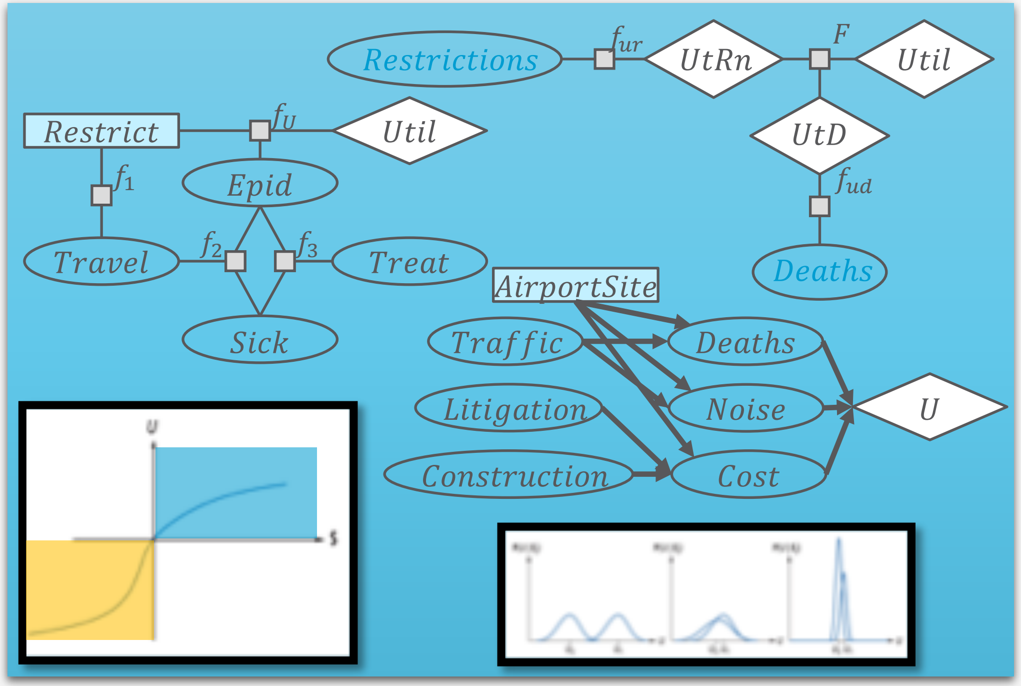
WWU
MÜNSTER

Wir tragen Maske!
We wear masks!

