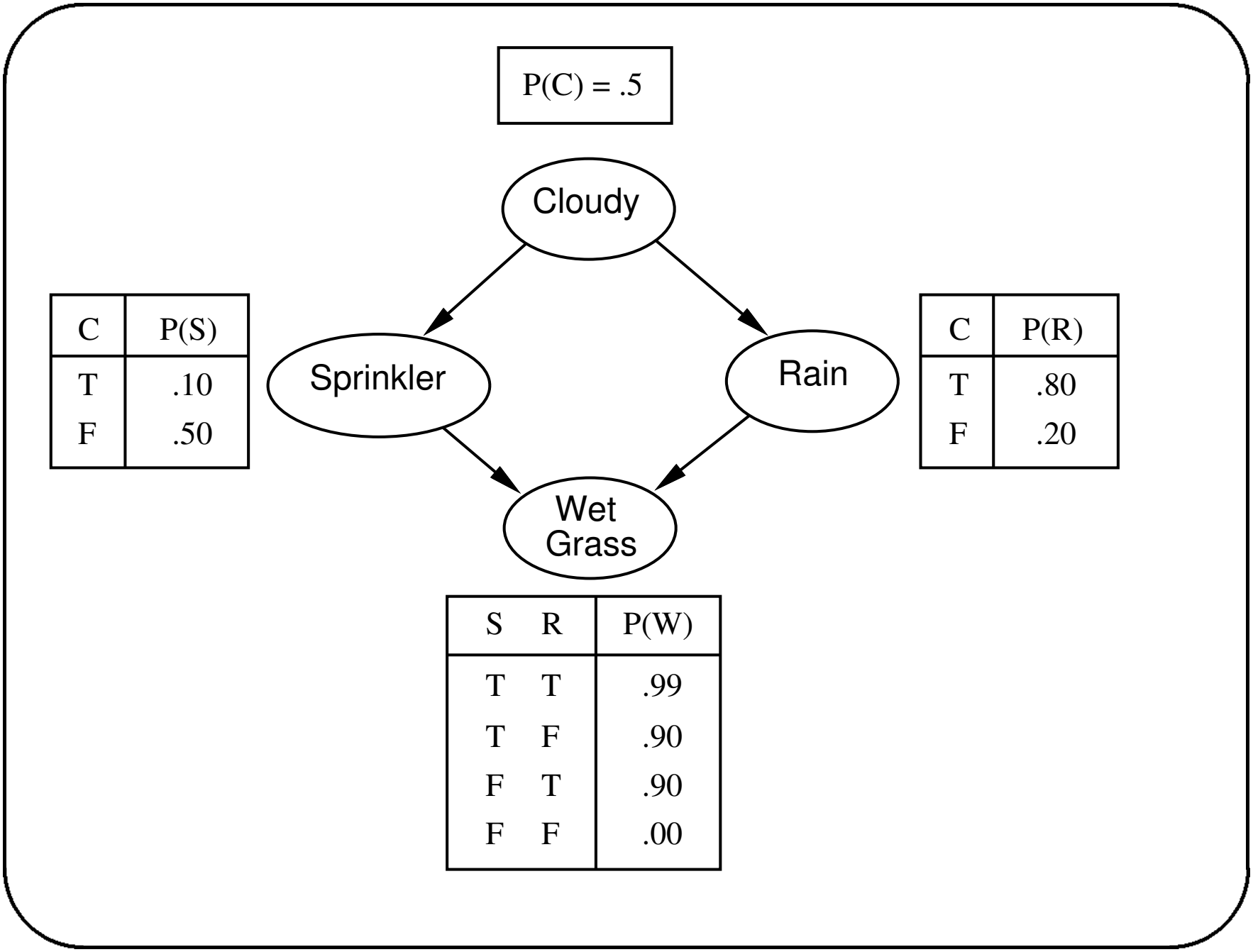
if EVIDENCE?(X) **then return** observed point distribution for X
else
  calculate $\mathbf{P}(E^-_{X\setminus V}|X)$ = EVIDENCE-EXCEPT(X, V)
  $\boldsymbol{U} \leftarrow$ PARENTS[X]
  if U is empty
    then return $\alpha\ \mathbf{P}(E^-_{X\setminus V}|X)\ \mathbf{P}(X)$
  else
    for each U_i **in** U
      calculate and store $\mathbf{P}(U_i|E_{U_i\setminus X})$ = SUPPORT-EXCEPT(U_i, X)
    return $\alpha\ \mathbf{P}(E^-_{X\setminus V}|X)\ \sum_{\mathbf{u}}\ \mathbf{P}(X|\mathbf{u})\ \prod_i\ \mathbf{P}(U_i|E_{u_i\setminus X})$

