Publieke verdediging

MASTERPROEF

Jasper De Bock

30 juni 2011

EstiHMM: Een efficiënt algoritme ter bepaling van de maximale sequenties in een imprecies hidden Markovmodel

Jasper De Bock

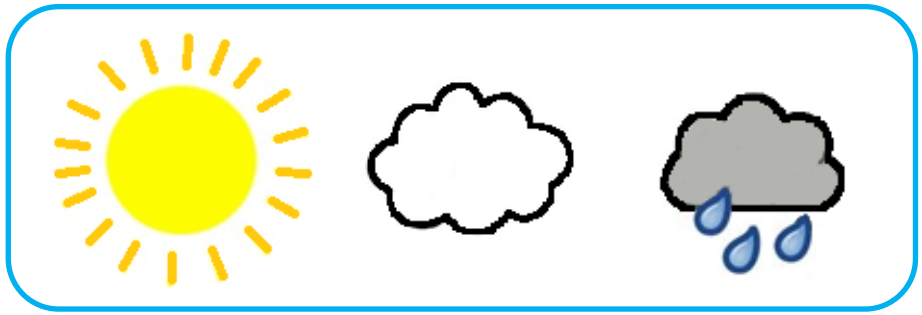
30 juni 2011

EstiHMM: Een efficiënt algoritme ter bepaling van de maximale **sequenties** in een imprecies hidden Markovmodel

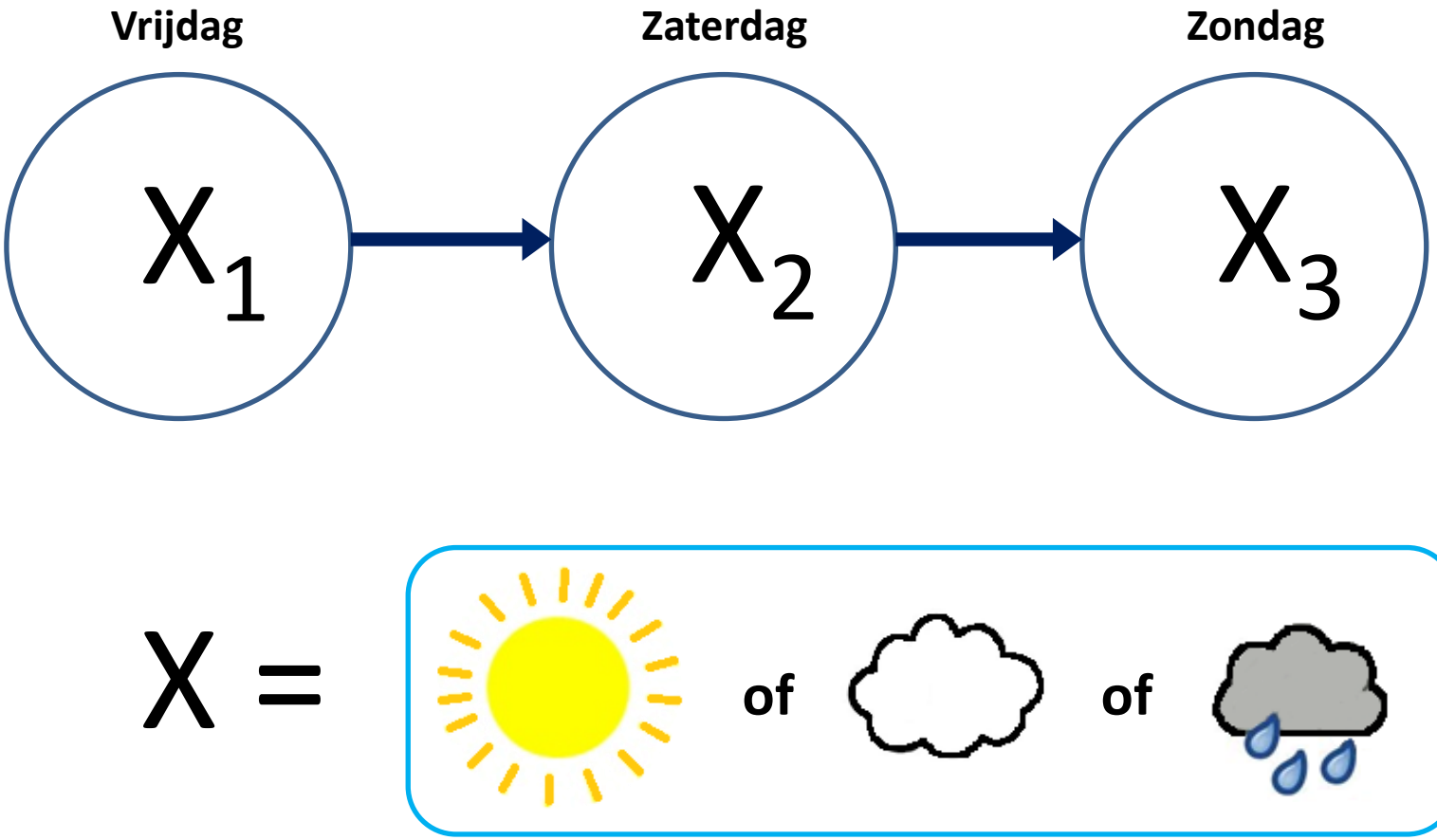
Sequenties

Vrijdag Zaterdag Zondag

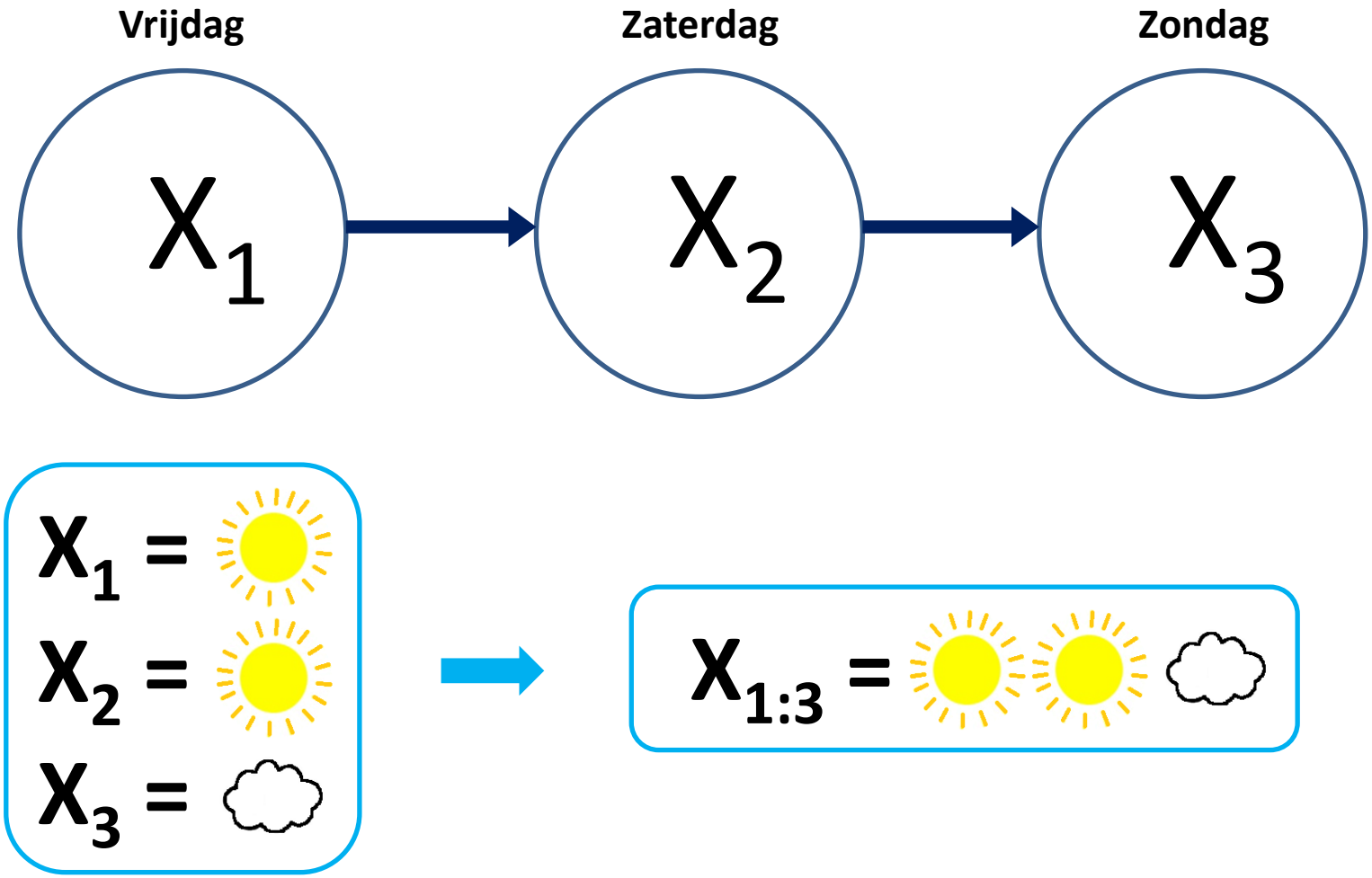
? ? ?



Sequenties



Sequenties



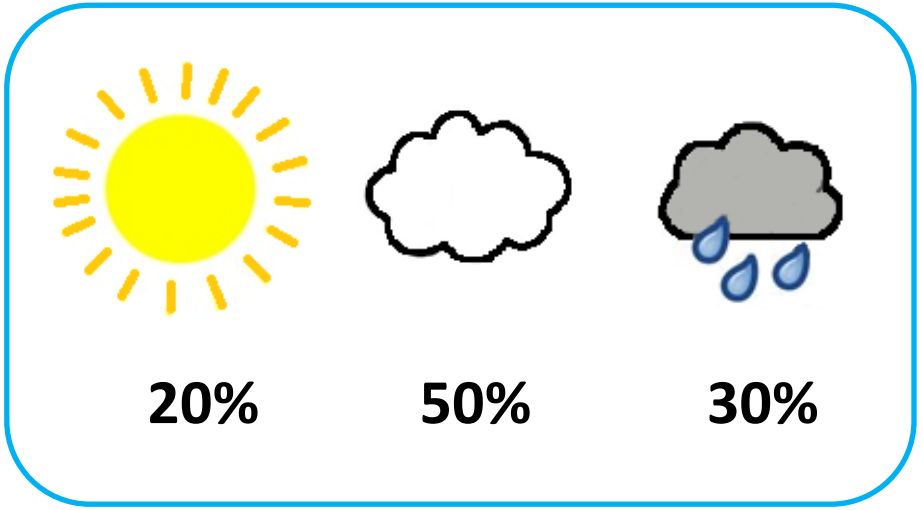
EstiHMM: Een efficiënt algoritme
ter bepaling van de maximale **sequenties**
in een imprecies hidden **Markovmodel**