# Entscheidungstheoretische PGMs und Inferenz

## Einführung in die Künstliche Intelligenz



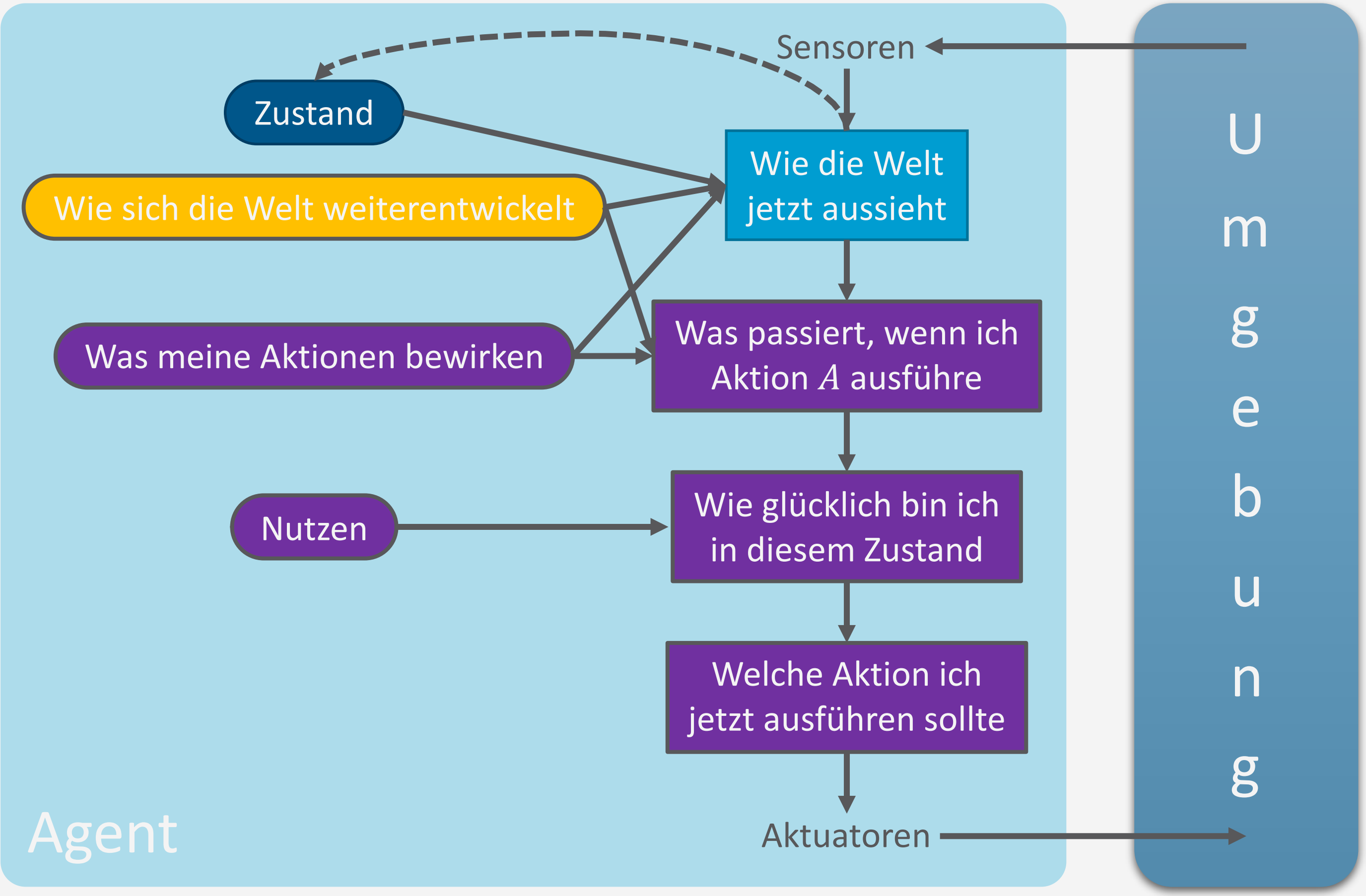



Tanya Braun
Arbeitsgruppe Data Science, Institut für Informatik

# Inhalte

# Einordnung der Vorlesung: *Modell- und nutzenbasierter Agent*

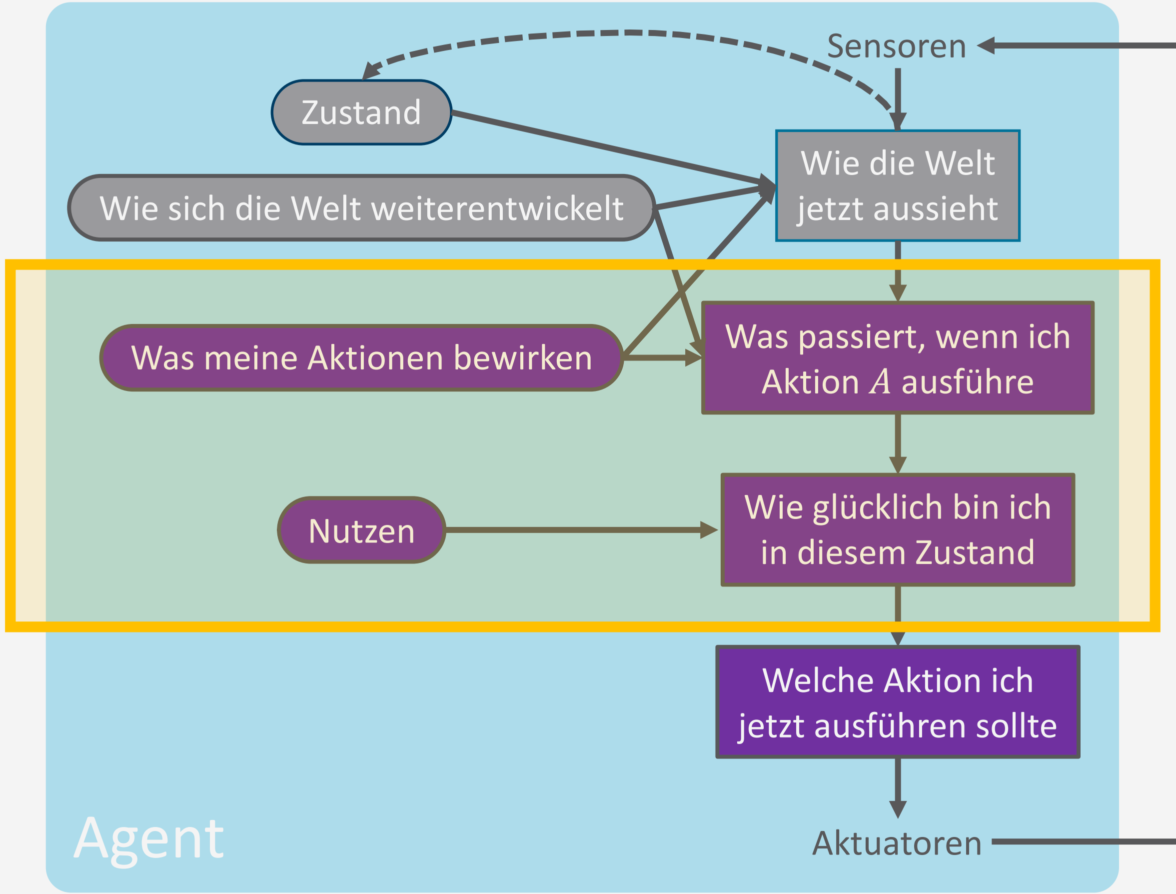
- Nachfolgende Themen der Vorlesung
  2. Episodische PGMs
  3. Exakte Inferenz in episodischen PGMs
  4. Approximative Inferenz in episodischen PGMs
  5. Lernalgorithmen für episodische PGMs
  6. Sequentielle PGMs und Inferenz
  7. Entscheidungstheoretische PGMs und Inferenz



# Erwarteter Nutzen (*Expected Utility*)

- Zufallsvariable $R$ mit $n$ Werten $r_1, \dots, r_n$ als Domäne, einer Verteilung $(p_1, \dots, p_n)$ folgend
  - $R$ repräsentiert den Zustand, den man erreicht, nachdem man Aktion $a$ unter Unsicherheit ausgeführt hat (Ergebniszustand)
- Funktion $U$ von $R$
  - Beispiel: $U$ ist der Nutzen des Zustandes
- Erwarteter Nutzen von $a$ ist

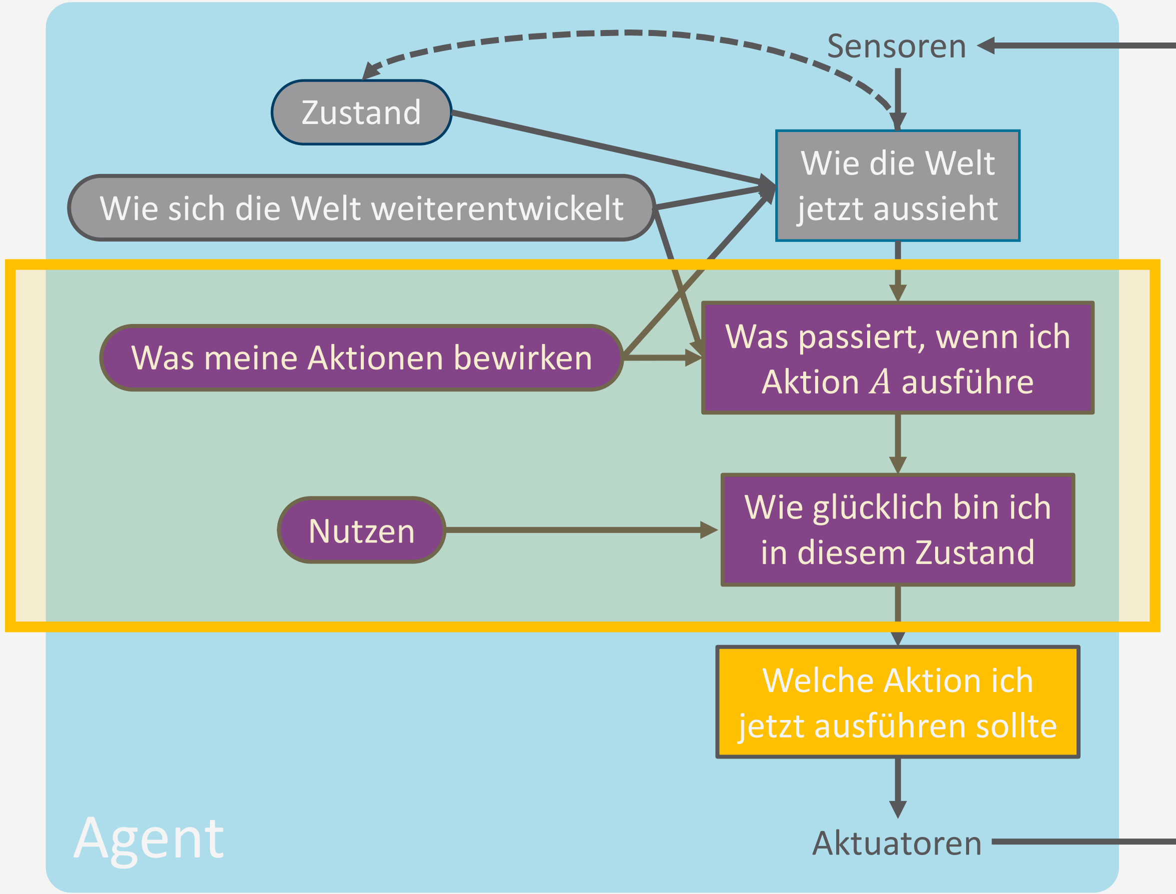
$$EU[a] = \sum_{i=1}^{n} P(r_i|a) \cdot U(r_i)$$