---

function EVIDENCE-EXCEPT(X, V) **returns** $\mathbf{P}(E^-_{X\setminus V}|X)$

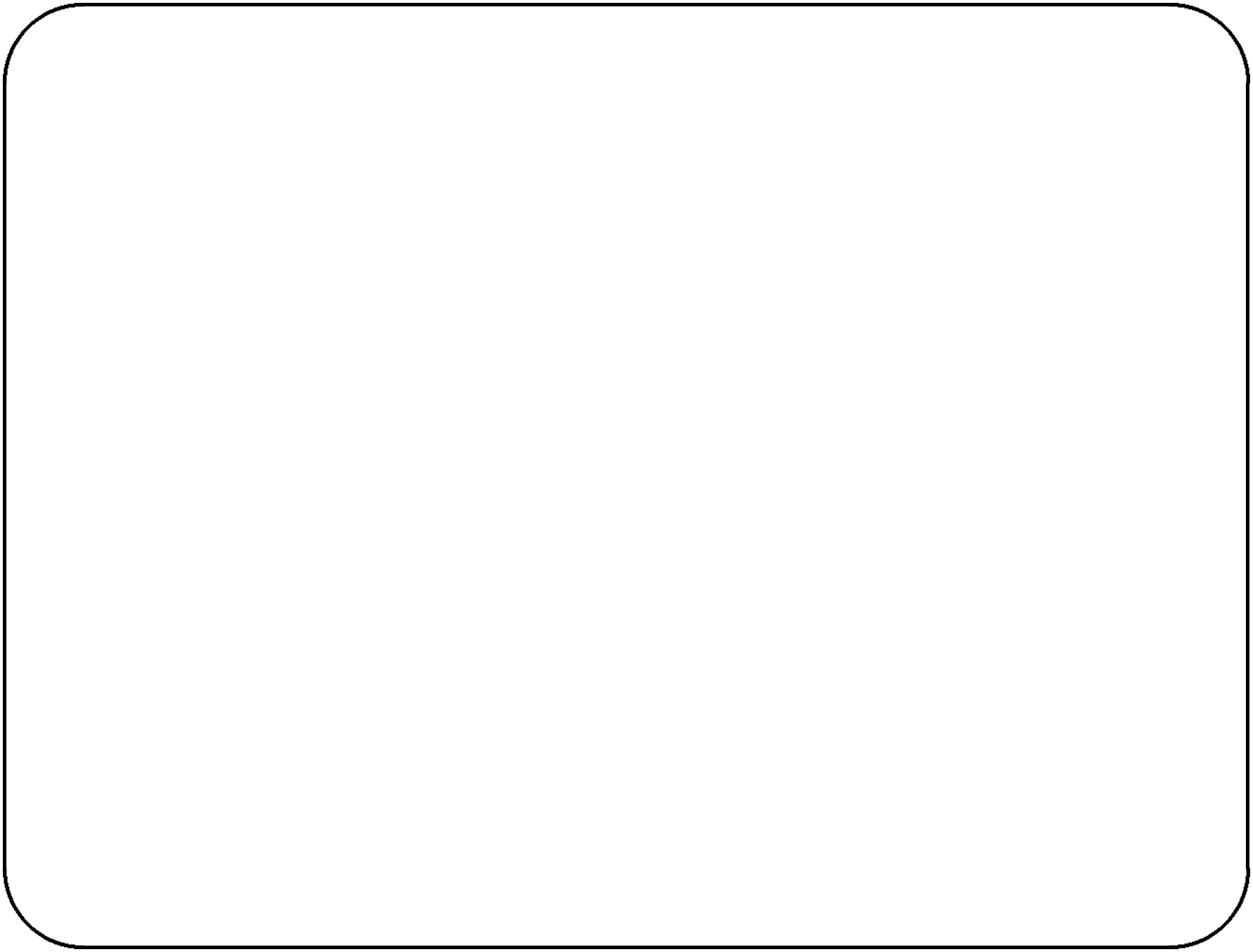
$\mathbf{Y} \leftarrow$ CHILDREN[X] $-$ V
if Y is empty
  then return a uniform distribution
else
  for each Y_i **in Y do**
    calculate $\mathbf{P}(E^-_{Y_i}|y_i)$ = EVIDENCE-EXCEPT(Y_i, null)
    $\mathbf{Z}_i \leftarrow$ PARENTS[Y_i] $-$ X
    for each Z_{ij} **in** $\mathbf{Z}_i$
      calculate $\mathbf{P}(Z_{ij}|E_{Z_{ij}\setminus Y_i})$ = SUPPORT-EXCEPT(Z_{ij}, Y_i)
  return $\beta\ \prod_i\ \sum_{y_i}\ P(E^-_{Y_i}|y_i)\ \sum_{\mathbf{z}_i}\ \mathbf{P}(y_i|X, \mathbf{z}_i)\ \prod_j\ P(z_{ij}|E_{Z_{ij}\setminus Y_i})$