Markovmodel

Vrijdag

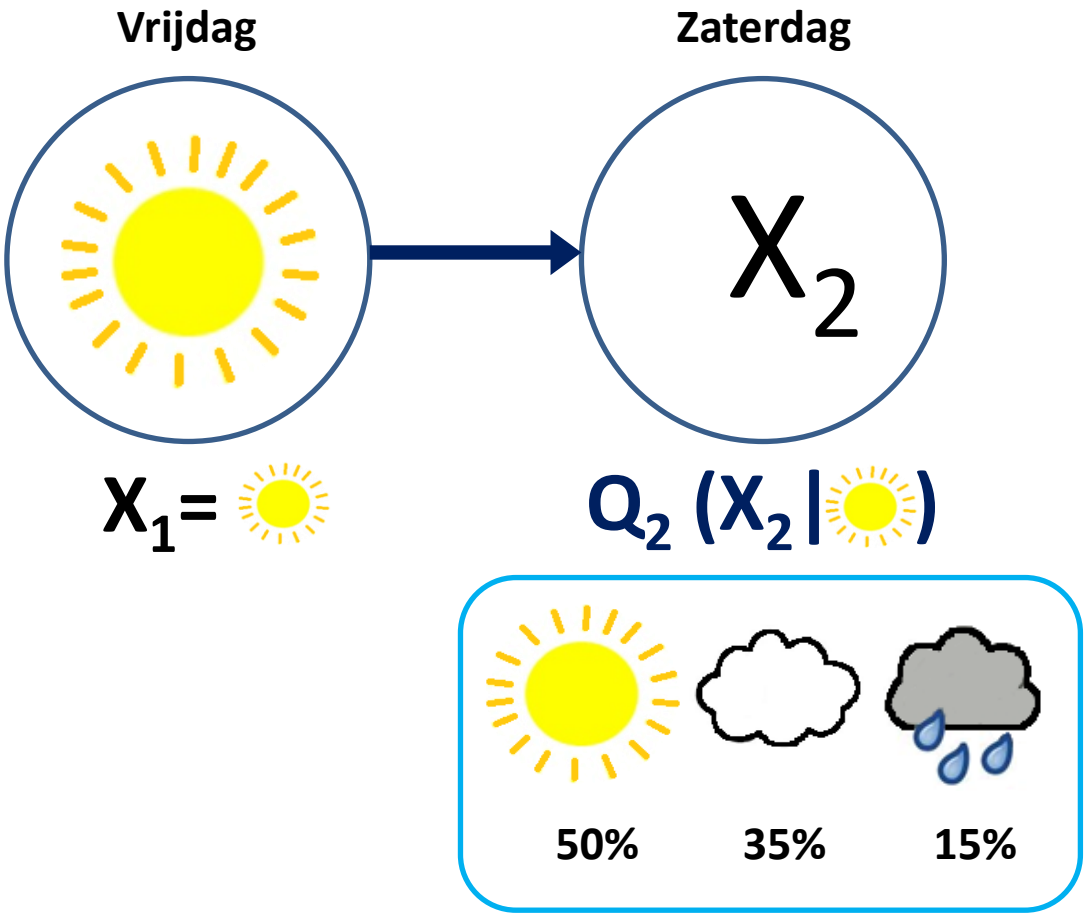
X_1

$Q_1 (X_1)$



lokaal model $Q_1(X_1)$

Markovmodel

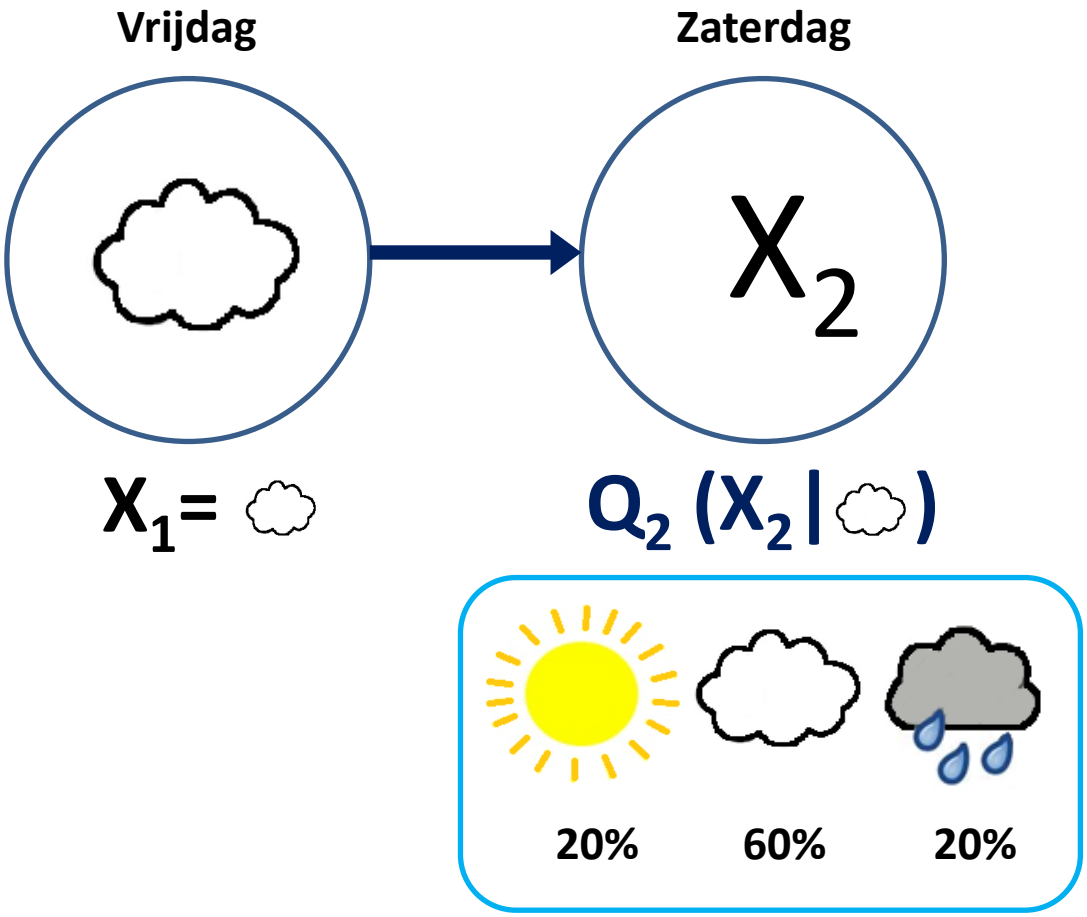


conditioneel

lokaal model

$Q_2 (X_2 | X_1)$

Markovmodel

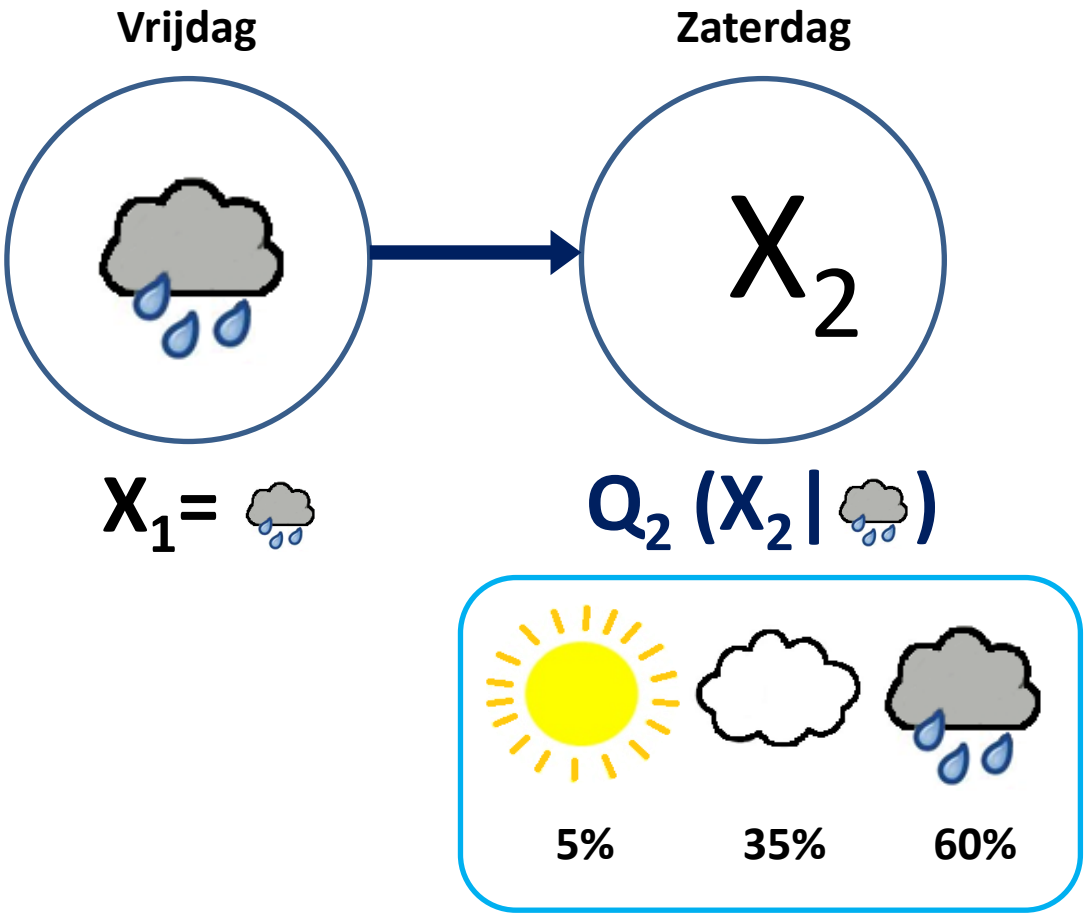


conditioneel

lokaal model

$Q_2 (X_2 | X_1)$

Markovmodel

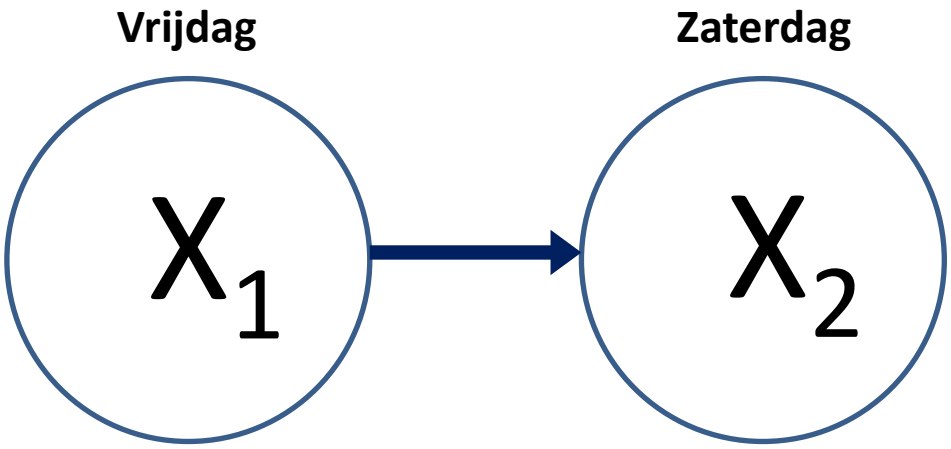


conditioneel

lokaal model

$Q_2(X_2|X_1)$

Markovmodel

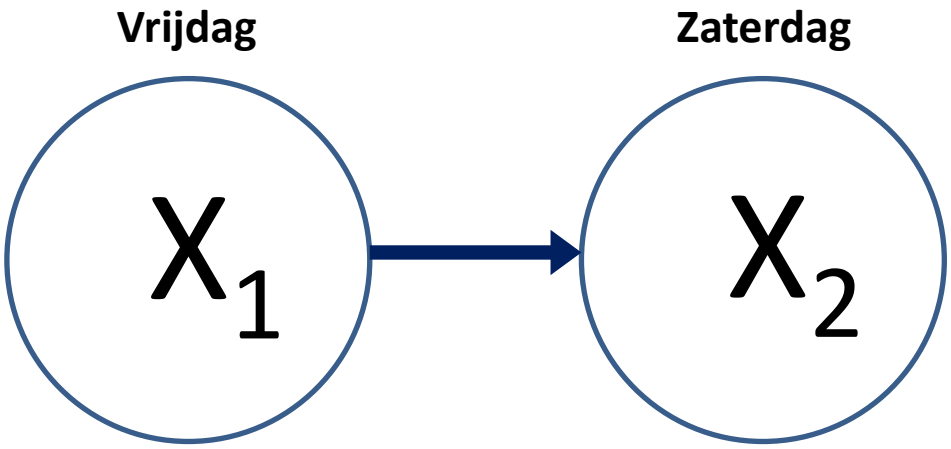


conditioneel

lokaal model

$Q_2 (X_2|X_1)$

Markovmodel



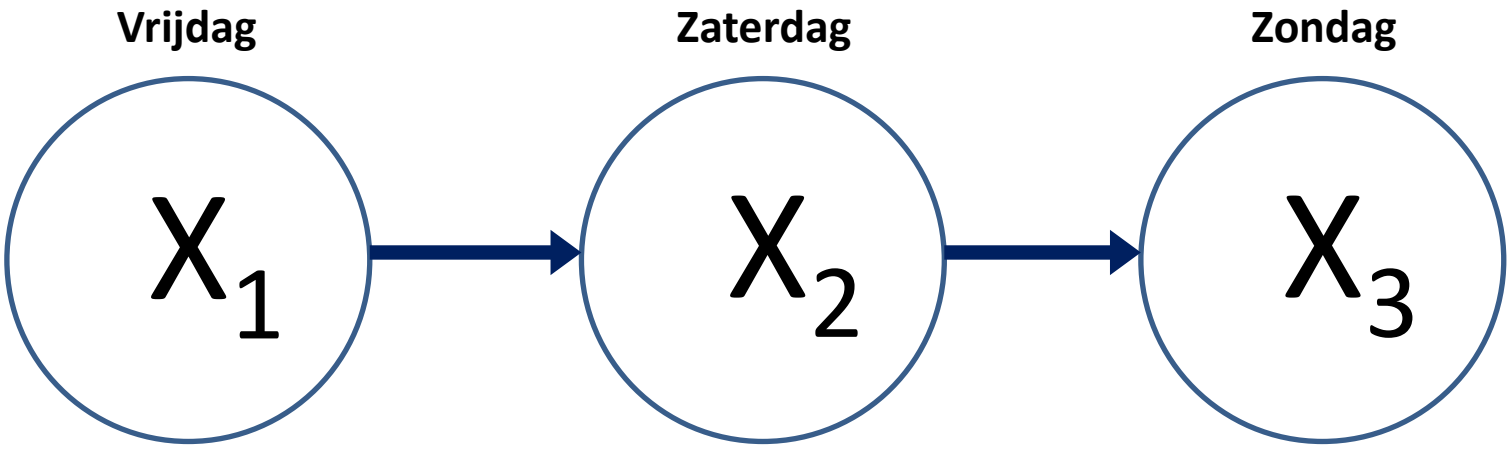
$Q_1 (X_1)$

lokaal model

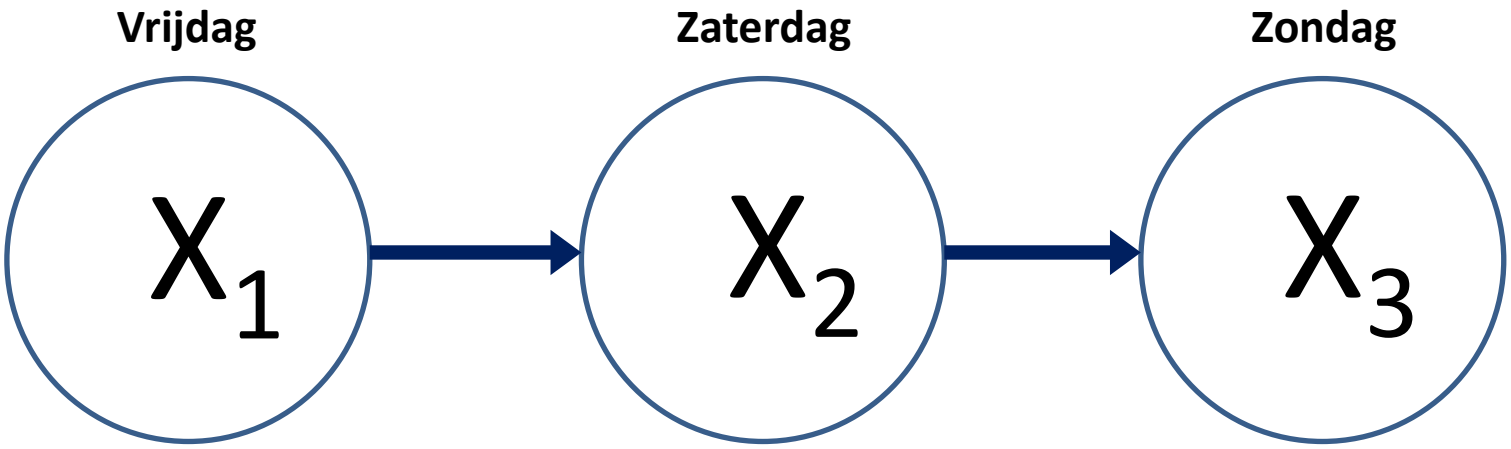
$Q_2 (X_2 | X_1)$

conditioneel lokaal model

Markovmodel



Markovmodel



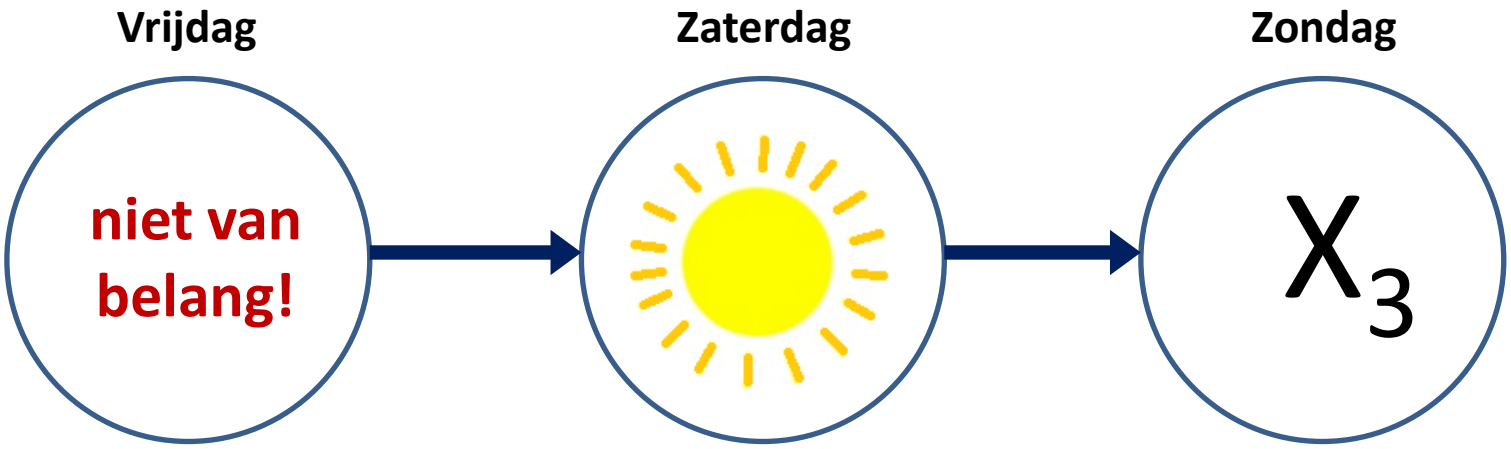
$Q_2 (X_3 | X_2)$

conditioneel lokaal model

Markovvoorwaarde!

Het weer op een bepaalde dag hangt enkel af van het weer de dag voordien

Markovmodel



$Q_2(X_3 | \text{☀})$

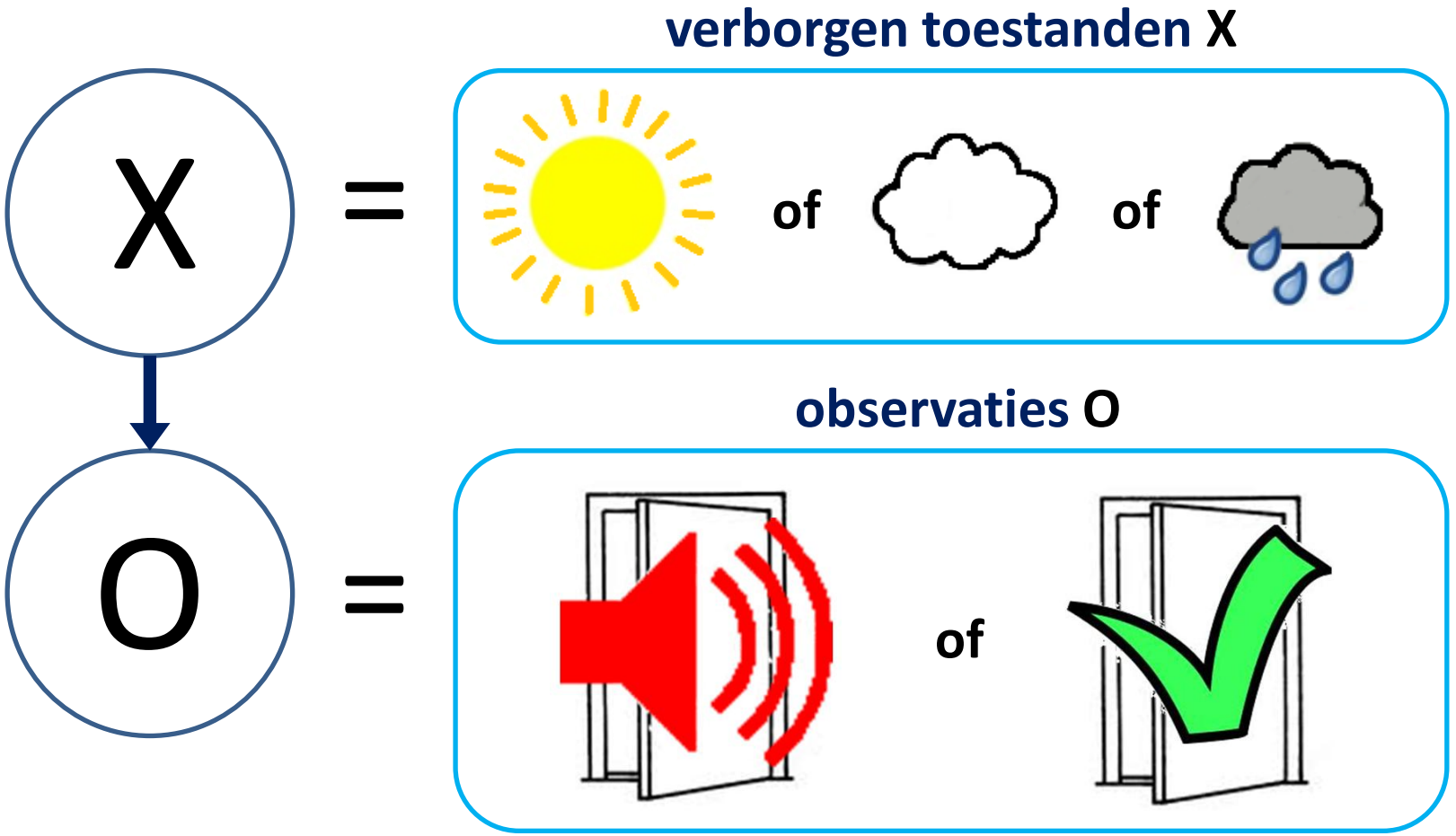
conditioneel lokaal model

Markovvoorwaarde!

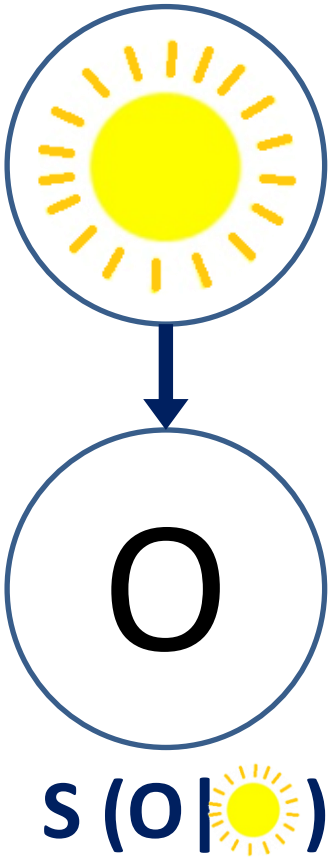
Het weer op een bepaalde dag hangt enkel af van het weer de dag voordien

EstiHMM: Een efficiënt algoritme
ter bepaling van de maximale **sequenties**
in een imprecies **hidden** **Markovmodel**

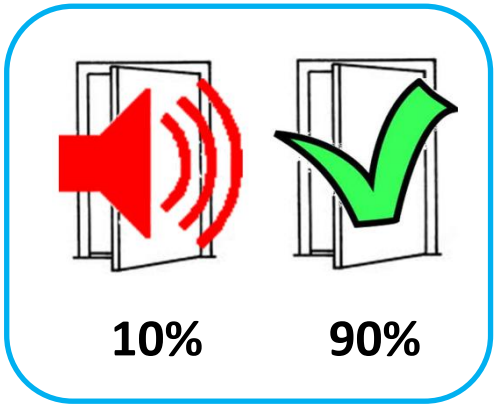
Hidden Markovmodel



Hidden Markovmodel



S (O|☀)

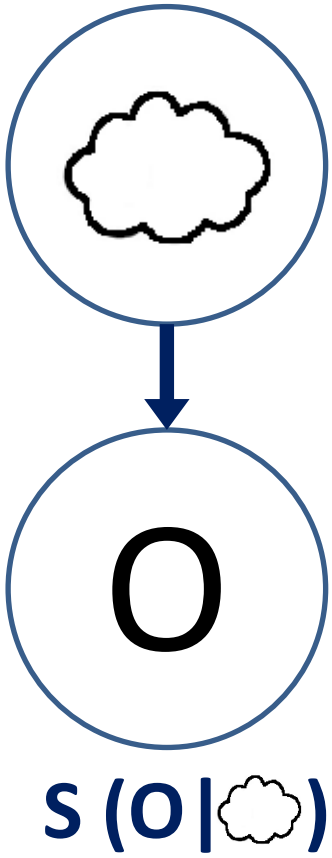


conditioneel

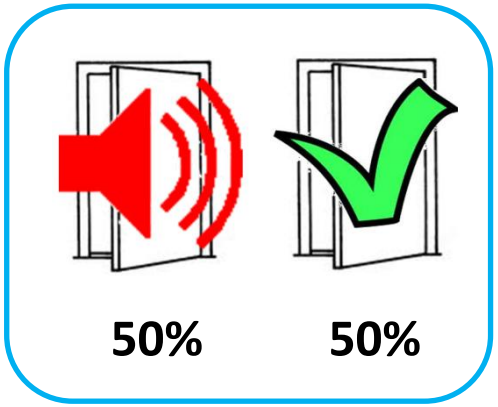
lokaal model

S (O|X)

Hidden Markovmodel



S (O|☁)