# Maximum-Expected-Utility (MEU) Prinzip

- Ein rationaler Agent sollte die Aktion wählen, welche den erwarteten Nutzen des Agenten maximiert
- Grundlage des Forschungsgebiets der Entscheidungstheorie
- MEU Prinzip bietet ein normatives Kriterium für rationale Entscheidungen

KI ist gelöst?!



# …nicht ganz

- Benötigt vollständige Informationen in einem Modell über
  - Aktionen
  - Nutzen
  - Zuständen
  - Teilweise schwierig zu beschaffen:
    - Berechnung des Nutzens jedes Zustandes bedingt häufig Suche oder Planung, weil ein Agent nicht weiß, wie gut ein Zustand ist, bis er weiß, wohin er von diesem Zustand aus gelangen kann
- Selbst, wenn man ein Modell darüber hat, kann das Problem intraktabel sein
- Tatsächlich würde ein wirklich rationaler Agent auch den Nutzen des Abwägens mit einbeziehen – beschränkte Rationalität
- Nichtsdestotrotz Formalisierung erlaubte große Fortschritte in dem Gebiet in den letzten Jahrzehnten, erlaubt komplexe entscheidungstheoretische Probleme zu lösen

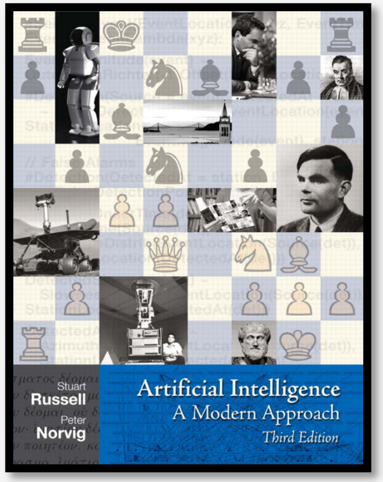
# Rahmen

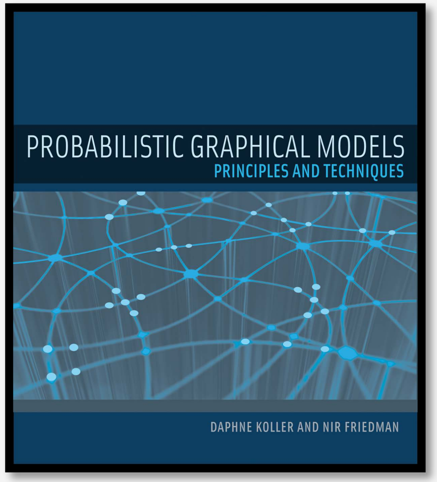
- Agent kann Aktionen in einer Umgebung ausführen
  - Umgebung
    - Anfangs episodisch, später sequentiell
    - Nicht-deterministisch
      - Ausgang von Aktionen nicht eindeutig
      - Mit Wahrscheinlichkeiten assoziiert (→ probabilistisches Modell)
    - Partiell beobachtbar
      - Latente, i.e., nicht beobachtbare, Zufallsvariablen
  - Agent hat Präferenzen über Zustände / Aktionsausgänge
    - In einer Nutzenfunktion kodiert → Nutzentheorie
- „Entscheidungstheorie = Nutzentheorie + Wahrscheinlichkeitstheorie“
  - Modelliere Präferenzen mit Nutzenfunktion, die Welt mit probabilistischem Modell
  - Finde die Aktion, welche zum maximalen erwarteten Nutzen führt
    - Auch Entscheidungsfindung (*decision making*) genannt

# Literaturhinweise

Inhalte dieses Themenblocks werden in den folgenden Kapiteln der Vorlesungsbücher behandelt

- AIMA(de)
  - Kap. 16: Einfache Entscheidungen
  - Kap. 17.1: Sequentielle Entscheidungsprobleme
    - Besonders. 17.1.1
- PGM
  - Kap. 22: Nutzen und Entscheidungen
  - Kap. 23: Strukturierte Entscheidungsprobleme
    - Besonders: 23.2, 23.4, 23.7



# Überblick: 7. Entscheidungstheoretische PGMs und Inferenz

A. *Nutzentheorie*
   - Präferenzen, Nutzen, Maximum-Expected-Utility Prinzip, Dominanz und Unabhängigkeiten

B. *Entscheidungstheoretische episodische PGMs und Inferenz*
   - Syntax: Knoten für Nutzen / Entscheidungen, Einflussdiagram, Entscheidungsmodell; Semantik
   - Inferenzaufgaben: Erwarteter Nutzen, Aktionen mit höchstem erwarteten Nutzen
   - Entscheidungstheoretische VE-Variante
   - Informationswerttheorie

C. *Entscheidungstheoretische sequentielle PGMs und Inferenz*
   - Zeitliche Inferenzaufgaben
   - VE-basierter Inferenzalgorithmus
   - Automatisiertes Handeln

# Präferenzen

- Agent wählt zwischen Gewinnen ($A$, $B$, etc.) und Lotterien, i.e., Situationen mit ungewissen Gewinnen
  - Ausgang einer nicht-deterministischen Aktion ist eine Lotterie, Aktion ist das Los
- Lotterie $L = [p, A; (1-p), B]$
  - $A$ und $B$ können wieder Lotterien sein
  - Gewinne sind spezielle Lotterien: $[1, Reward; 0, \text{not } Reward]$
  - Mehr als zwei Ausgänge:
    - $L = [p_1, S_1; p_2, S_2; \cdots; p_n, S_n], \sum_{i=1}^{n} p_i = 1$
- Notation
  - $A \succ B$ $\quad$ $A$ bevorzugt gegenüber $B$
  - $A \sim B$ $\quad$ Unentschlossenheit zwischen $A$ und $B$
  - $A \succsim B$ $\quad$ $B$ nicht bevorzugt gegenüber $A$