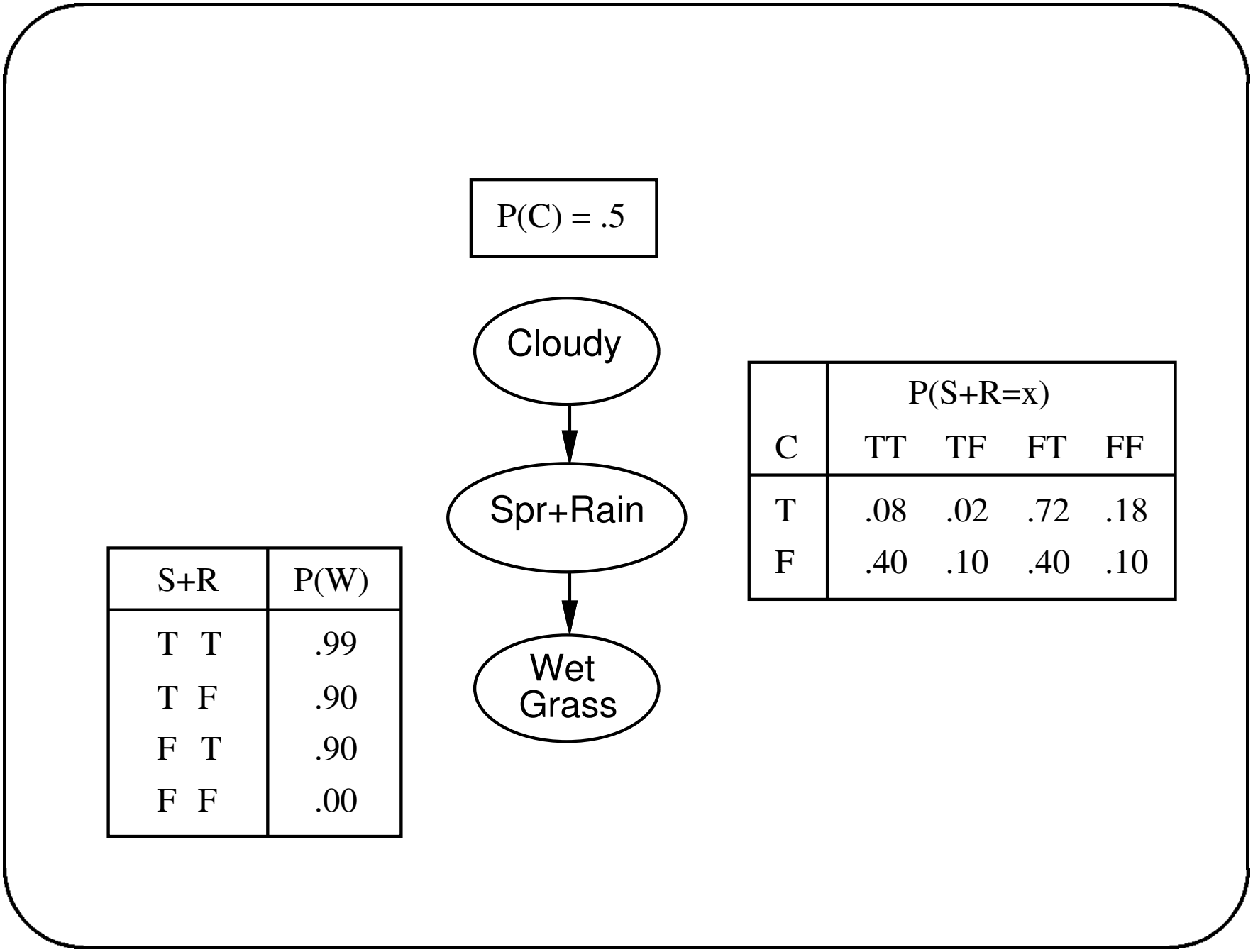
Inferência em Redes de Crenças Multiplamente conectadas