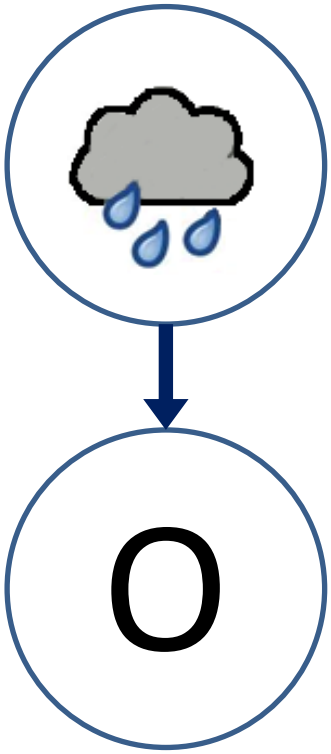


conditioneel

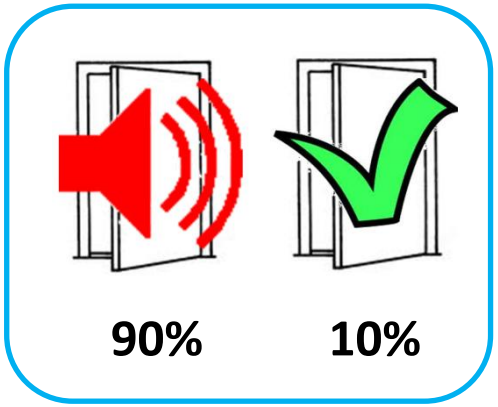
lokaal model

S (O|X)

Hidden Markovmodel



S (O |)

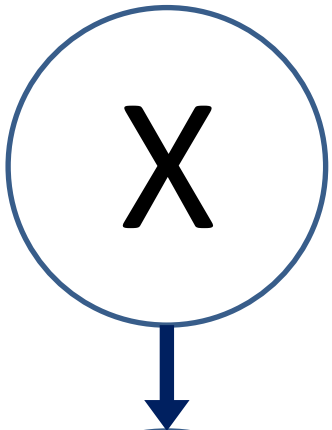


conditioneel

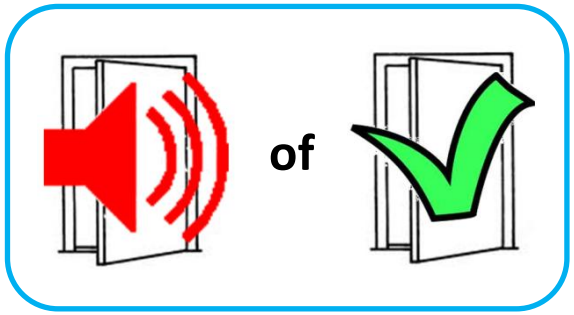
lokaal model

S (O | X)

Hidden Markovmodel



S (O|X)

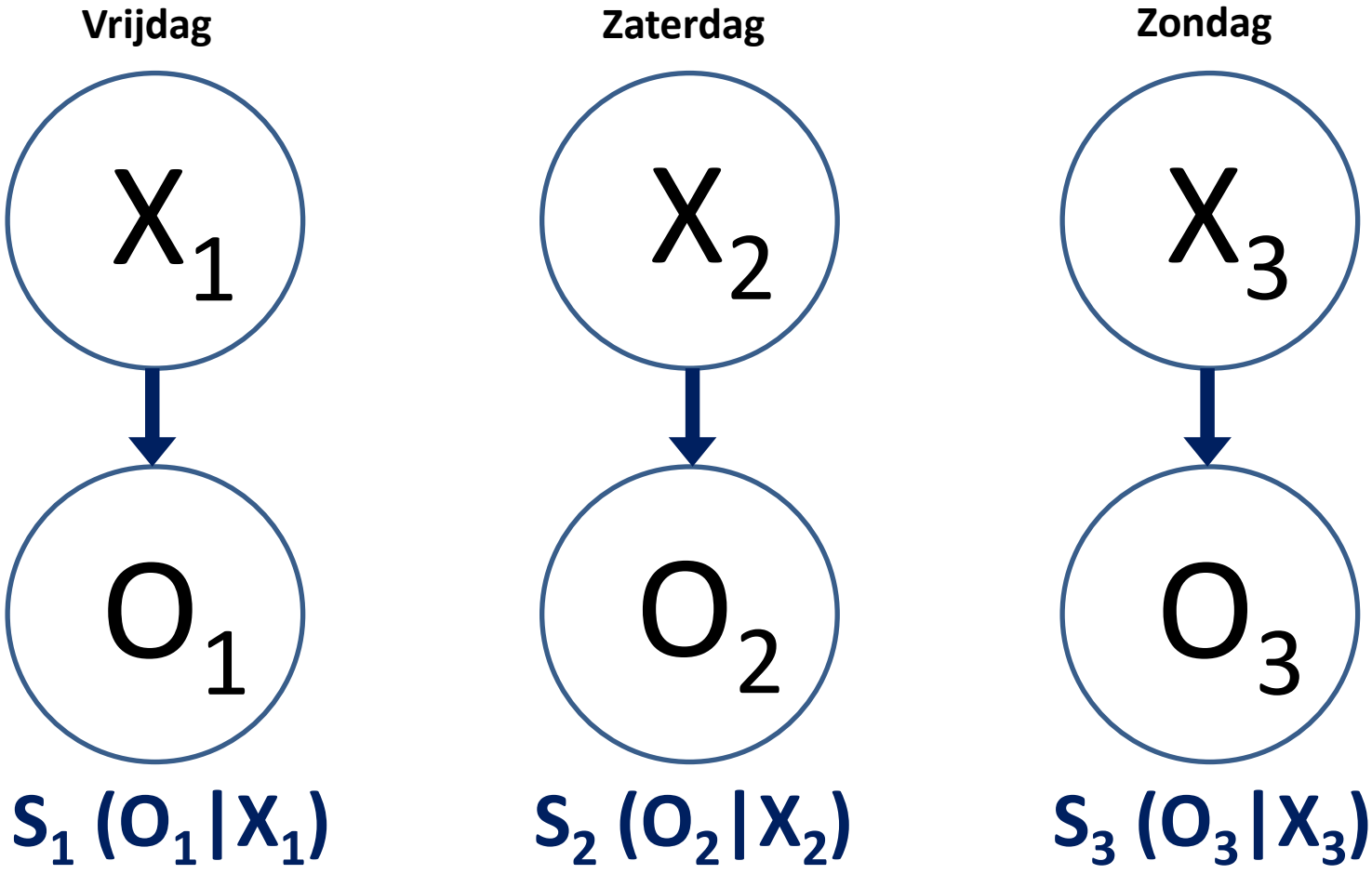


conditioneel

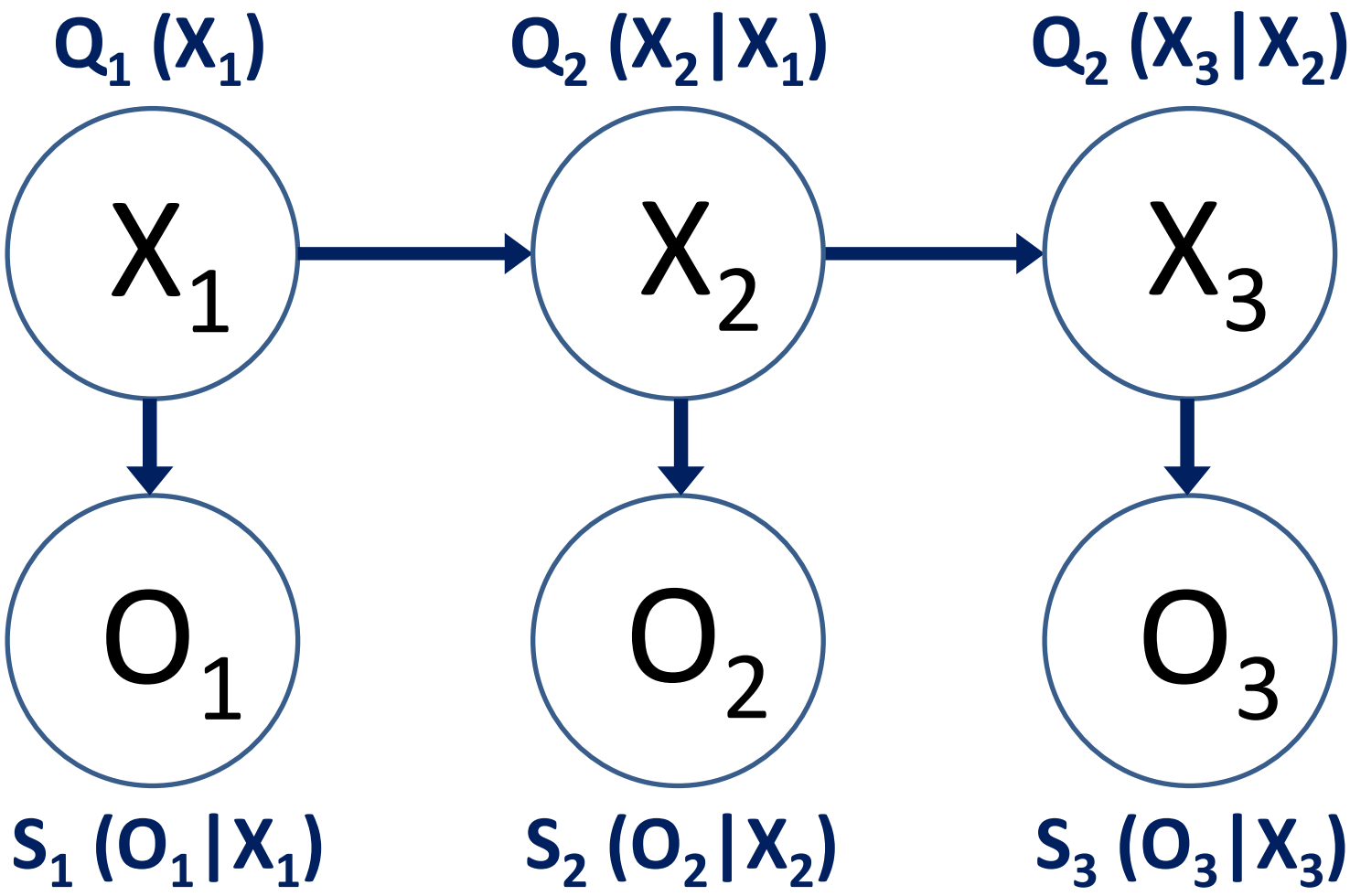
lokaal model

S (O|X)

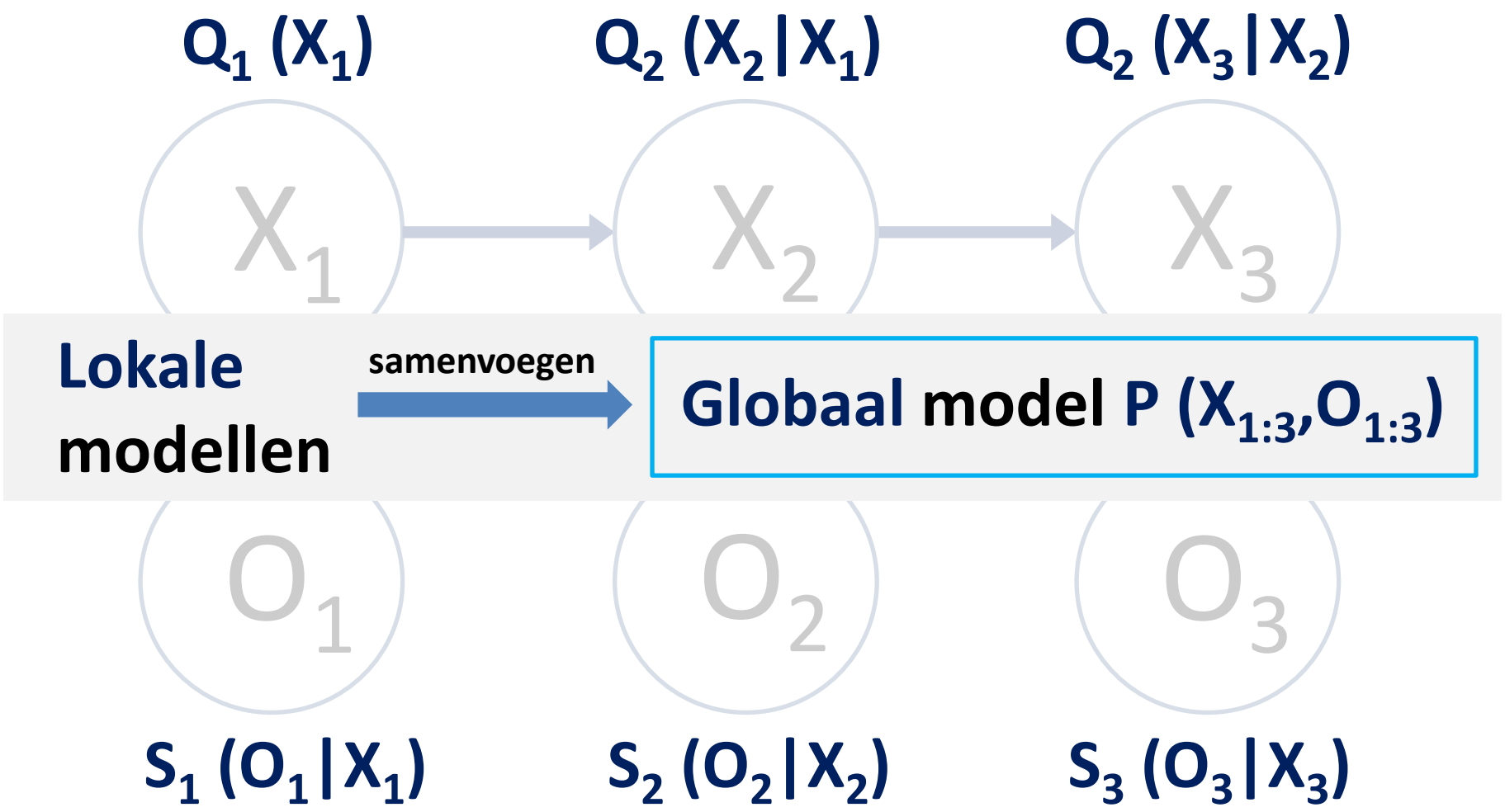
Hidden Markovmodel



Hidden Markovmodel



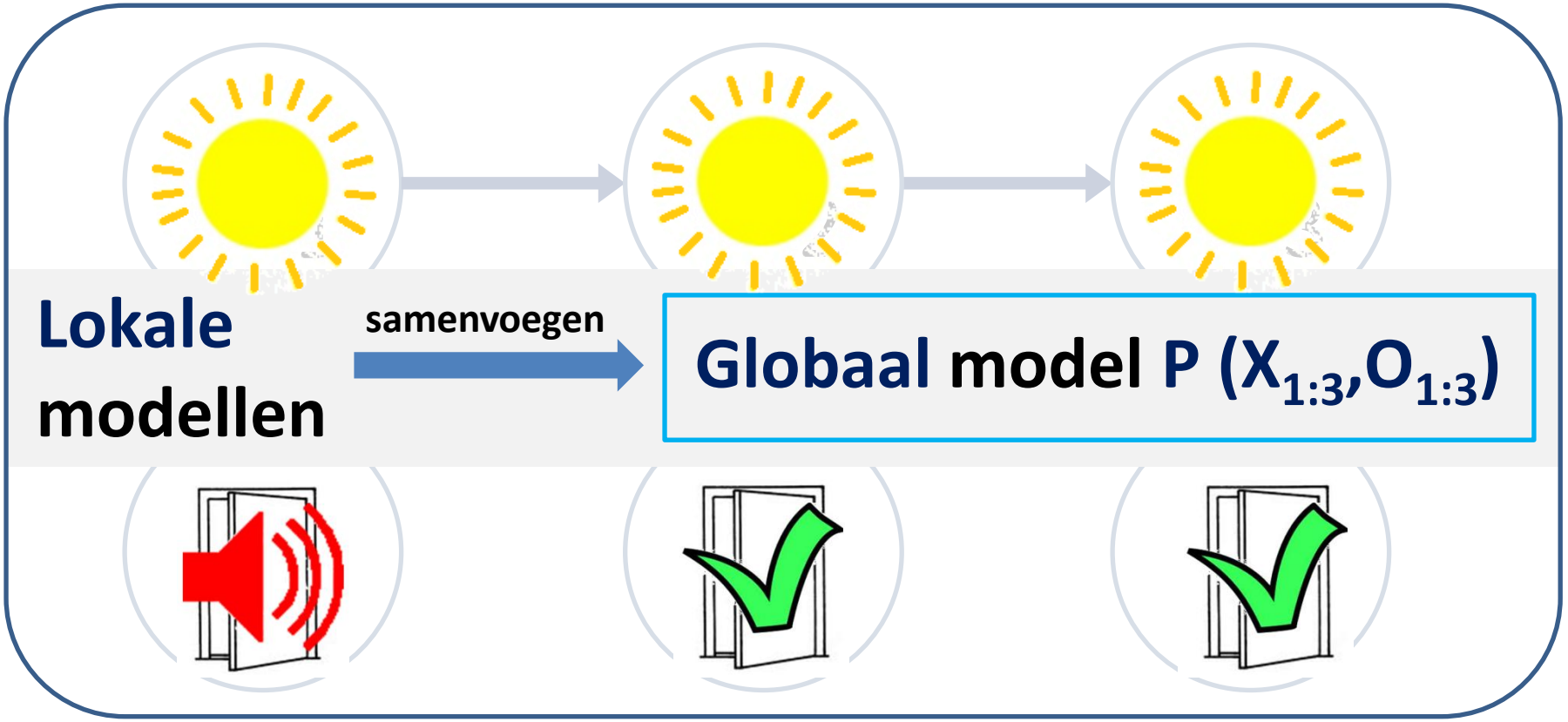
Hidden Markovmodel



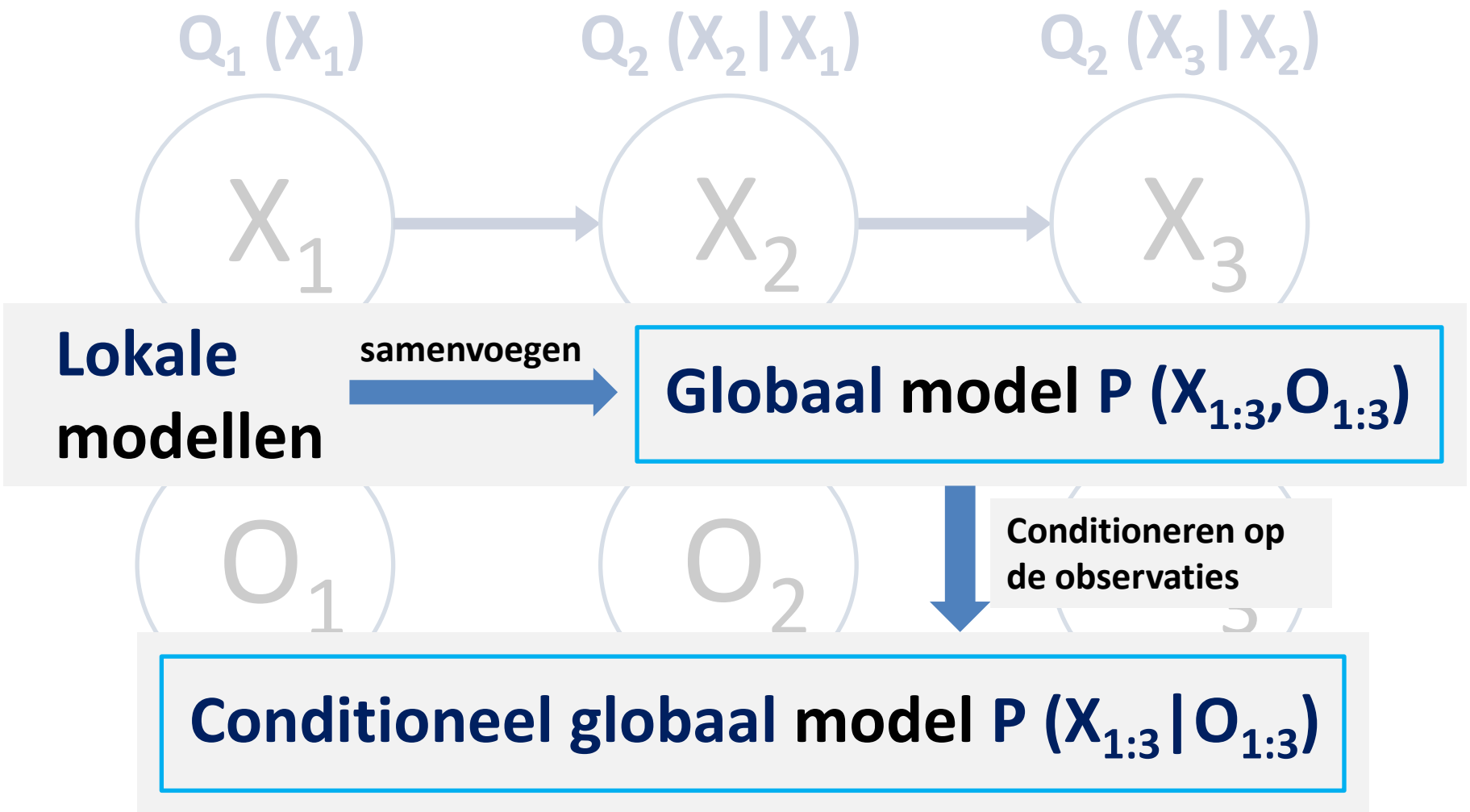
Hidden Markovmodel

P ( , )