# Rational Präferenzen

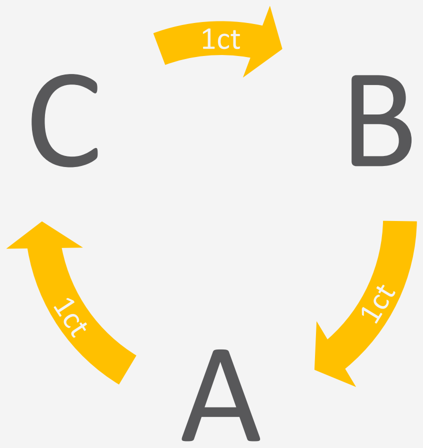
- Idee: Präferenzen eines rationalen Agenten müssen Einschränkungen folgen
  - Als Voraussetzung für sinnvolle Präferenzrelationen
- Rationale Präferenzen
  $\Rightarrow$ Verhalten des Agenten mittels Maximierung des erwarteten Nutzens beschreibbar

# Rationale Präferenzen

- Verletzen von Einschränkungen führt zu offensichtlicher Irrationalität

- Beispiel
  - Agent mit intransitiven Präferenzen kann dazu gebracht all sein Geld wegzugeben

  - Wenn $B \succ C$, dann würde ein Agent, der $C$ hat, (sagen wir mal) 1 Cent für $B$ zahlen
  - Wenn $A \succ B$, dann würde ein Agent, der $B$ hat, (sagen wir mal) 1 Cent für $A$ zahlen
  - Wenn $C \succ A$, dann würde ein Agent, der $A$ hat, (sagen wir mal) 1 Cent für $C$ zahlen

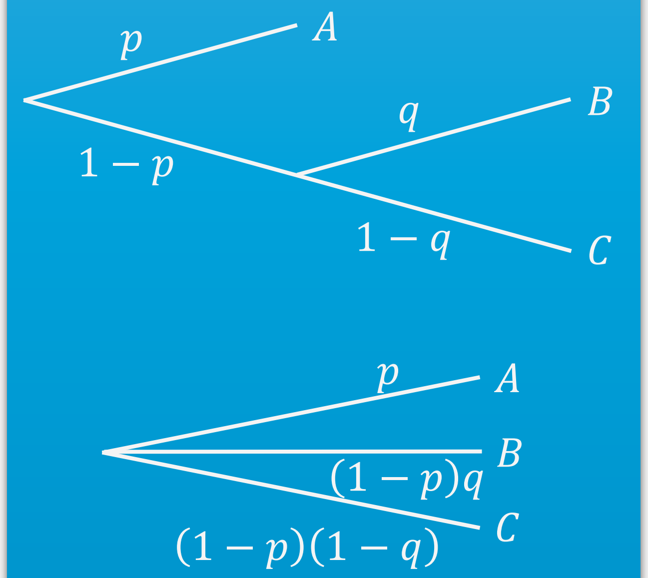


# Axiome der Nutzentheorie über Präferenzen

1. Sortierbarkeit (*orderability*)
   - $(A \succ B) \vee (A \prec B) \vee (A \sim B)$
     - $\{\prec, \succ, \sim\}$ zusammen exhaustiv, paarweise disjunkt
2. Transitivität (*transitivity*)
   - $(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C)$
3. Stetigkeit (*continuity*)
   - $A \succ B \succ C \Rightarrow \exists p\ [p, A;\ 1-p, C] \sim B$
4. Ersetzbarkeit (*substitutability*)
   - $A \sim B \Rightarrow [p, A; 1-p, C] \sim [p, B; 1-p, C]$
     - Gilt auch, wenn man $\sim$ mit $\succ$ ersetzt
5. Monotonie (*monotonicity*)
   - $A \succ B \Rightarrow (p \geq q \Leftrightarrow [p, A; 1-p, B] \succsim [q, A; 1-q, B])$
6. Zerlegbarkeit (*decomposability*)
   - $\big[p, A;\ 1-p, [q, B;\ 1-q, C]\big] \sim [p, A;\ (1-p)q, B; (1-p)(1-q), C]$

*Zerlegbarkeit:*
Kein Spaß beim Spielen.
*There is no fun in gambling.*

Äquivalente Lotterien:



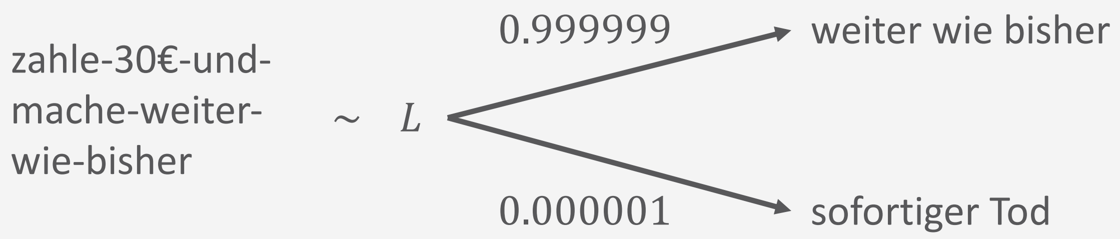
# Und dann gab es Nutzen

- Theorem (Ramsey, 1931; von Neumann und Morgenstern, 1944):
  - Gegeben Präferenzen unter diesen Einschränkungen gibt es eine reell-wertige Funktion $U$ mit
    $$U(A) \geq U(B) \Leftrightarrow A \gtrsim B$$
    - Existenz der Nutzenfunktion
  - Erwarteter Nutzen einer Lotterie:
    $$EU([p_1, S_1; \dots; p_n, S_n]) = \sum_{i=1}^{n} p_i U(S_i)$$
- MEU Prinzip
  - Wähle die Aktion, welche den erwarteten Nutzen maximiert

# Nutzenfunktionen

- Nutzenfunktionen bilden Zustände auf reelle Zahlen ab.
  Welche Zahlen?