P(C) = .5
Cloudy
Sprinkler
Rain
Wet Grass
C P(S)
T .10
F .50
C P(R)
T .80
F .20
S R P(W)
T T .99
T F .90
F T .90
F F .00

Inferência em Redes de Crenças Multiplamente conectadas

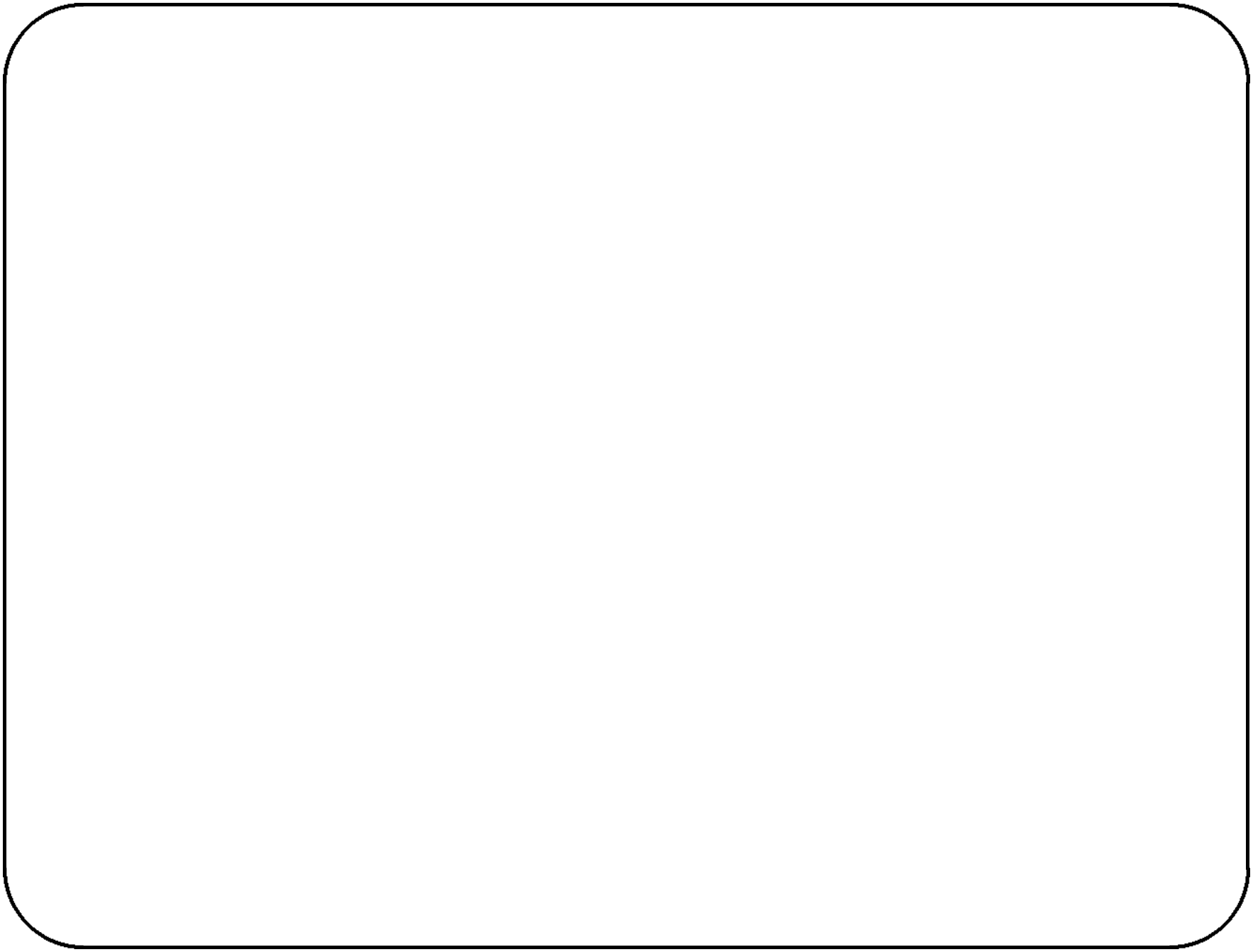
- Três classes de algoritmos:
  - **Clustering**: transforma a rede em uma poli-árvore probabilisticamente equivalente, mas com topologia diferente.
  - **Condicionamento**: oposto de clustering, transforma a rede em várias poli-árvores através da instanciação de valores para as variáveis aleatórias. Avalia a poli-árvore para cada instância diferente.
  - **Simulação estocástica** (ou amostragem lógica): calcula uma prob aproximada através de simulações repetidas do mundo descrito pela rede, observando a frequência com que eventos relevantes acontecem.
  - Em geral: inferência exata em redes de crenças é um problema NP-difícil.