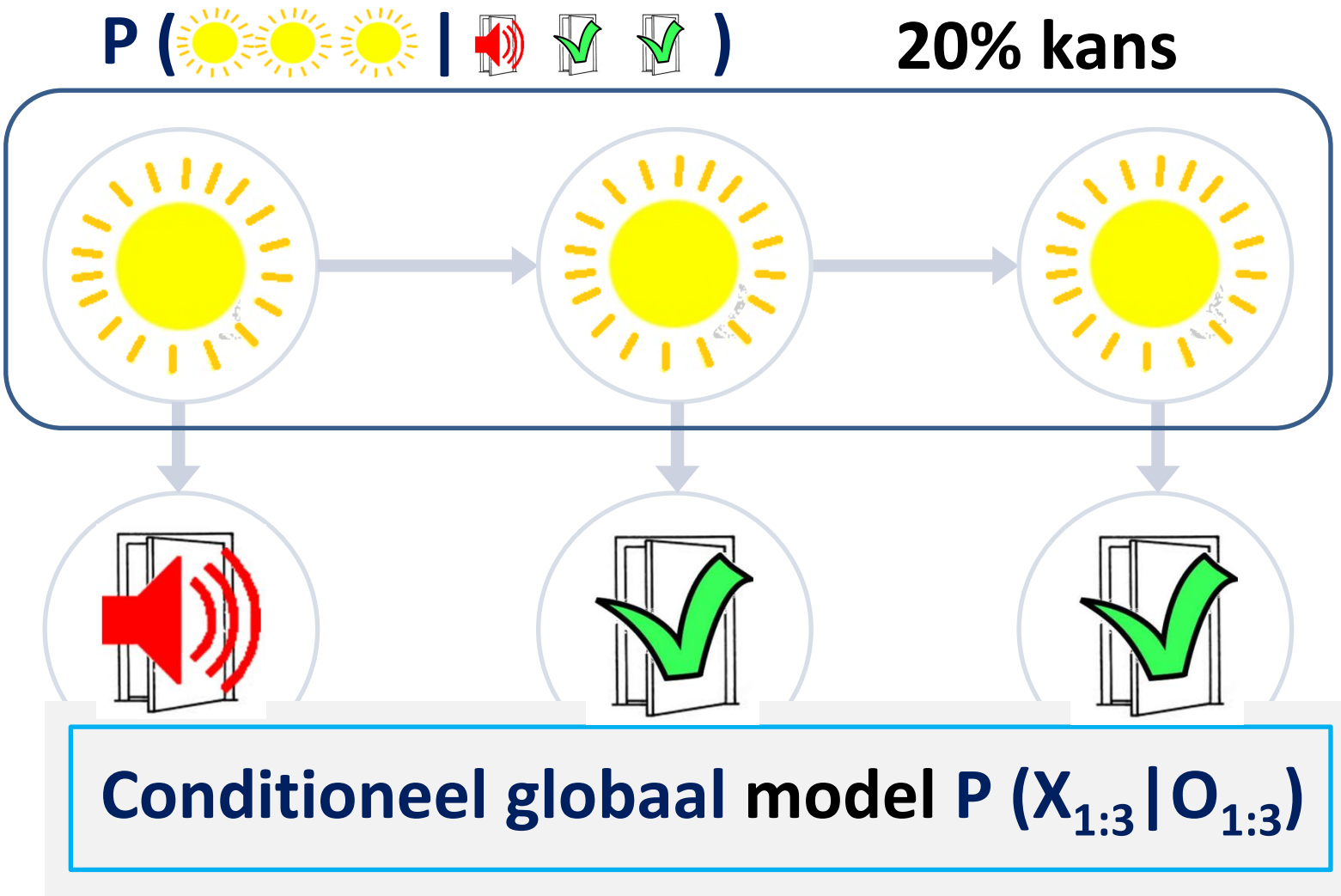
8% kans



Hidden Markovmodel

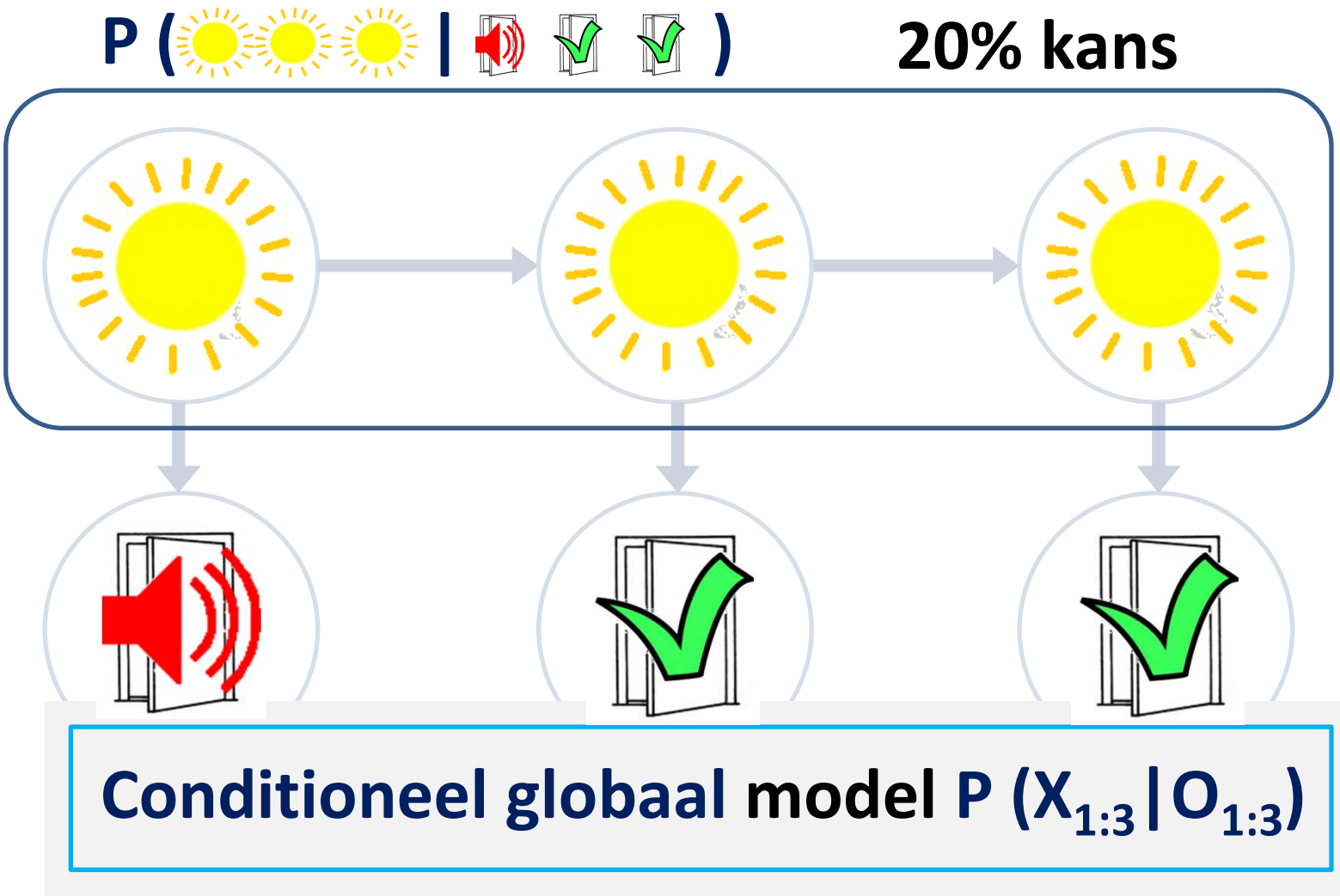


Hidden Markovmodel

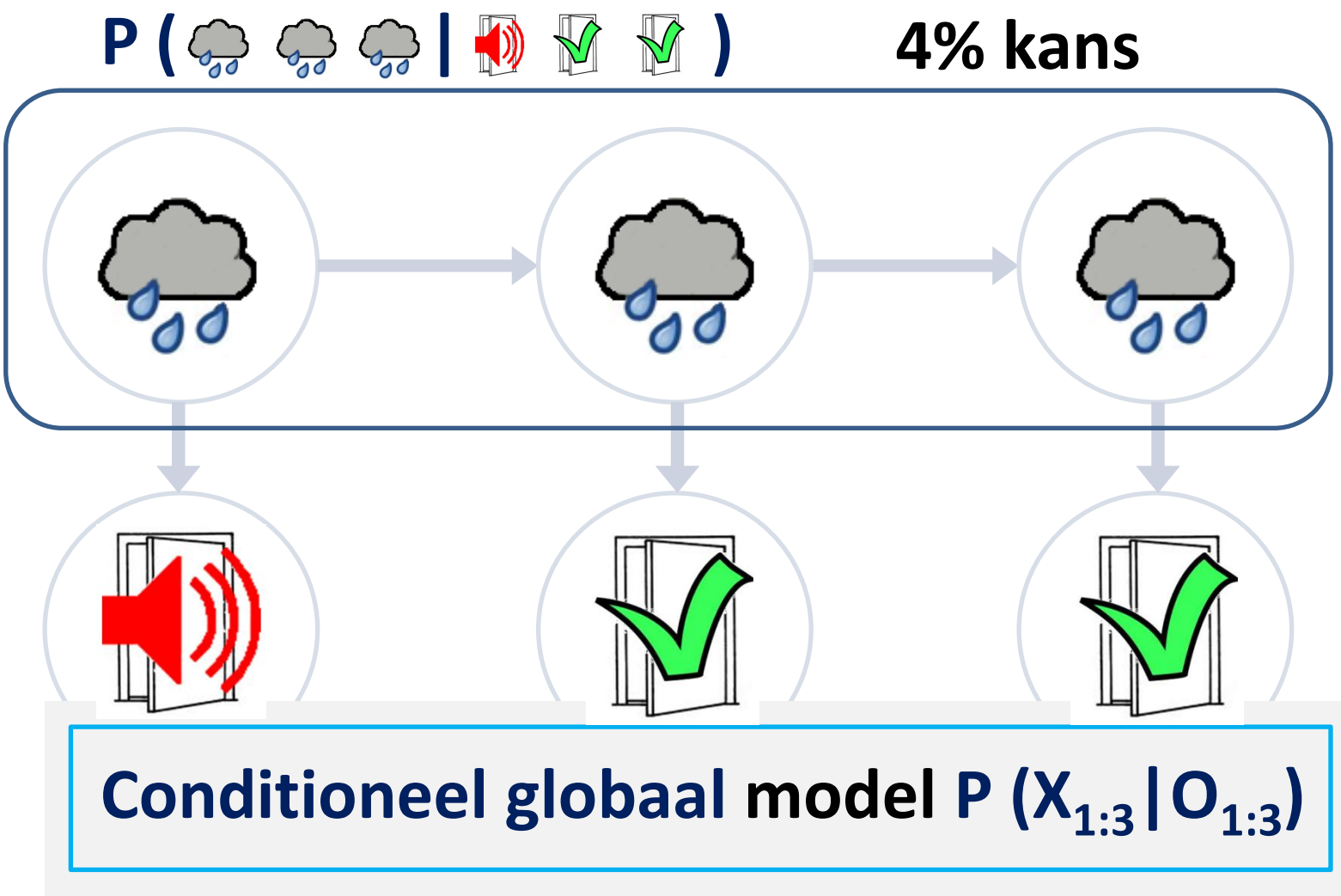


EstiHMM: Een efficiënt algoritme ter bepaling van de maximale sequenties in een imprecies hidden Markovmodel