- Standardvorgehen zur Einschätzung von Nutzen für einen Menschen:
  - Vergleiche gegebenen Zustand $A$ mit einer Standardlotterie $L_p$ (zwei Ausgänge $\top, \bot$), in der gilt:
    - "bestmöglicher Ausgang" $\top$ mit Wahrscheinlichkeit $p$
    - "schlimmste mögliche Katastrophe" $\bot$ mit Wahrscheinlichkeit $(1-p)$
  - Wahrscheinlichkeit $p$ anpassen, bis $A \sim L_p$ (Agent bzw. Mensch unentschlossen)



# Skalen für Nutzen

- Normalisierter Nutzen: $u_\top = 1.0, u_\perp = 0.0$
  - Nutzen der Lotterie $L \sim$ (zahle-30€-und-mache-weiter-wie-bisher): $U(L) = u_\top \cdot 0.999999 + u_\perp \cdot 0.000001 = 0.999999$
- Micromort: Todeswahrscheinlichkeit von 1 zu 1 Million
  - Nützlich für russisches Roulette, Bezahlen um Produktrisiken zu minimieren
  - Beispiel bei kleinen Risiken:
    - 370km mit Auto ≈ 1 Mikromort → 150.000km Auto Lebensdauer ≈ 400 Mikromort
    - Viele Menschen sind bereit 10.000$ für einen sicheren Wagen auszugeben, der Todeswahrscheinlichkeit halbiert → 50$ pro Mikromort
- QALYs: quality-adjusted life years
  - Äquivalent mit einem Jahr bei guter Gesundheit ohne irgendwelche Gebrechen
  - Nützlich bei medizinischen Entscheidungen, welche substanzielle Risiken mit sich bringen

# Nutzenfunktionen

- Verhalten ist invariant bezogen auf positive lineare Transformation
$$U'(r) = k_1 U(r) + k_2$$
  - Keine eindeutige Nutzenfunktion; $U'(r)$ und $U(r)$ führen zum gleichen Verhalten
- Ordinale Nutzenfunktion, auch Wertefunktion (*value function*) genannt
  - Mit rein deterministischen Ausgängen (ohne Lotteriewahlmöglichkeiten) kann nur ein ordinaler Nutzen bestimmt werden, i.e., eine Ordnung über Ausgänge
    - Ordnung total, wenn keine Ausgänge mit gleichem Nutzen vorkommen
    - Rangliste von Alternativen (Zuständen), aber keine bedeutsame metrische Skala (Zahlen sind egal)
- Nebenbemerkung:
Agent kann völlig rational handeln (konsistent mit dem MEU Prinzip), ohne jemals Nutzen und Wahrscheinlichkeiten zu repräsentieren oder zu manipulieren
  - Beispiel: Nachschlagetabelle für perfektes Tic-Tac-Toe

# Geld

- Geld verhält sich nicht wie eine Nutzenfunktion
- Gegeben eine Lotterie $L$ mit erwartetem monetärem Wert $EMV(L)$ gilt normalerweise $U(L) < U\left(S_{EMV(L)}\right)$, i.e., Menschen sind risikoscheu
  - $S_n$: Zustand, in dem man ein Gesamtvermögen von \$n hat
  - Nutzenkurve
    - Bei was für einer Wahrscheinlichkeit $p$ bin ich unentschlossen zwischen einem Gewinn $x$ und einer Lotterie $[p, \$M; (1-p), \$0]$ für große $M$?
    - Rechts: Typische empirische Daten extrapoliert mit risikobereitem Verhalten für negatives Vermögen

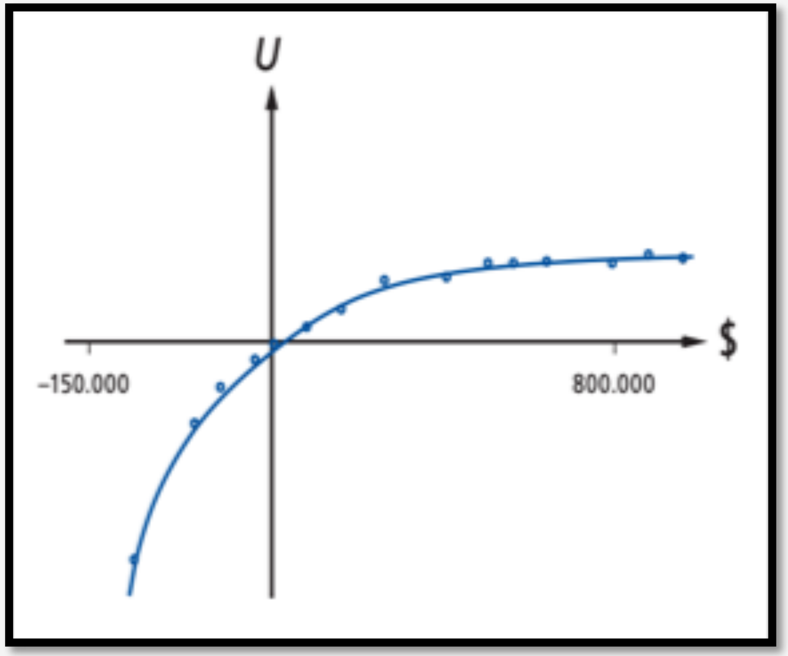



Abbildung: AIMAde, S. 720

# Geld versus Nutzen

- Geld $\neq$ Nutzen
  - Mehr Geld ist besser, aber nicht in einer linearen Beziehung zur Geldmenge
- Erwarteter monetärer Wert
  - Risikoscheu
    - $U(L) < U(S_{EMV(L)})$
  - Risikobereit
    - $U(L) > U(S_{EMV(L)})$
  - Risikoneutral
    - $U(L) = U(S_{EMV(L)})$
    - Lineare Kurve
    - Für kleine Änderungen im Vermögen relativ zum gegenwertigen Vermögen