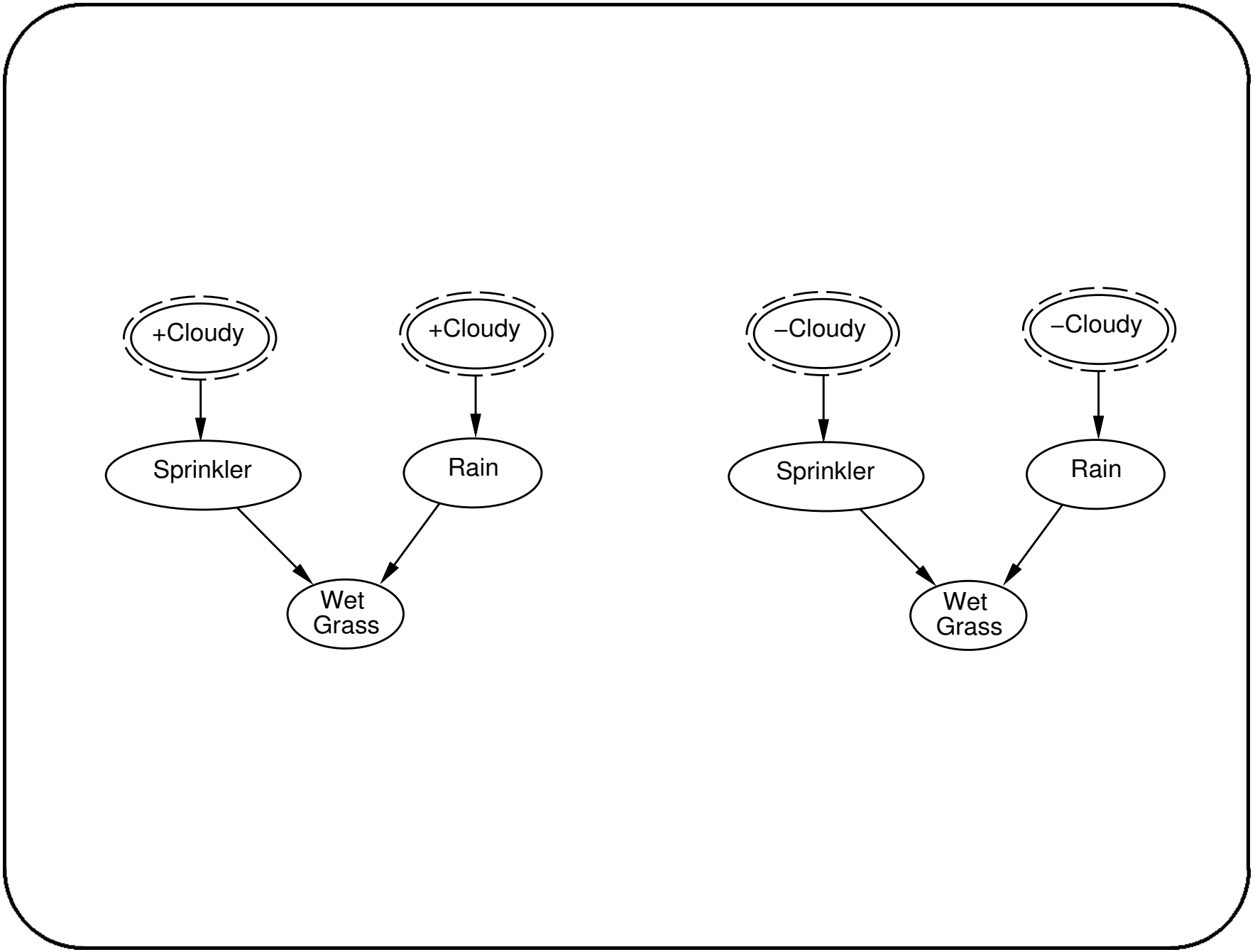




C	P(S+R=x) TT	TF	FT	FF
T	.08	.02	.72	.18
F	.40	.10	.40	.10

S+R	P(W)
T T	.99
T F	.90
F T	.90
F F	.00




+Cloudy
+Cloudy
Sprinkler
Rain
Wet
Grass
–Cloudy
–Cloudy
Sprinkler
Rain
Wet
Grass

Engenharia do Conhecimento para Raciocínio Probabilístico

- Decidir sobre o que falar.
- Decidir sobre o vocabulário de variáveis aleatórias.
- Codificar o conhecimento sobre as relações de dependências entre as variáveis.
- Codificar uma descrição de uma instância específica do problema.
- Colocar consultas ao procedimento de inferência.
- Estudo de caso: Pathfinder, leitura p/ casa!

Outras Abordagens

- Por que? Falta de credibilidade em métodos probabilísticos para solução de problemas devido ao crescimento exponencial das tabelas e algoritmos de redes de crenças ainda não conhecidos.
- Probabilidade é numérica (!). Raciocício humano é mais "qualitativo" que "quantitativo". Principal abordagem qualitativa: **default reasoning**, conclusões não são confiáveis com certo grau, mas acreditadas até que uma razão melhor é encontrada para acreditar em outra coisa.
- Sistemas **baseados em regras**: baseados em sistemas lógicos baseados em regras, mas com anotações nos programas para representar incerteza.

Outras Abordagens

- **Dempster-Shafter theory**: utiliza graus de confiança em intervalos (além de tratar de incerteza, trata de ignorância).
- **Lógica Nebulosa**: representa o conhecimento "vago". Um evento pode ser +/- verdadeiro. Por exemplo, $T(A \wedge B) = min(T(A), T(B))$, onde T é uma função de verdade nebulosa. $T(A \vee \neg A) \neq T(True)$. Ex de linguagem: FRIL (Univ of Bristol, Jim Baldwin, Lotfi Zadeh).