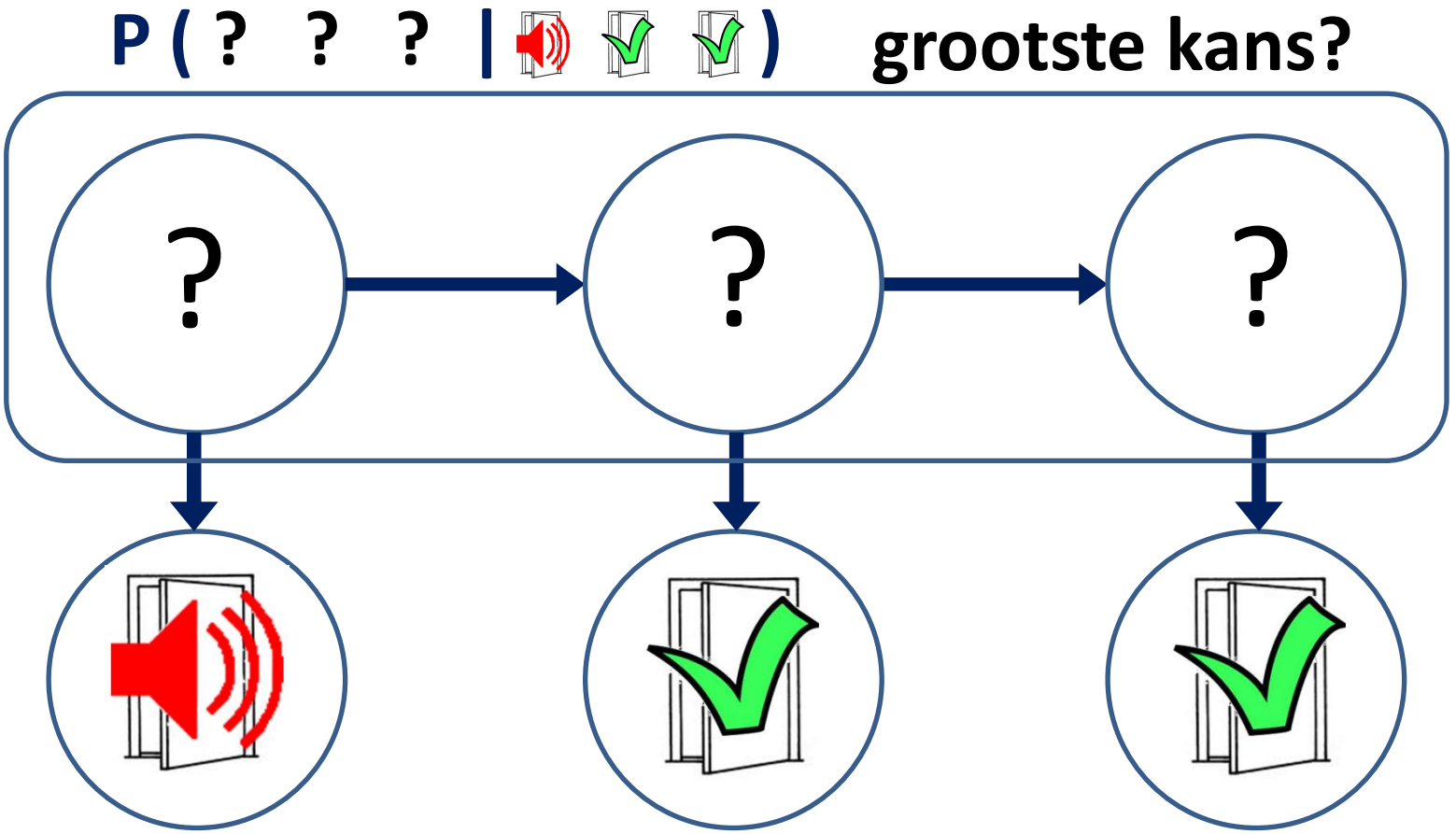
Maximale sequenties



Maximale sequenties

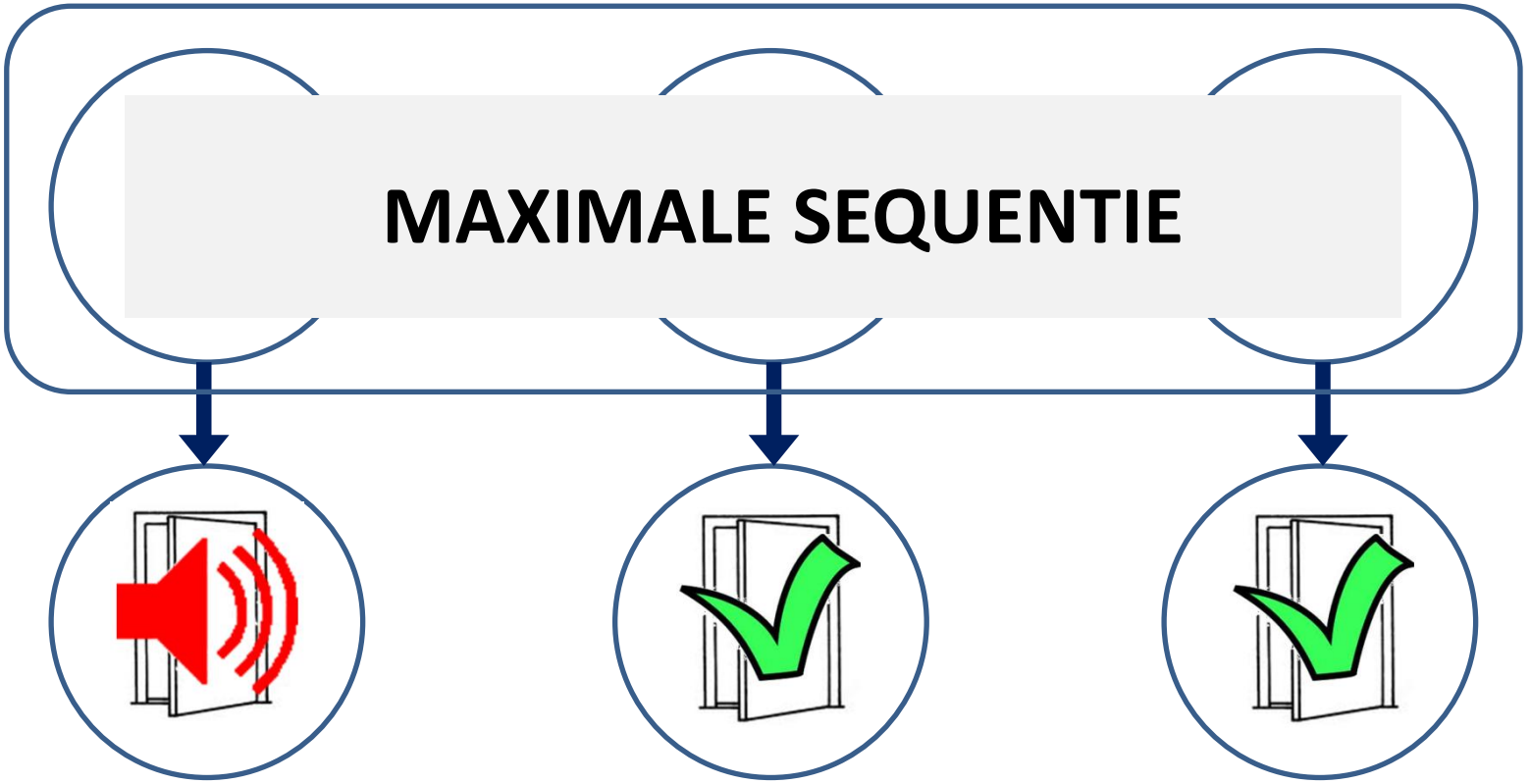


Maximale sequenties



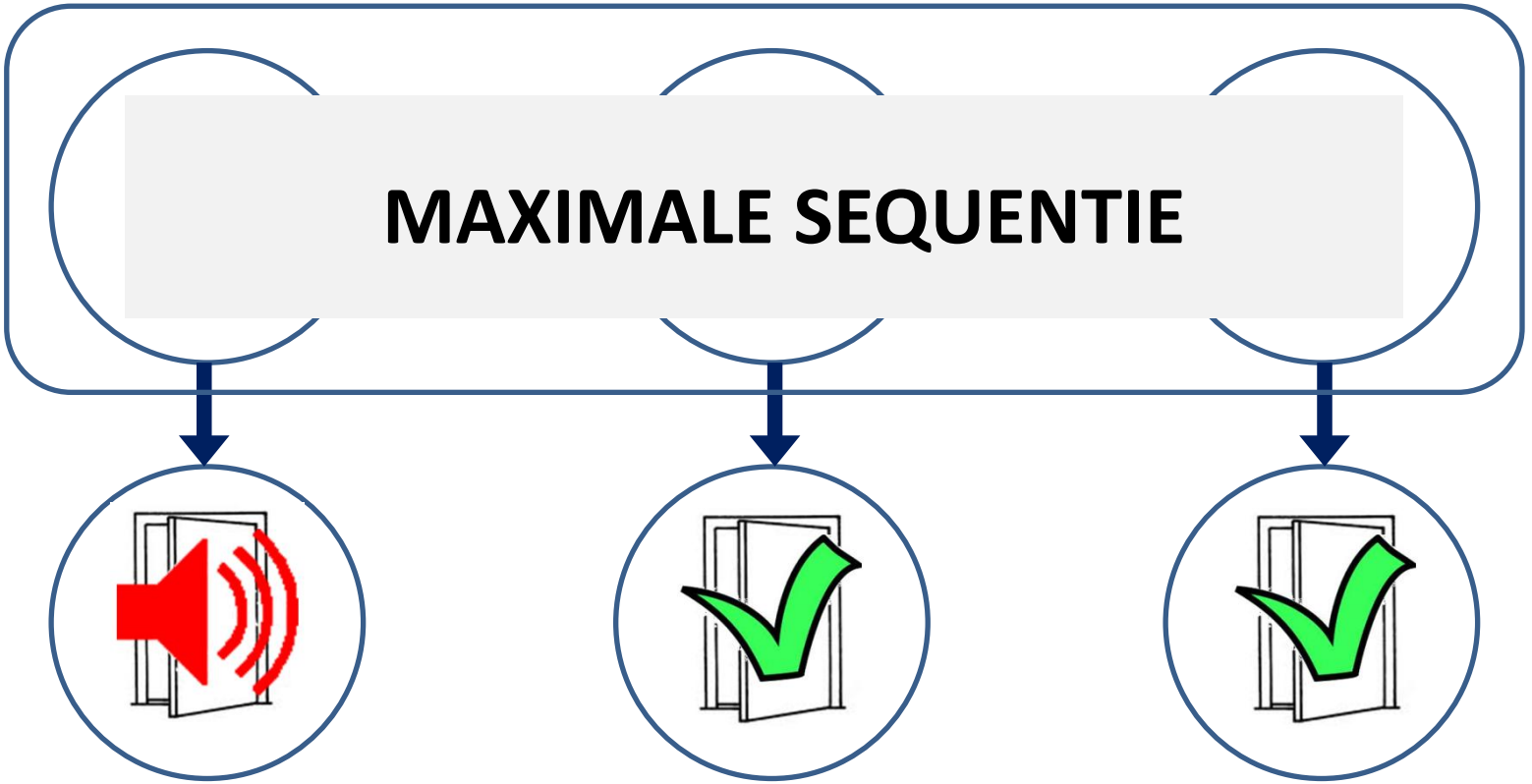
Maximale sequenties

P (? ? ? |   ) grootste kans?



Maximale sequenties

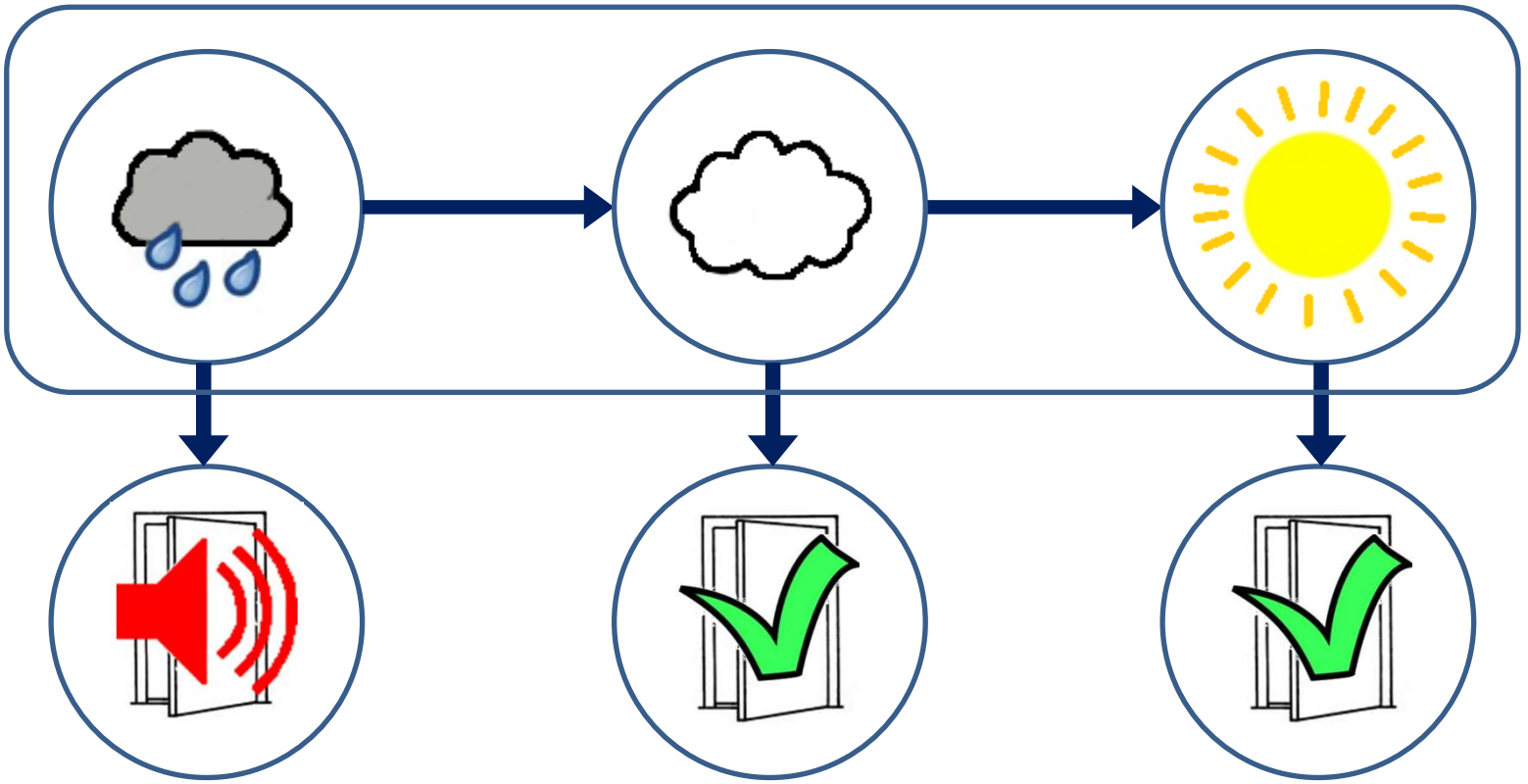
P (? ? ? |   ) grootste kans?



Viterbi-algoritme

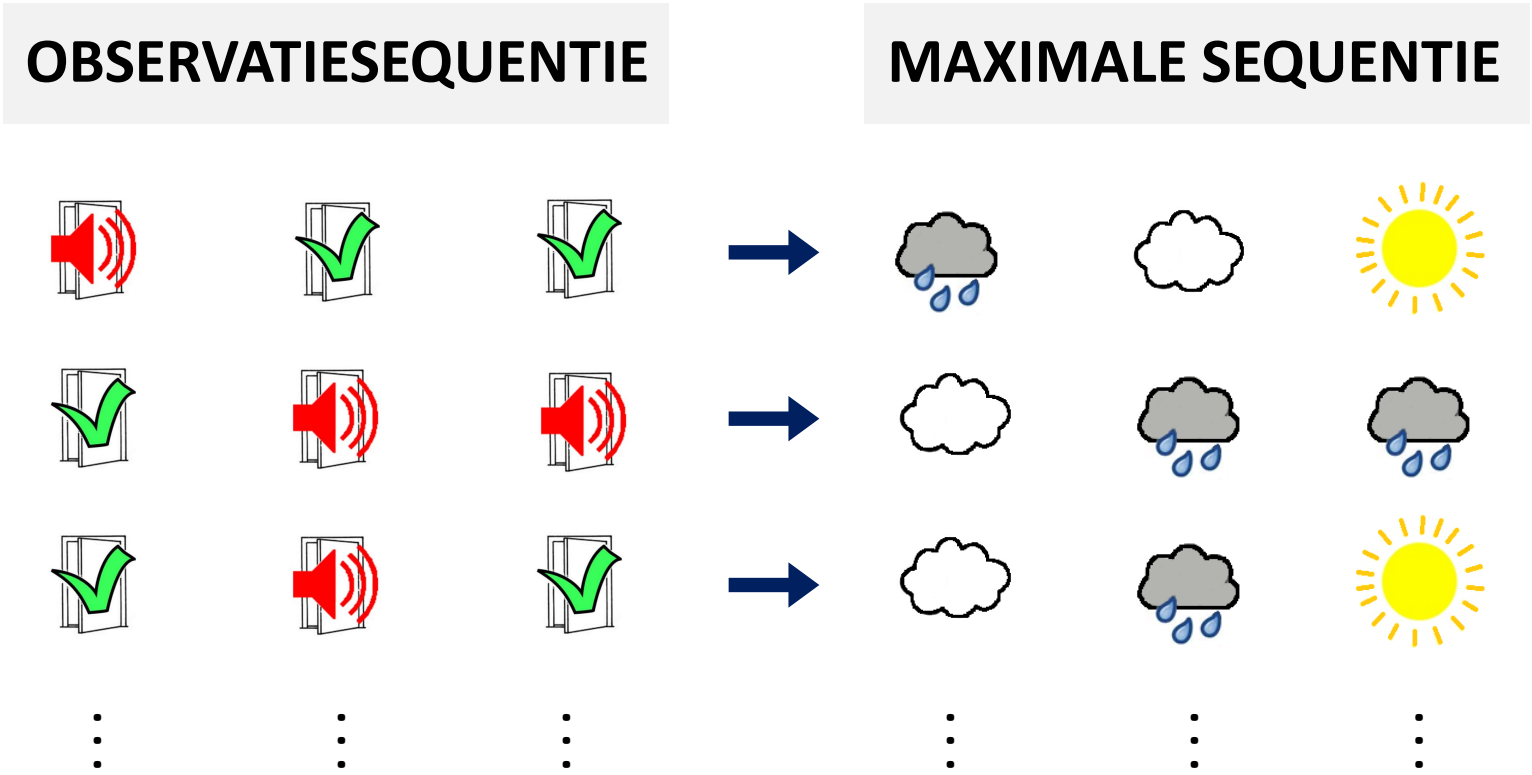
Maximale sequenties

P (? ? ? |   ) grootste kans?



Viterbi-algoritme

Maximale sequenties



Viterbi-algoritme

EstiHMM: Een efficiënt algoritme
ter bepaling van de **maximale sequenties**
in een **imprecies hidden Markovmodel**

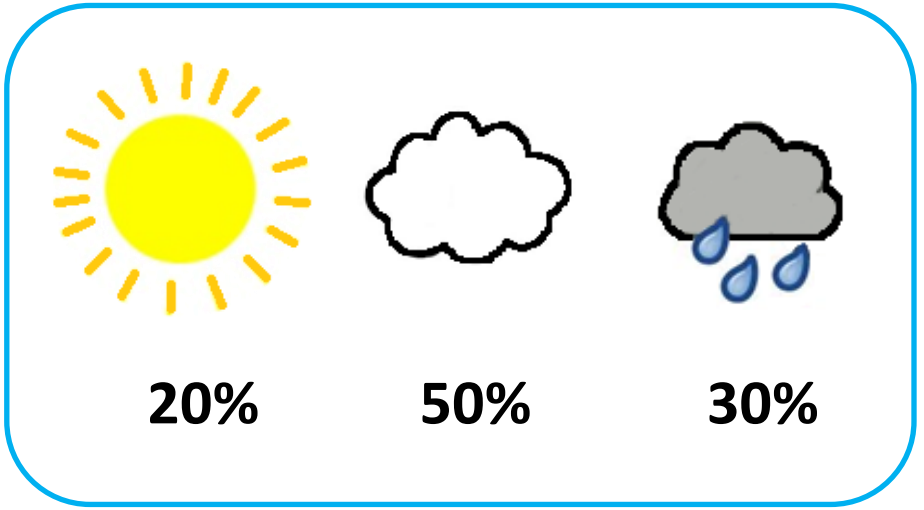
Imprecies hidden Markovmodel

Vrijdag

X_1

$Q_1 (X_1)$

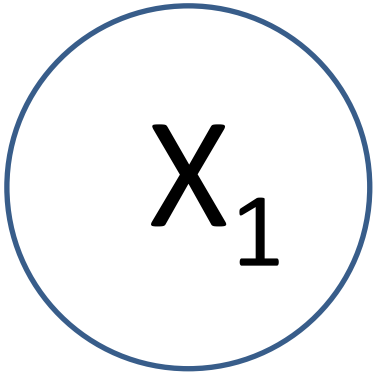
Ben ik daar wel zeker van?



Precies lokaal model $Q_1(X_1)$

Imprecies hidden Markovmodel

Vrijdag

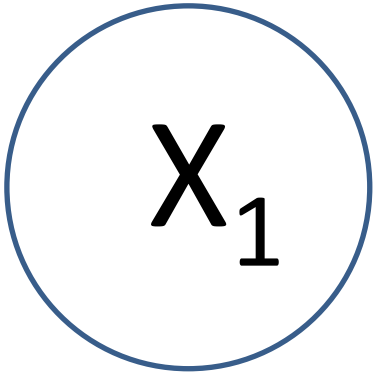


Ben ik daar wel zeker van?