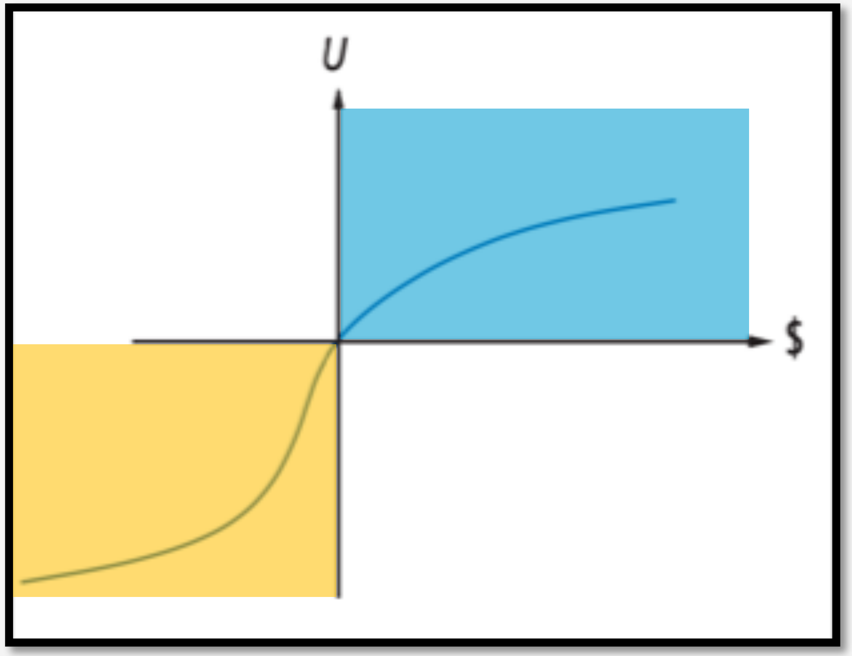



Abbildung: AIMAde, S. 720

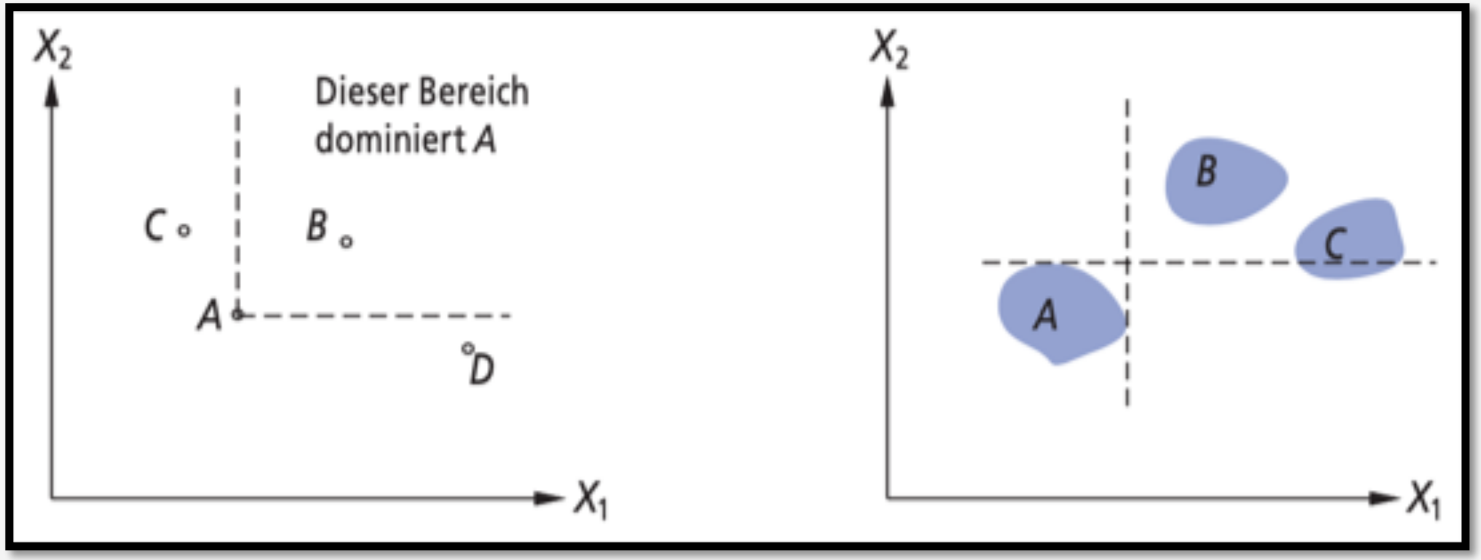
# Nutzentheorie mit Mehrfachattributen

- Ein gegebener Zustand kann mehrere Nutzen-Attribute haben
  - ...weil es mehrere Bewertungskriterien gibt
  - ...weil es mehrere Agenten (interessierte Parteien) mit verschiedenen Nutzenfunktionen gibt
  - Beispiel: Erstellung eines neuen Flughafens → Wahl des Standortes
    - Kriterien: Probleme die der Bau verursacht, welche Kosten für das Bauland anfallen, welcher Abstand zu Ballungsgebieten vorliegt, welchen Lärm der Flugbetrieb verursacht, welche Sicherheitsrisiken aus der lokalen Topografie und den Wetterbedingungen entstehen usw.
- Wir schauen uns an
  - Fälle, in denen Entscheidungen gemacht werden können *ohne* die Attributwerte in einen einzigen Nutzenwert zu kombinieren
    - Strenge Dominanz
  - Fälle, in denen der Nutzen von Attributkombinationen sehr knapp spezifiziert werden kann

# Strenge Dominanz

- Typischerweise Attribute so definieren, dass $U$ monoton in jeder Dimension ist →
- Strenge Dominanz
  - Wahl $B$ dominiert Wahl $A$ strikt genau dann, wenn

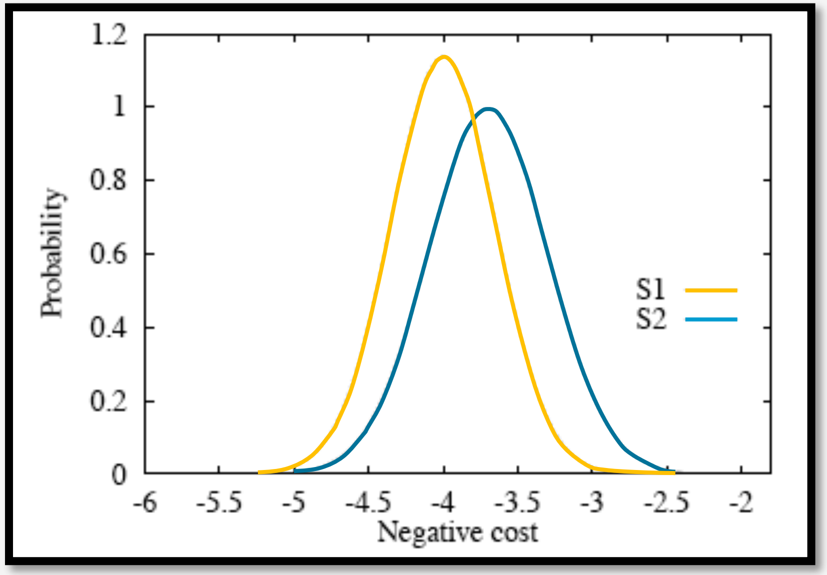
$$\forall\, i : X_i(B) \geq X_i(A) \text{ (und damit } U(B) \geq U(A))$$

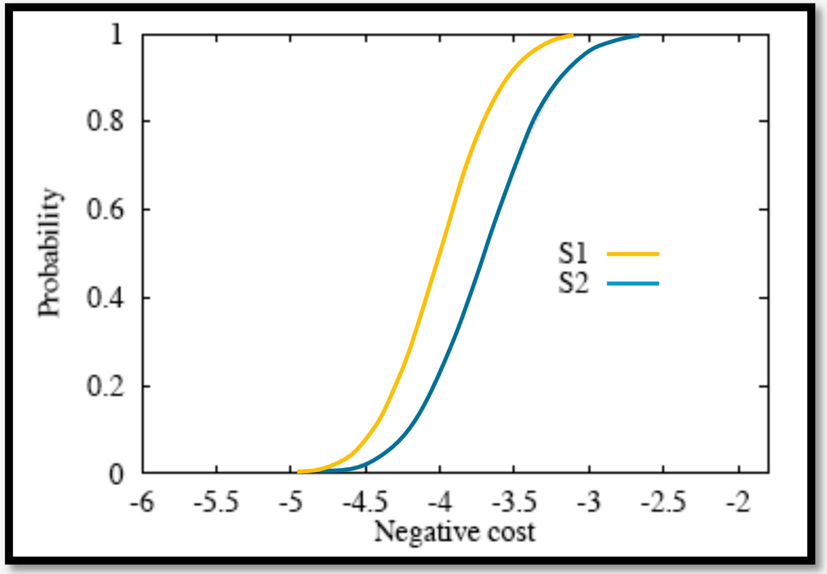



Deterministische Attribute | Unsichere Attribute

# Stochastische Dominanz

- Stochastische Dominanz erster Ordnung: Kumulative Verteilung $p_1$ dominiert Verteilung $p_2$ genau dann, wenn
$$\forall x: \; p_1(x) \leq p_2(x)$$
  - Mit einer strengen Ungleichung für ein Intervall
  - Dann gilt $E_{p_1} > E_{p_2}$ ($E$ bezieht sich auf den Erwartungswert)
    - Gegenteil nicht zwangsweise wahr
  - Impliziert nicht, dass jede mögliche Ausgabe der überlegenen Verteilung größer ist als jede mögliche Ausgabe der unterlegenen Verteilung
- Beispiel:
  - S2 dominiert S1 mit
$$\forall x: \; p_{S_2}(x) \leq p_{S_1}(x)$$



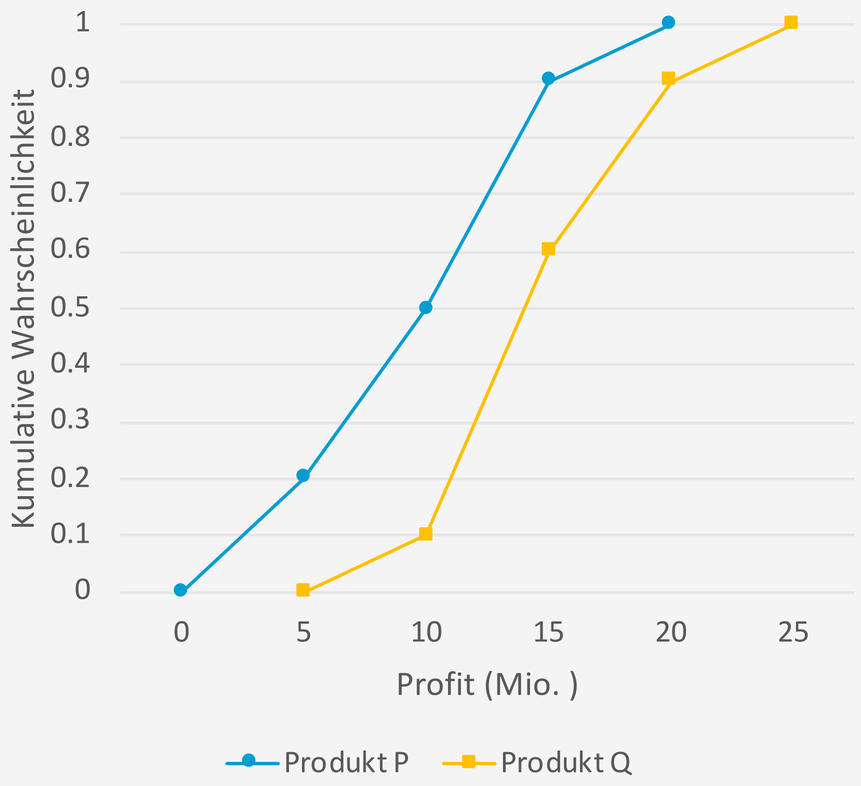
# Beispiel

- Produkt P

| Profit ($m) | Wahrscheinlichkeit |
|---|---|
| 0 bis unter 5 | 0.2 |
| 5 bis unter 10 | 0.3 |
| 10 bis unter 15 | 0.4 |
| 15 bis unter 20 | 0.1 |

- Produkt Q

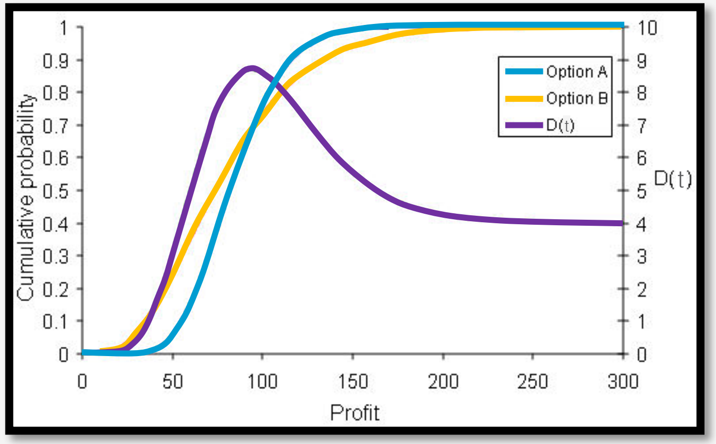
| Profit ($m) | Wahrscheinlichkeit |
|---|---|
| 0 bis unter 5 | 0.0 |
| 5 bis unter 10 | 0.1 |
| 10 bis unter 15 | 0.5 |
| 15 bis unter 20 | 0.3 |
| 20 bis unter 25 | 0.1 |

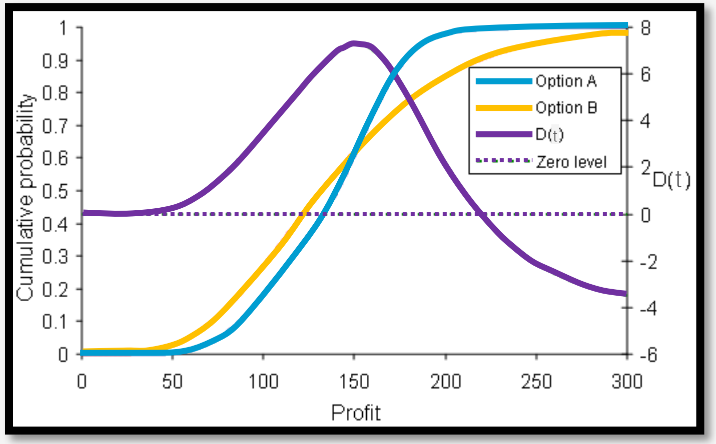



Q dominiert P stochastisch in erster Ordnung.