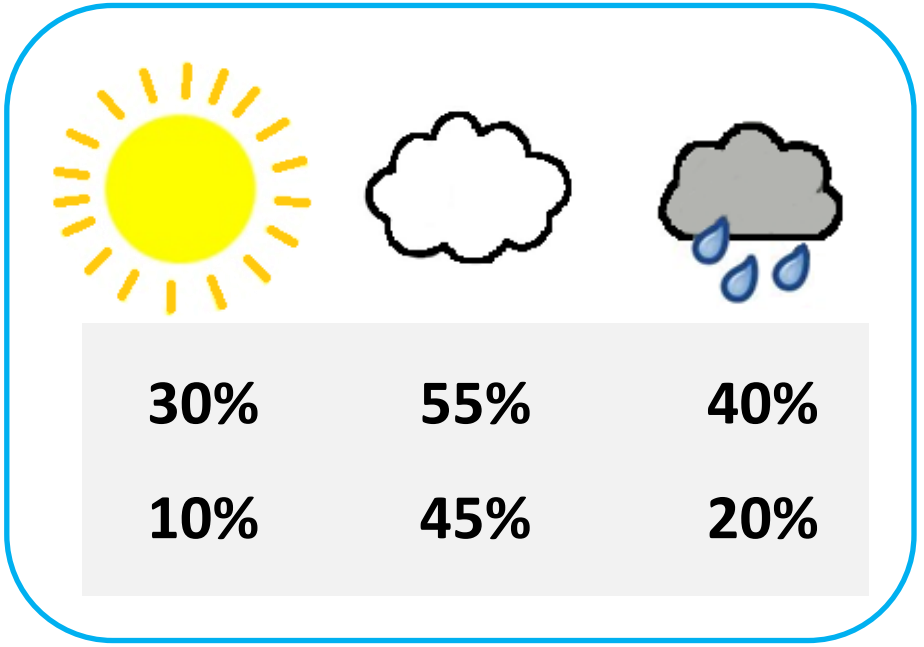
30%	55%	40%
20%	50%	30%
10%	45%	20%

Imprecies hidden Markovmodel

Vrijdag

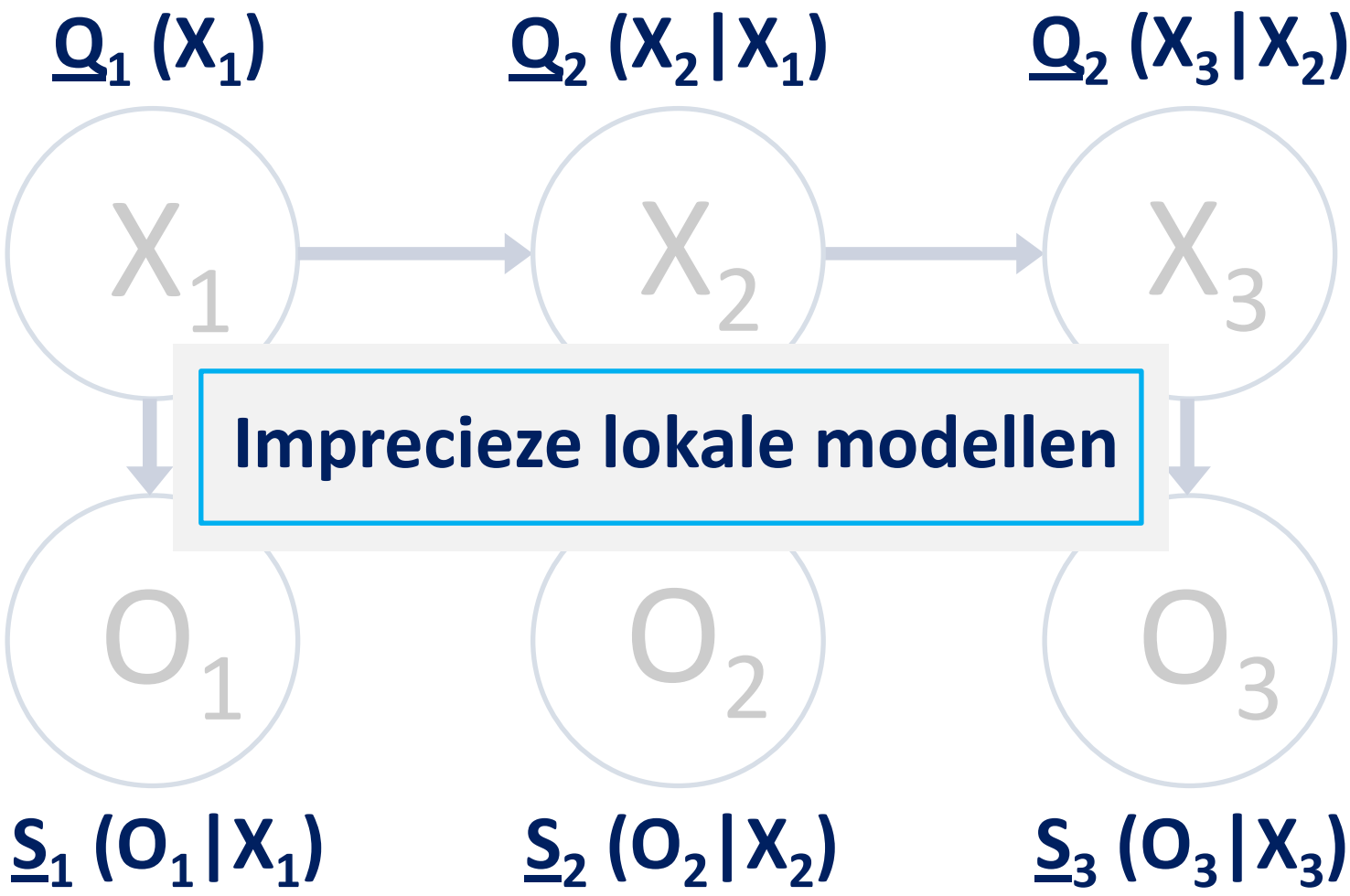


$\underline{Q_1}\ (X_1)$

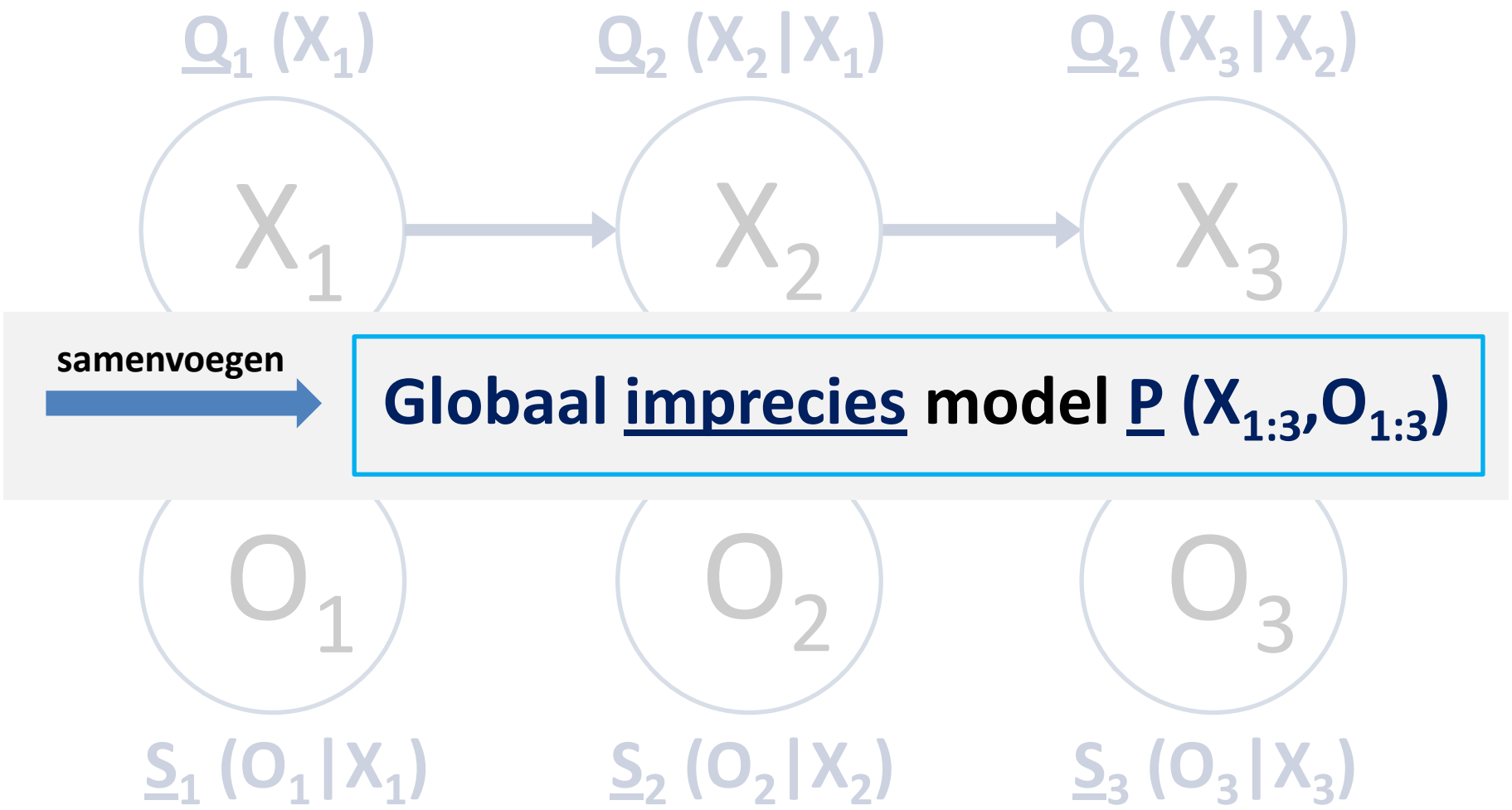


<u>Imprecies</u> lokaal model $\underline{Q_1}(X_1)$

Imprecies hidden Markovmodel



Imprecies hidden Markovmodel



Imprecies hidden Markovmodel

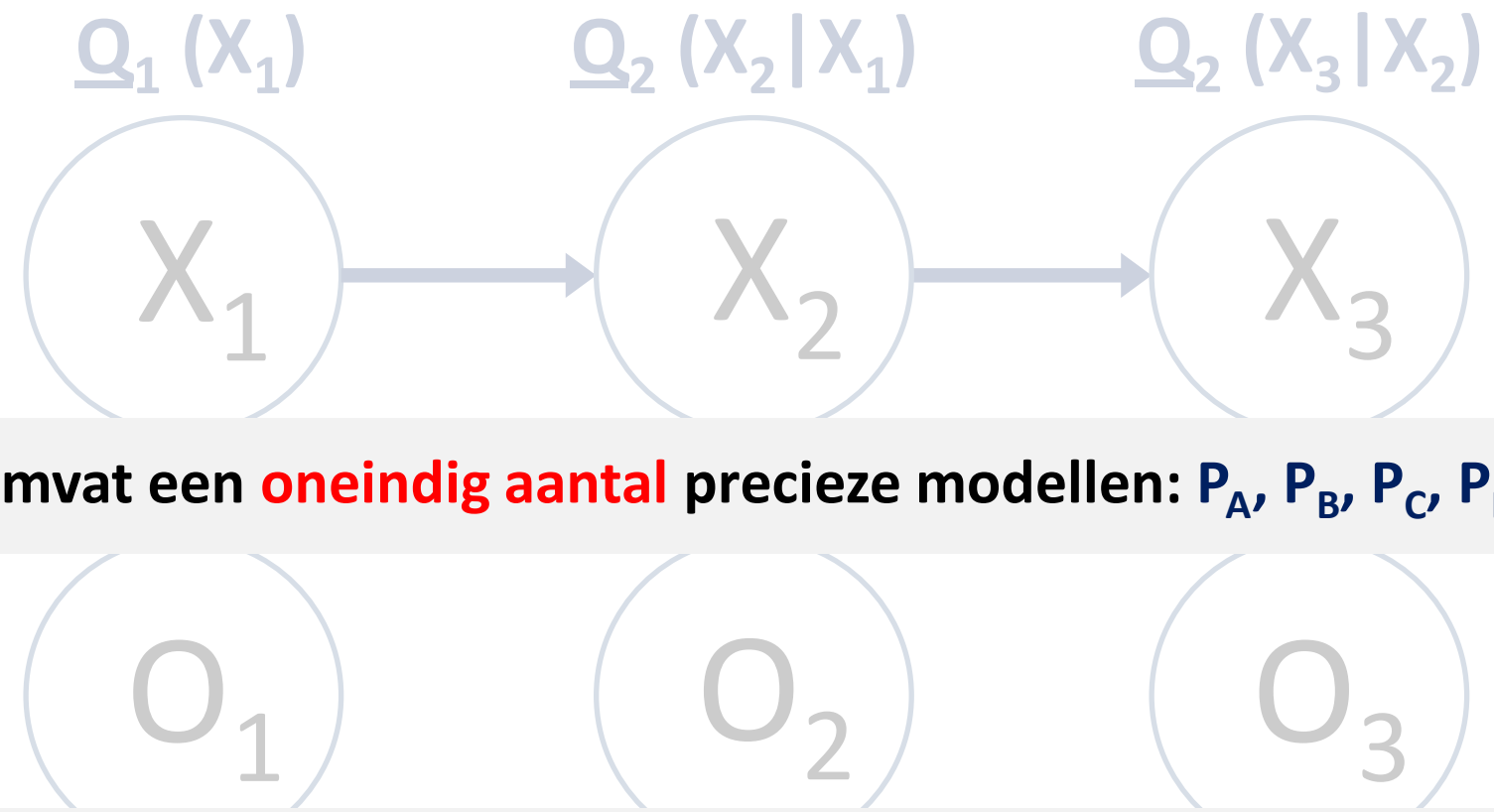
samenvoegen

Globaal <u>imprecies</u> model $\underline{P}(X_{1:3}, O_{1:3})$

Conditioneren op de observaties

Conditioneel globaal <u>imprecies</u> model $\underline{P}(X_{1:3}|O_{1:3})$

Imprecies hidden Markovmodel

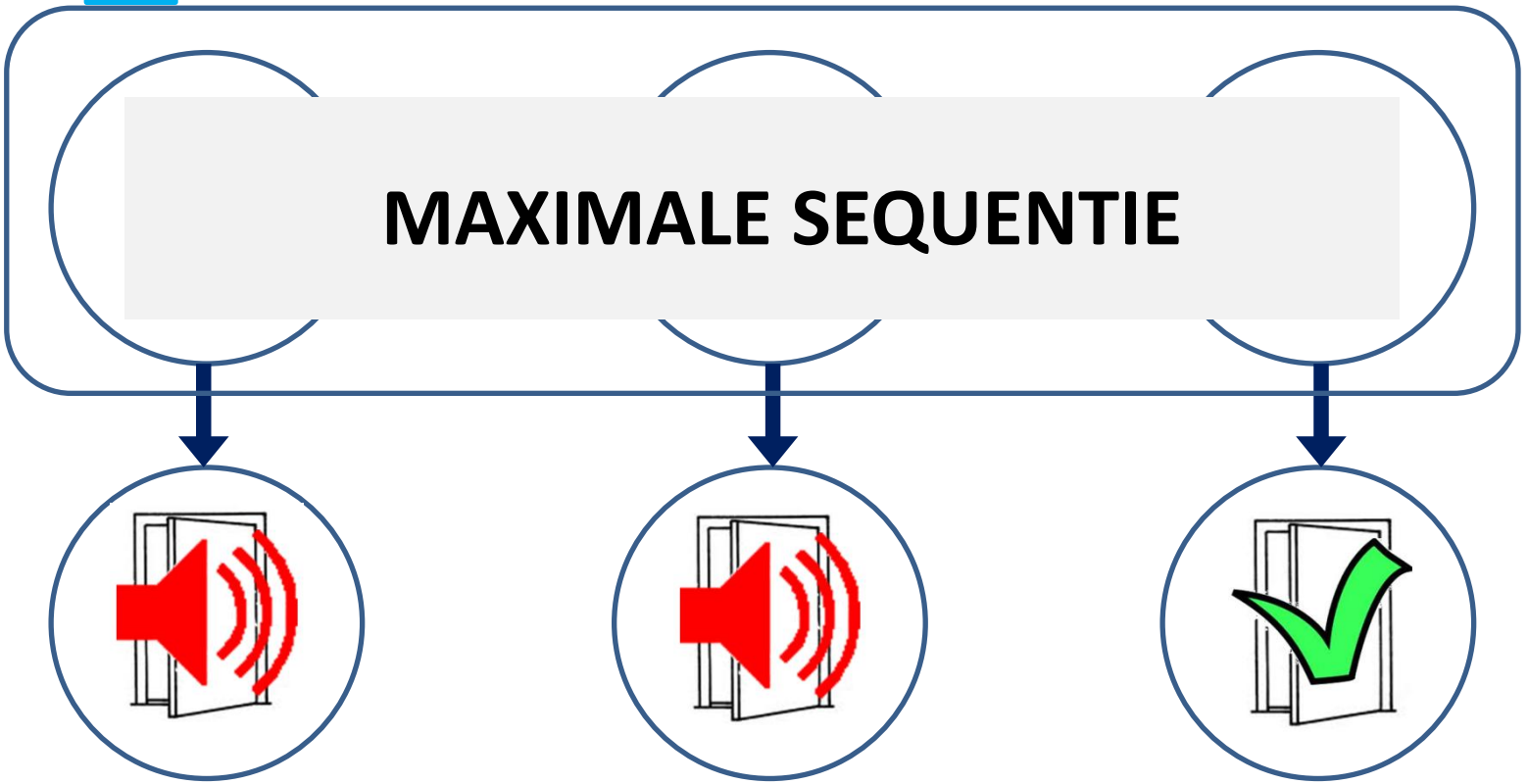


$\underline{P}$ Omvat een oneindig aantal precieze modellen: P_A, P_B, P_C, P_D, ...

Conditioneel globaal $\underline{\text{imprecies}}$ model $\underline{P}(X_{1:3}|O_{1:3})$

Imprecies hidden Markovmodel

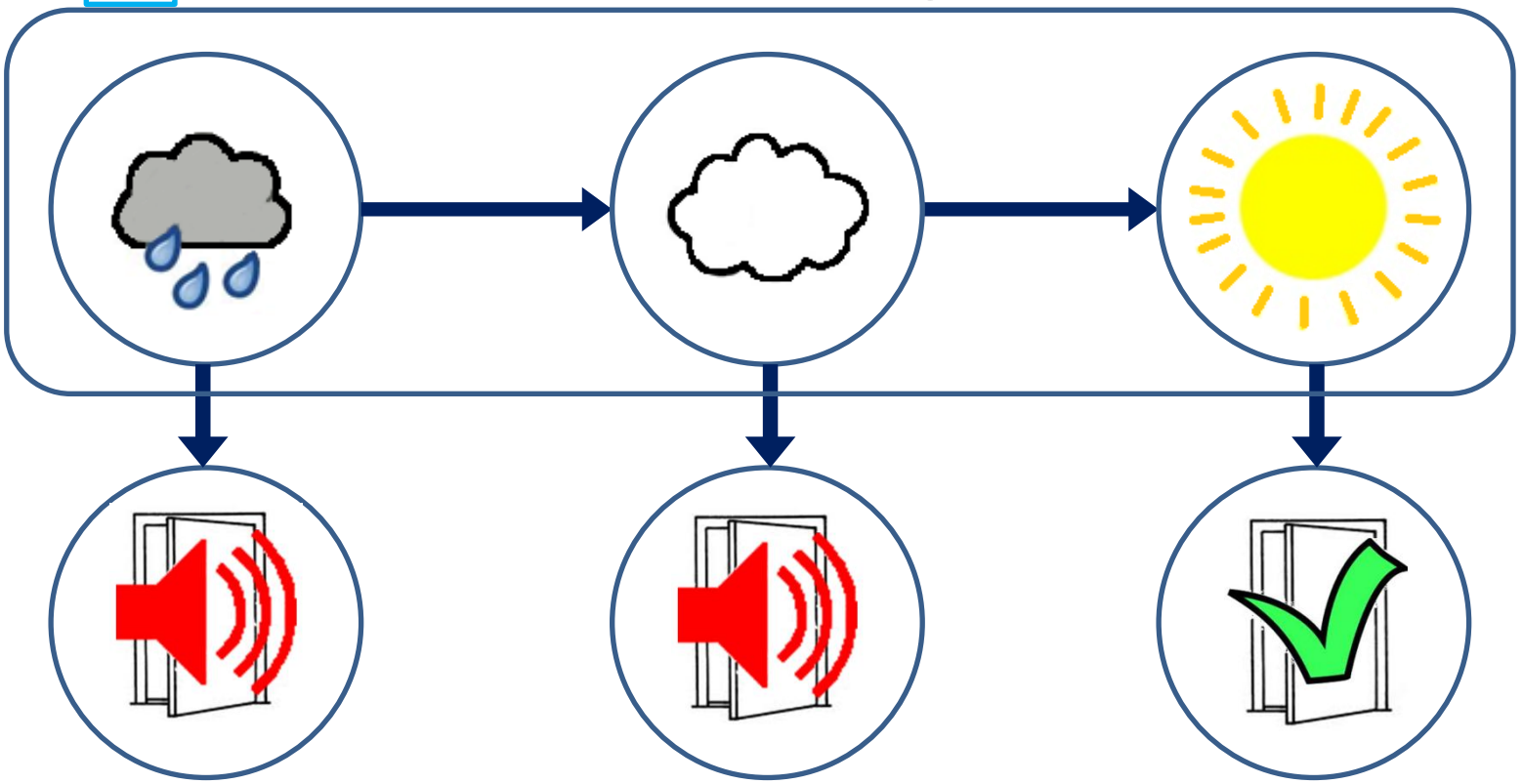
P_A (? ? ? | ) grootste kans?



P Omvat een oneindig aantal precieze modellen: P_A, P_B, P_C, P_D, ...

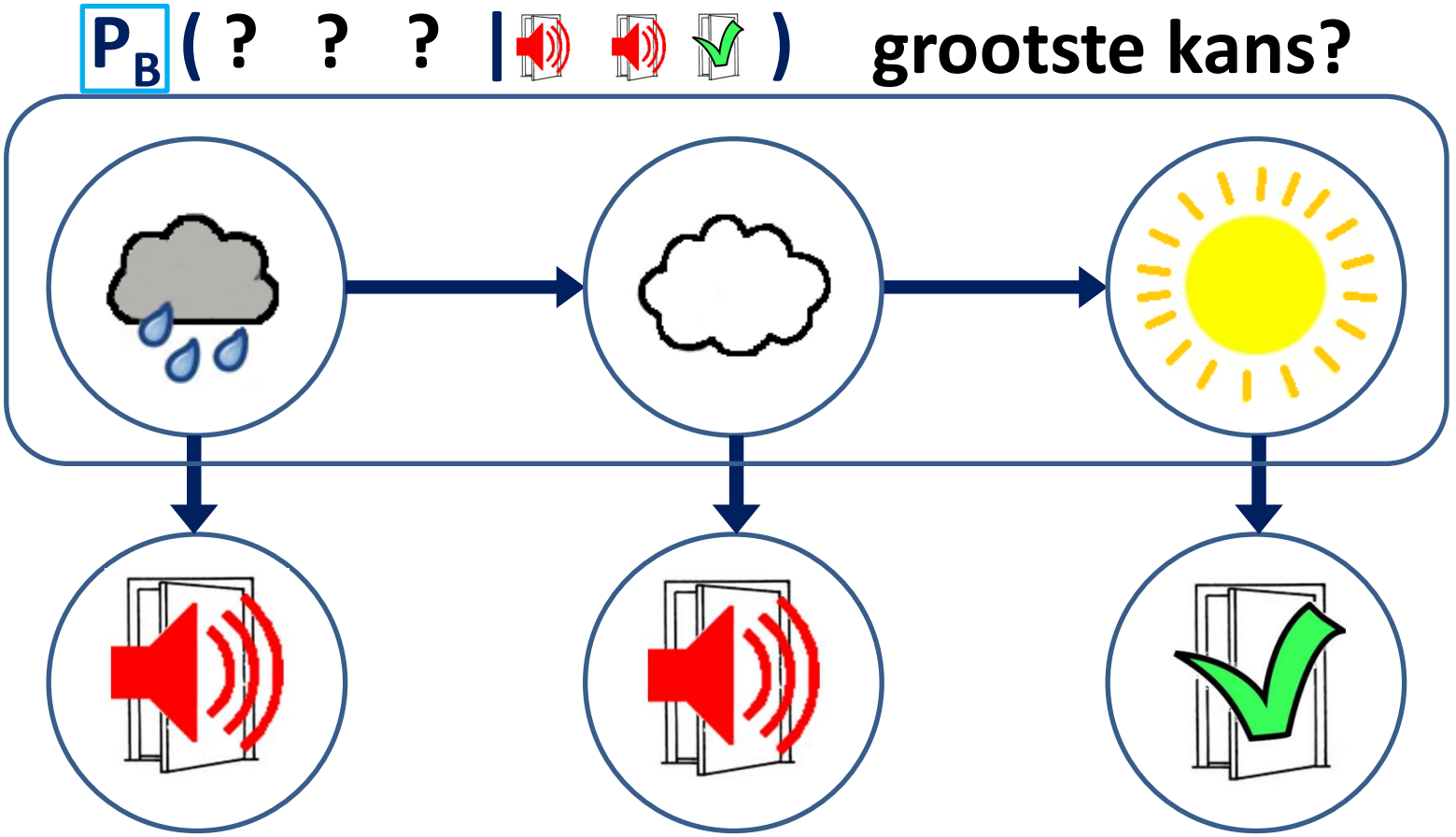
Imprecies hidden Markovmodel

P_A(? ? ? | ) grootste kans?



P Omvat een oneindig aantal precieze modellen: P_A, P_B, P_C, P_D, ...

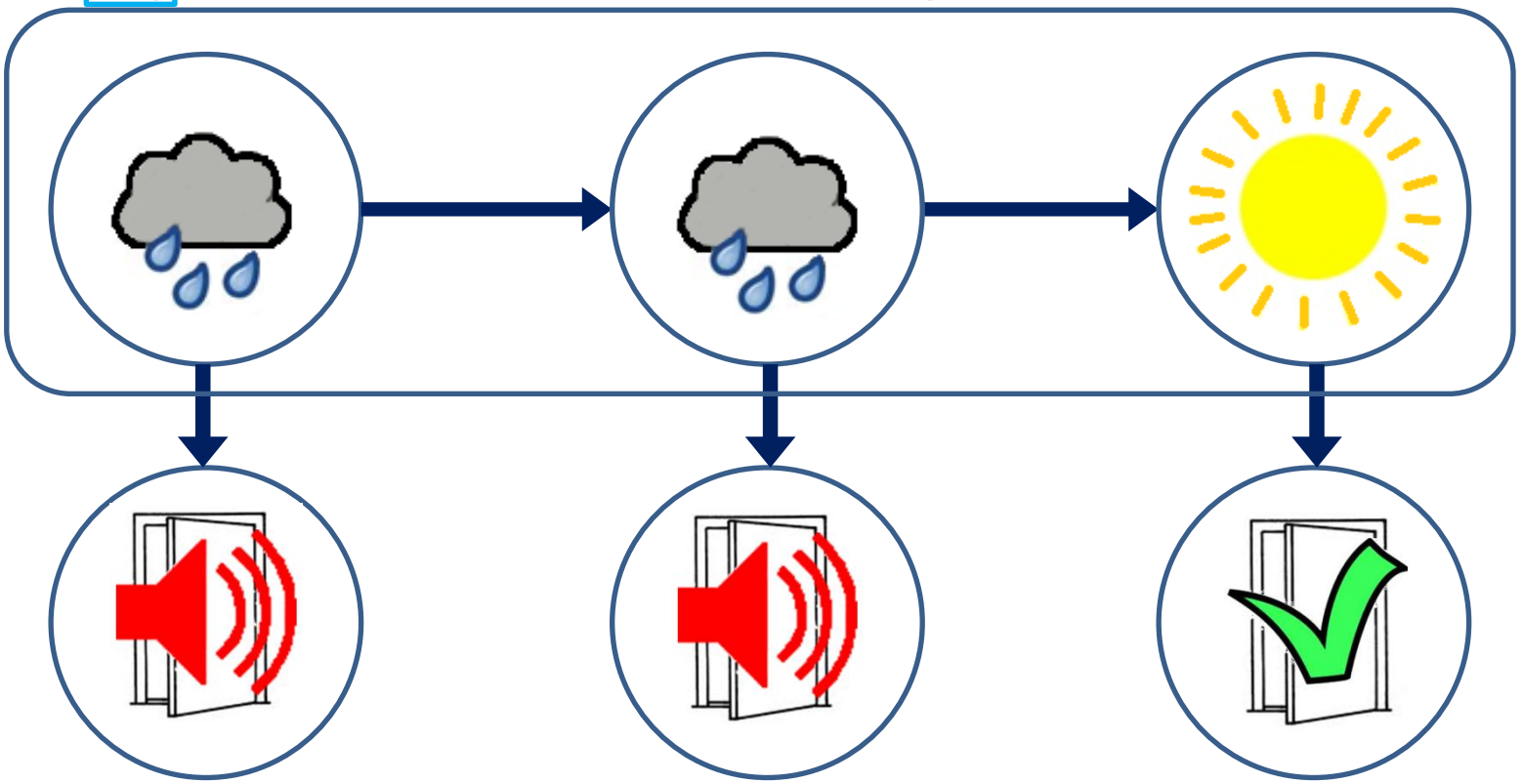
Imprecies hidden Markovmodel



P Omvat een oneindig aantal precieze modellen: P_A, P_B, P_C, P_D, ...

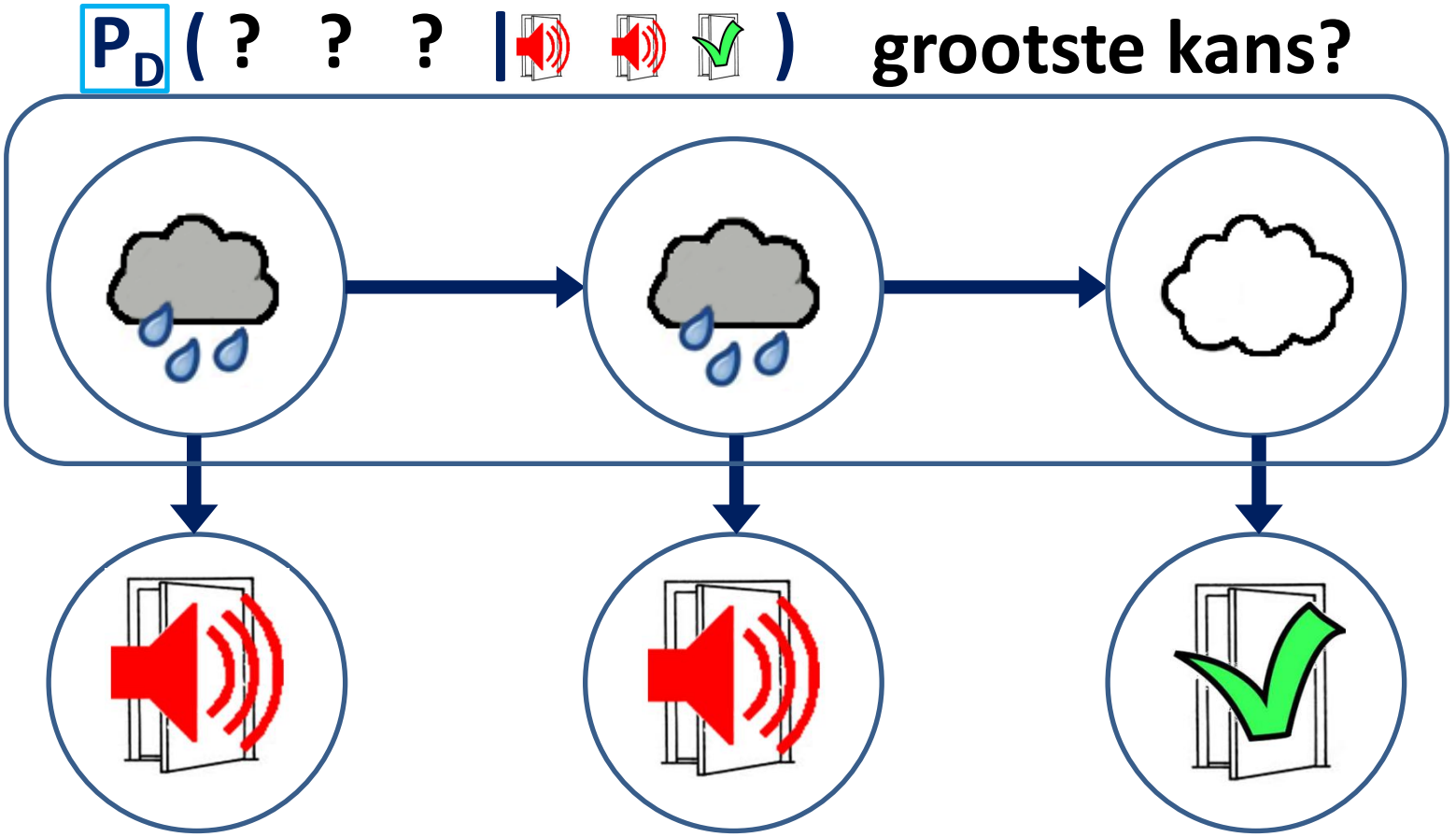
Imprecies hidden Markovmodel

P_C (? ? ? | ) grootste kans?



P Omvat een oneindig aantal precieze modellen: P_A, P_B, P_C, P_D, ...

Imprecies hidden Markovmodel



P Omvat een oneindig aantal precieze modellen: P_A, P_B, P_C, P_D, ...

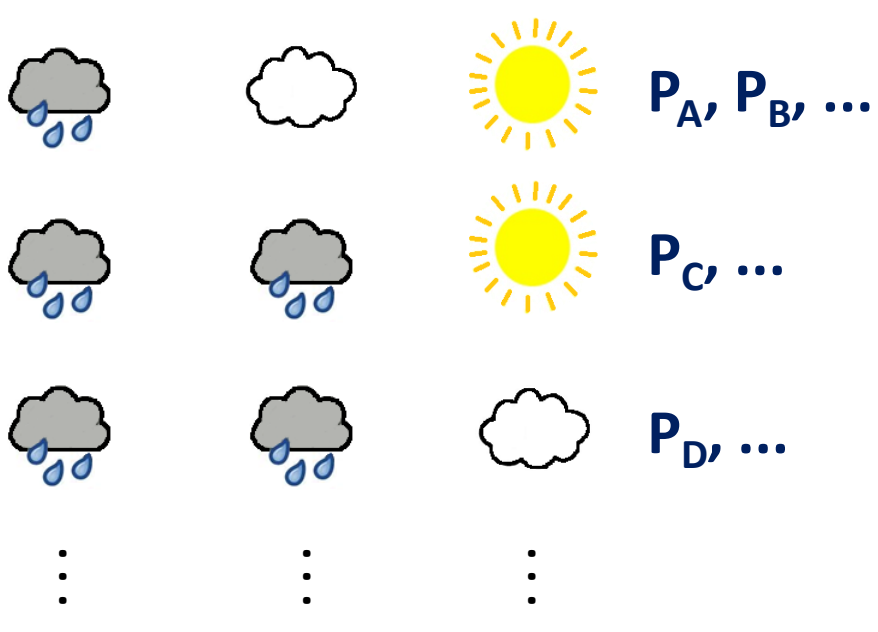
Imprecies hidden Markovmodel

OBSERVATIESEQUENTIE

$\underline{P}\,(\,X_{1:3} \mid$  $)$

≠ maximale sequenties

MAXIMALE SEQUENTIES



Imprecies hidden Markovmodel

OBSERVATIESEQUENTIE



MAXIMALE SEQUENTIES

$P(X_{1:3} |$    $)$

1 maximale sequentie

Imprecies hidden Markovmodel

OBSERVATIESEQUENTIE



MAXIMALE SEQUENTIES

P_A, P_B, P_C, P_D, …

$\underline{P}\,(\,X_{1:3} \mid$    $)$

≠ maximale sequenties mogelijk
in dit geval maar 1 oplossing !

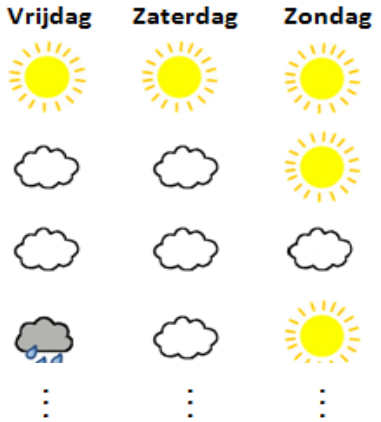
→ **ROBUUST!**

EstiHMM: Een **efficiënt** algoritme
ter **bepaling van de maximale sequenties**
in een **imprecies hidden Markovmodel**

Een efficiënte oplossing

twee grote moeilijkheden

Zeer veel
mogelijkheden



imprecies model

$\underline{P}$

omvat

P_A, P_B, P_C, P_D, …

oneindig aantal
precieze modellen

$$\theta_k(x_k,\hat{x}_k|z_{k-1}) = \min\left\{a \geq 0\colon \underline{Q}_k(\mathbb{I}_{\{x_k\}} - a\mathbb{I}_{\{\hat{x}_k\}}|z_{k-1}) \leq 0\right\}$$

$$\frac{\alpha_n(\hat{x}_n)}{\beta_n^{\max}(x_n)} \geq \theta_n(x_n,\hat{x}_n|z_{n-1})$$

$$\begin{aligned}\rho_k(\varepsilon) &= \underline{P}_k(\mathbb{I}_{\{o_{k:n}\}}[\mathbb{I}_{\{x_{k:n}\}} - \mathbb{I}_{\{\hat{x}_{k:n}\}} - \varepsilon]|x_{k-1}) \\ &= \underline{Q}_k(\underline{E}_k(\mathbb{I}_{\{o_{k:n}\}}[\mathbb{I}_{\{x_{k:n}\}} - \mathbb{I}_{\{\hat{x}_{k:n}\}} - \varepsilon]|X_k)|x_{k-1}) \\ &= \underline{Q}_k\Big(\mathbb{I}_{\{x_k\}}\underline{E}_k(\mathbb{I}_{\{o_{k:n}\}}[\mathbb{I}_{\{x_{k+1:n}\}} - \mathbb{I}_{\{\hat{x}_{k+1:n}\}} - \varepsilon]|x_k) \\ &\qquad + \sum_{z_k \neq x_k} \mathbb{I}_{\{z_k\}}\underline{E}_k(-\varepsilon\mathbb{I}_{\{o_{k:n}\}}|z_k)\Big|\, x_{k-1}\Big)\end{aligned}$$

$$\begin{aligned}\hat{x}_{1:n} \in \operatorname{opt}(\mathscr{X}_{1:n}|o_{1:n}) &\Leftrightarrow (\forall x_{1:n} \in \mathscr{X}_{1:n})\ x_{1:n} \not\prec \hat{x}_{1:n} \\ &\Leftrightarrow (\forall x_{1:n} \in \mathscr{X}_{1:n})\ \underline{P}_1(\mathbb{I}_{\{x_{1:n}\}} - \mathbb{I}_{\{\hat{x}_{1:n}\}}|o_{1:n}) \leq 0 \\ &\Leftrightarrow (\forall x_{1:n} \in \mathscr{X}_{1:n})\ \underline{P}_1(\mathbb{I}_{\{o_{1:n}\}}[\mathbb{I}_{\{x_{1:n}\}} - \mathbb{I}_{\{\hat{x}_{1:n}\}}]) \leq 0\end{aligned}$$

$$\alpha_k^{\text{opt}}(z_{k:s}|z_{k-1}) = \frac{\alpha_k^{\text{opt}}(z_{k:s-1}|z_{k-1})}{\overline{S}_{s-1}(\{o_{s-1}\}|z_{s-1})\overline{Q}_s(\{z_s\}|z_{s-1})}$$

EstiHMM: Een efficiënt algoritme ter bepaling van de maximale sequenties in een imprecies hidden Markovmodel

$$\operatorname{opt}(\mathscr{X}_n|z_{n-1},o_n) = \left\{\hat{x}_n \in \mathscr{X}_n\colon \alpha_n(\hat{x}_n) \geq \alpha_n^{\text{opt}}(\hat{x}_n|z_{n-1})\right\}$$

$$\operatorname{opt}(\mathscr{X}_{k:n}|z_{k-1},o_{k:n}) = \left\{\hat{x}_{k:n} \in \operatorname{mog}(\mathscr{X}_{k:n}|z_{k-1},o_{k:n})\colon \alpha_k(\hat{x}_{k:n}) \geq \alpha_k^{\text{opt}}(\hat{x}_k|z_{k-1})\right\}$$

$$\operatorname{opt}(\mathscr{X}_{k:n}|z_{k-1},o_{k:n}) \subseteq \operatorname{mog}(\mathscr{X}_{k:n}|z_{k-1},o_{k:n})$$

$$\operatorname{mog}(\mathscr{X}_{k:n} = \left(\bigcup_{z_k \in \text{Pos}_k(z_{k-1})} z_k \oplus \operatorname{opt}(\mathscr{X}_{k+1:n}|z_k,o_{k+1:n})\right) \cup \left(\bigcup_{z_k \notin \text{Pos}_k(z_{k-1})} z_k \oplus \mathscr{X}_{k+1:n}\right)$$

$$\begin{aligned}\alpha_k^{\max}(x_k) &\geq \alpha_k^{\text{opt}}(x_k|z_{k-1}) \\ \alpha_{k+1}^{\max}(x_{k+1}) &\geq \alpha_k^{\text{opt}}(\hat{x}_k \oplus x_{k+1}|z_{k-1}) \\ \alpha_{k+2}^{\max}(x_{k+2}) &\geq \alpha_k^{\text{opt}}(\hat{x}_{k:k+1} \oplus x_{k+2}|z_{k-1}) \\ \alpha_n^{\max}(x_n) &\geq \alpha_k^{\text{opt}}(\hat{x}_{k:n-1} \oplus x_n|z_{k-1})\end{aligned}$$

EstiHMM: Een efficiënt algoritme
ter bepaling van de maximale sequenties
in een imprecies hidden Markovmodel

EstiHMM

Estimation in imprecise hidden Markov models

HMM

iHMM

=

imprecies hidden Markovmodel

bepalen van de maximale sequenties

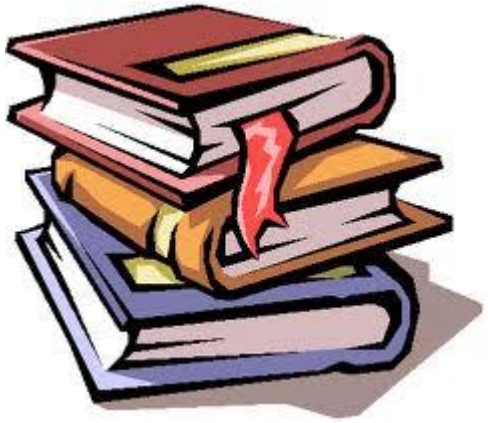
EstiHMM: Een efficiënt algoritme ter bepaling van de maximale sequenties in een imprecies hidden Markovmodel

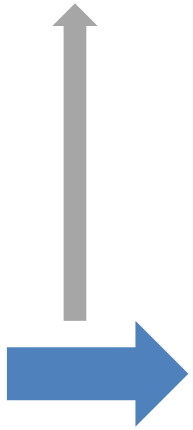
“Wat kunt ge daar nu eigenlijk mee doen?”

Spraakherkenning
Dynamica van aardbevingen
Dynamica van bouwkundige constructies
Optische karakterherkenning
Bio-infomatica
. . .

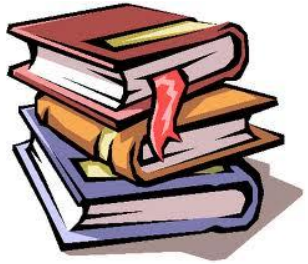
Een toepassing

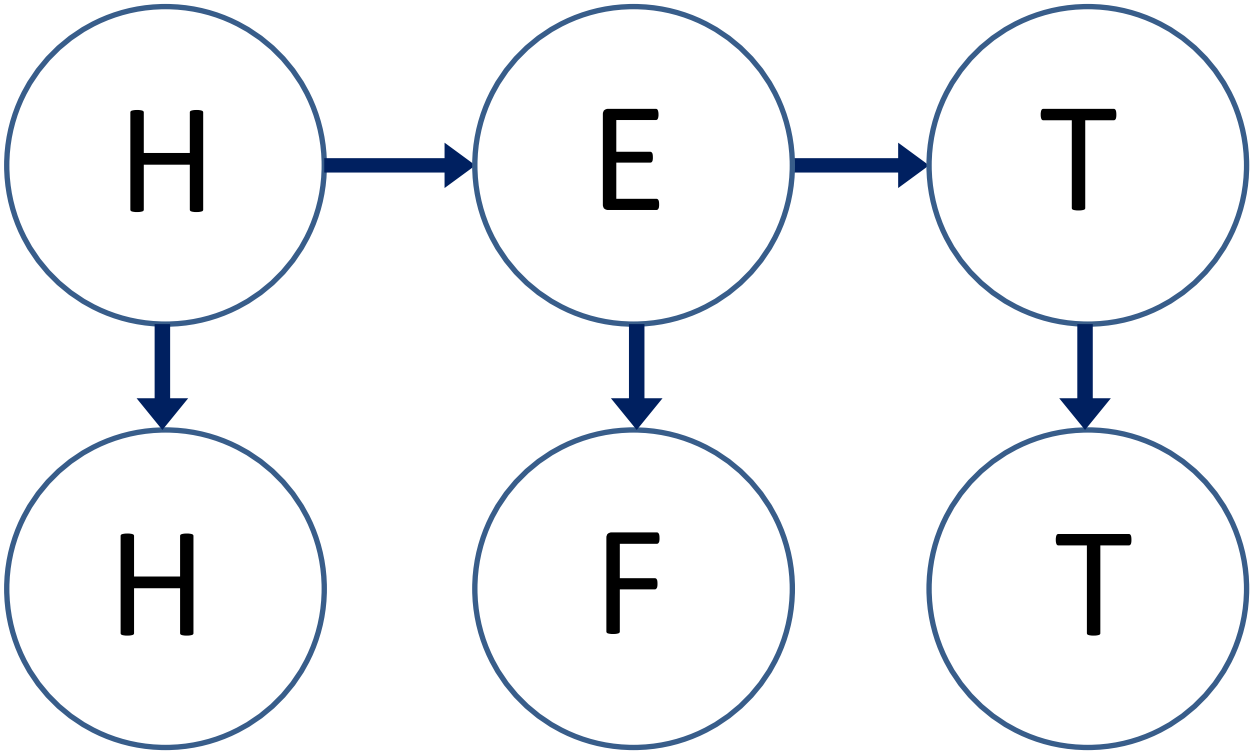
Optische karakterherkenningssoftware



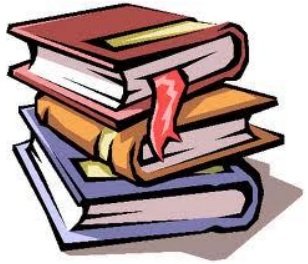


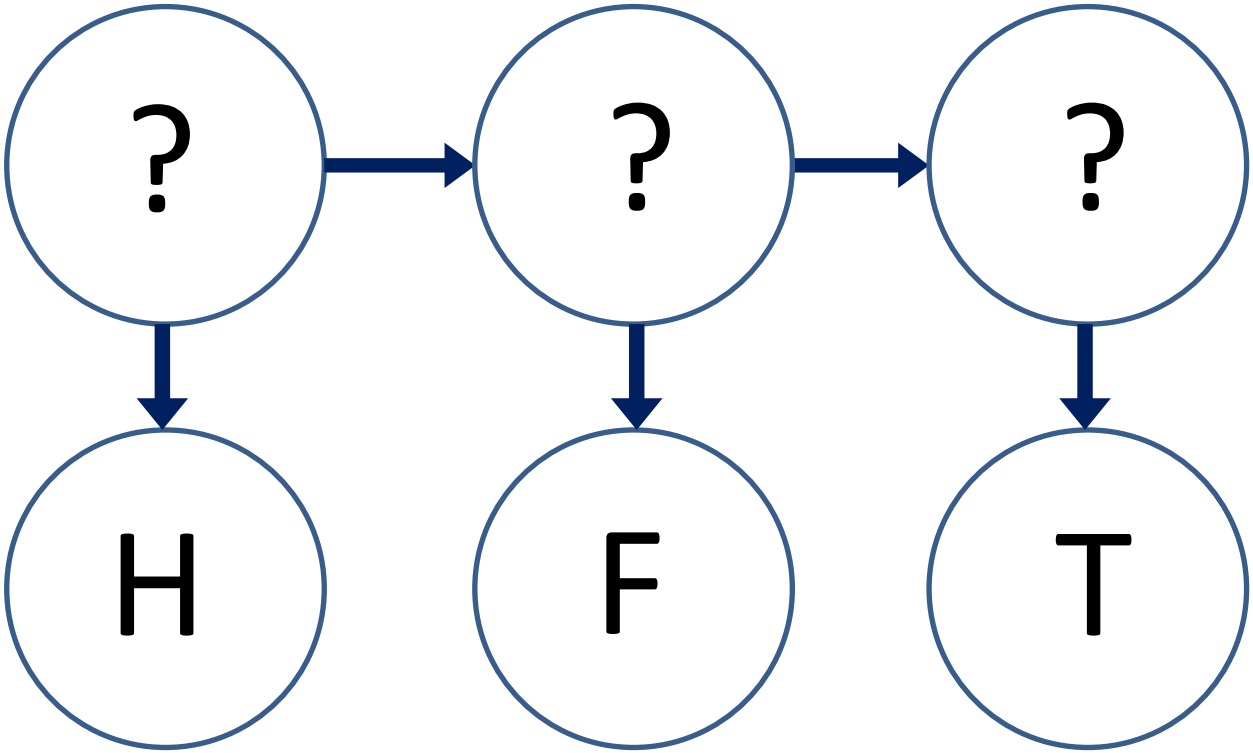
Een toepassing





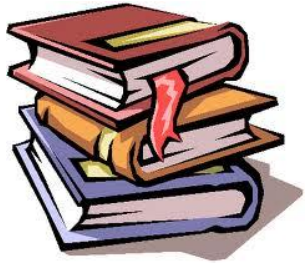
Een toepassing

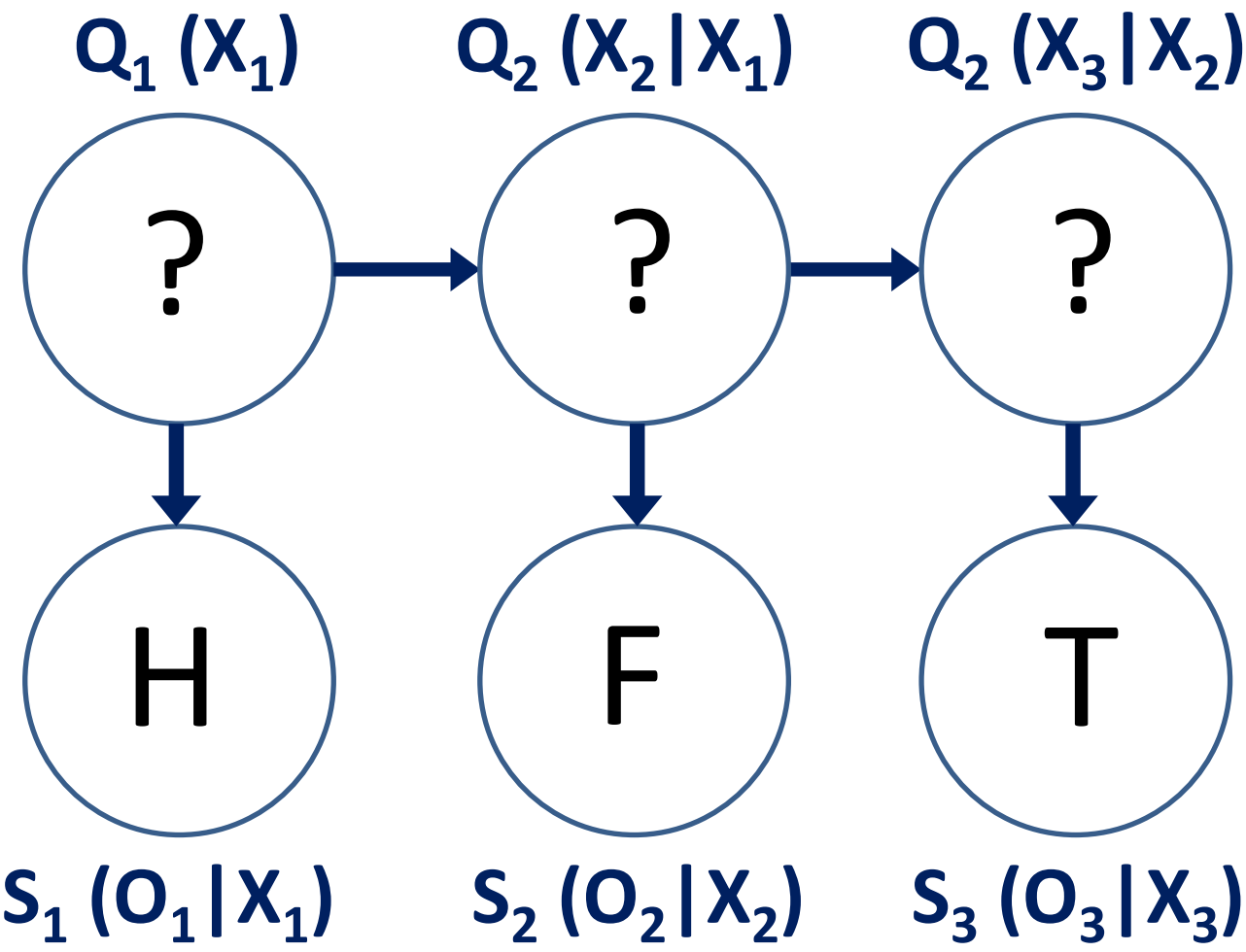




Een toepassing

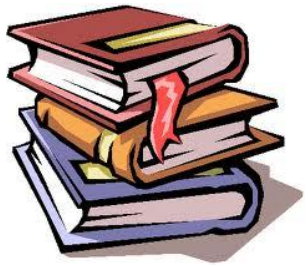
Viterbi-algoritme

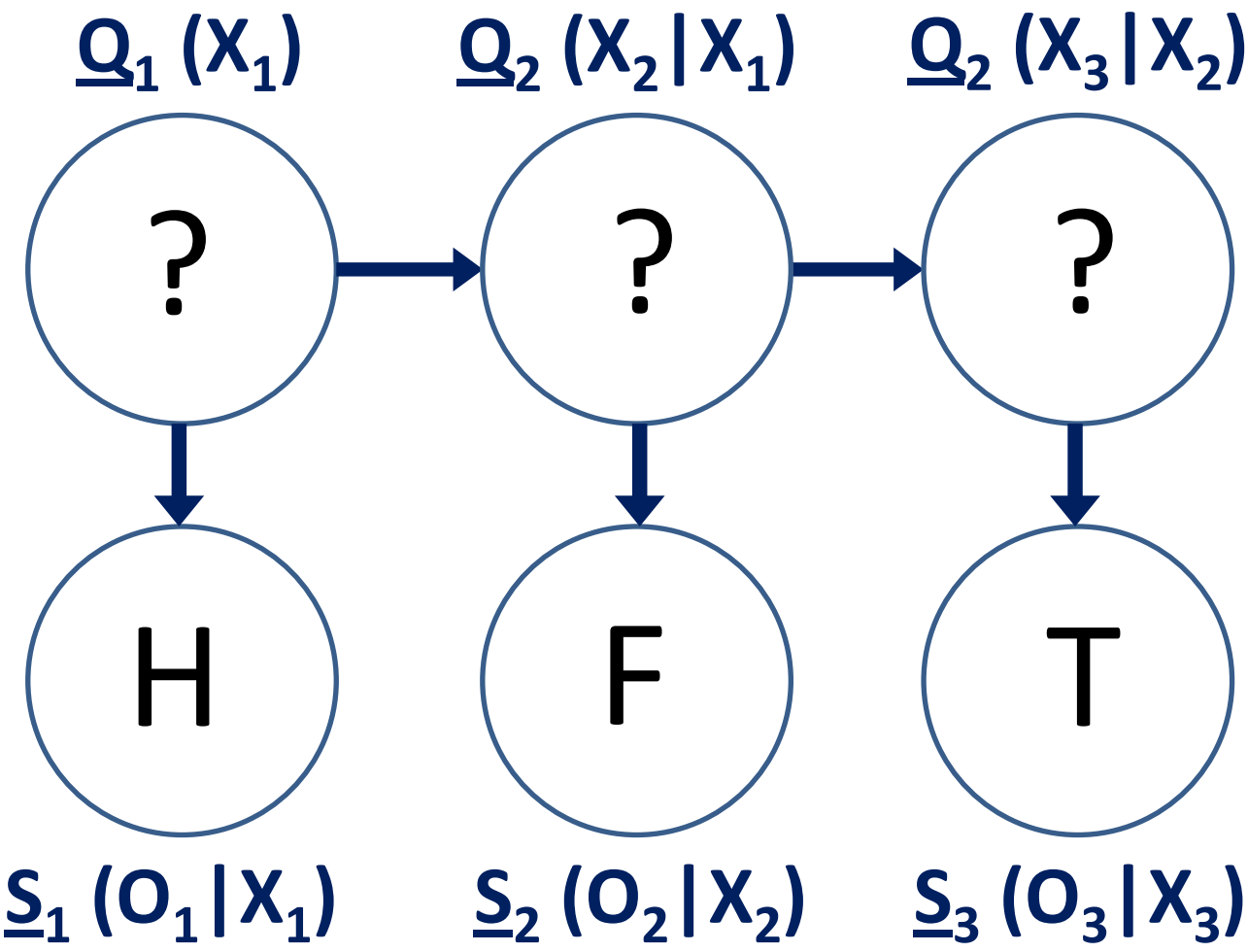




Een toepassing

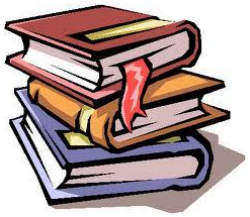
EstiHMM-algoritme





Een toepassing

La Divina Commedia



origineel

CON

foutief ingelezen

digitaal

CCN

Oplossing Viterbi

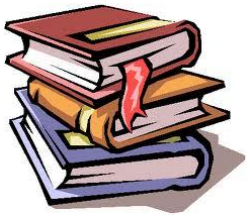
CON

Oplossing(en) EstiHMM-algoritme

CON

Een toepassing

La Divina Commedia



origineel

EH

correct ingelezen

digitaal

EH

Oplossing Viterbi

EN

Oplossing(en) EstiHMM-algoritme

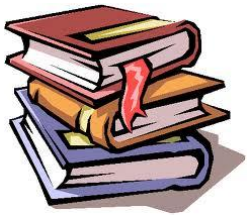
CH

EH

EN

Een toepassing

La Divina Commedia



origineel

IO

foutief ingelezen

digitaal

ZO

Oplossing Viterbi

LO

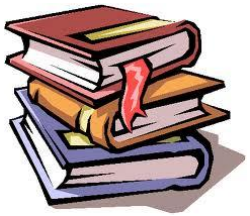
Oplossing(en) EstiHMM-algoritme

LO

IO

Een toepassing

La Divina Commedia



origineel

CHE

foutief ingelezen

digitaal

CNE

Oplossing Viterbi

ONE

Oplossing(en) EstiHMM-algoritme

CBE **CHE**

CNE **CZE**

ONE

Bedankt voor jullie aandacht!