# Stochastische Dominanz

- Stochastische Dominanz zweiter Ordnung: Kumulative Verteilung $p_1$ dominiert Verteilung $p_2$ genau dann, wenn
$$\forall\, t : \int_{-\infty}^{t} p_1(x)\, dx \leq \int_{-\infty}^{t} p_2(x)\, dx$$
  - Oder: $D(t) = \int_{-\infty}^{t} p_2(x) - p_1(x)\, dx \geq 0$
  - Mit einer strengen Ungleichung für ein Intervall
  - Dann gilt $E_{p_1} \geq E_{p_2}$ ($E$ bezieht sich auf den Erwartungswert)
- Beispiel:
  - Oben: A dominiert B stochastisch in zweiter Ordnung
    - Keine stochastische Dominanz erster Ordnung
  - Unten: Weder A noch B dominiert ($D(t) < 0$)






https://people.duke.edu/~dgraham/ECO_463/Handouts/StochasticDominance.pdf

Abbildungen: https://www.vosesoftware.com/riskwiki/Stochasticdominancetests.php (t=z)

# Präferenzstruktur

- Um eine vollständige Nutzenfunktion $U(r_1, \ldots, r_n)$ zu spezifizieren, brauchen wir $d^n$ Nutzenwerte im Worst-Case
  - $n$ Attribute
  - Jedes Attribut mit $d$ individuellen möglichen Werten
  - Bedeutung Worst-Case: Präferenzen des Agenten haben überhaupt keine Regelmäßigkeiten
- Annahme in Nutzentheorie mit Mehrfachattributen
  - Präferenzen von typischen Agenten haben sehr viel mehr Struktur
- Vorgehen
  - Identifizieren von Regelmäßigkeiten im Präferenzverhalten
  - So genannte Repräsentationstheoreme nutzen um zu zeigen, dass ein Agent mit einer bestimmten Art von Präferenzstruktur eine Nutzenfunktion
$$U(r_1, \ldots, r_n) = F[f_1(r_1), \ldots, f_n(r_n)]$$
    - hat, wobei $F$ hoffentlich eine einfache Funktion wie Addition ist

# Präferenzstruktur : Deterministisch

- $R_1$ und $R_2$ präferenzunabhängig (*preferentially independent*, PI) von $R_3$ genau dann, wenn
  - Präferenz zwischen $\langle r_1, r_2, r_3 \rangle$ und $\langle r_1', r_2', r_3 \rangle$ nicht von $r_3$ abhängig ist
  - Beispiel: Flughafenstandort mit Attributen $\langle Noise, Cost, Deaths \rangle$
    - $\langle 20{,}000\ suffer, \$4.6\ billion, 0.06\ deaths/month \rangle$
    - $\langle 70{,}000\ suffer, \$4.2\ billion, 0.06\ deaths/month \rangle$

- Theorem (Leontief, 1947)
  - Wenn jedes Paar von Attributen PI von seinem Komplement ist, dann ist jede Untermenge von Attributen PI von seinem Komplement
    - Gegenseitige PI (*mutual PI*, MPI) genannt

# Präferenzstruktur : Deterministisch

- Theorem (Debreu, 1960):
  - MPI ⇒ ∃ *additive* Wertefunktion

$$V(r_1, \dots, r_n) = \sum_{i=1}^{n} V_i(r_i)$$

  - Damit $n$ eindimensionale Wertefunktionen benötigt
    - Zerlegung von $V$ in Menge von *Summanden*

    von der Idee her gleich zur

    - Zerlegung von $P_{\boldsymbol{R}}$ in Menge von *Faktoren*
  - Häufig gute Approximation
  - Beispiel:

$$V(Noise, Cost, Deaths) = -Noise \cdot 10^4 - Cost - Deaths \cdot 10^{12}$$

# Präferenzstruktur: Stochastisch

- Präferenzen über Lotterien
- $R$ ist nutzenunabhängig (*utility-independent*, UI) von $S$ genau dann, wenn
  - Präferenzen über Lotterien in $R$ nicht von $s$ abhängen
- Theorem (Keeney, 1974):
  - Jede Untermenge ist UI von ihrem Komplement (gegenseitige UI ,*mutual UI*)
    $\Rightarrow \exists$ *multiplikative* Nutzenfunktion
  - Für $n = 3$:
$$\begin{aligned} U = {} & k_1U_1 + k_2U_2 + k_3U_3 \\ & + k_1k_2U_1U_2 + k_2k_3U_2U_3 + k_3k_1U_3U_1 \\ & + k_1k_2k_3U_1U_2U_3 \end{aligned}$$
  - I.e., benötigt $n$ eindimensionale Nutzenfunktionen und $n$ Konstanten

# Zwischenzusammenfassung

- Präferenzen
  - Präferenzen eines rationalen Agenten müssen Einschränkungen erfüllen
- Nutzen
  - Rationale Präferenzen = Maximierung des erwarteten Nutzens möglich
  - Axiome
  - MEU Prinzip
- Dominanz
  - Strenge Dominanz
  - Stochastische Dominanz erster und zweiter Ordnung
- Präferenzstruktur
  - (Gegenseite) Präferenzunabhängigkeit
  - (Gegenseite) Nutzenunabhängigkeit

# Überblick: 7. Entscheidungstheoretische PGMs und Inferenz

A. *Nutzentheorie*
- Präferenzen, Nutzen, Maximum-Expected-Utility Prinzip, Dominanz und Unabhängigkeiten

B. ***Entscheidungstheoretische episodische PGMs und Inferenz***
- Syntax: Knoten für Nutzen / Entscheidungen, Einflussdiagram, Entscheidungsmodell; Semantik
- Inferenzaufgaben: Erwarteter Nutzen, Aktionen mit höchstem erwarteten Nutzen
- Entscheidungstheoretische VE-Variante
- Informationswerttheorie

C. *Entscheidungstheoretische sequentielle PGMs und Inferenz*
- Zeitliche Inferenzaufgaben
- VE-basierter Inferenzalgorithmus
- Automatisiertes Handeln

# Entscheidungsmodelle

- Erweitere ein PGM um Aktionen und Nutzen zu berücksichtigen
  - Entscheidungsvariablen
  - Nutzenvariablen
- Auch Entscheidungsnetze bzw. für BNs, Einflussdiagramme (*influence diagrams*) genannt
- Gegeben ein Entscheidungsmodell, nutze beliebigen Inferenzalgorithmus um Aktionen zu finden, die zum höchsten erwarteten Nutzen führen

- Ermöglicht auch so genannte *Informationswert*-Berechnungen
  - Ist es es wert Ressourcen darauf zu verwenden mehr Information (in Form von Evidenz) zu besorgen?

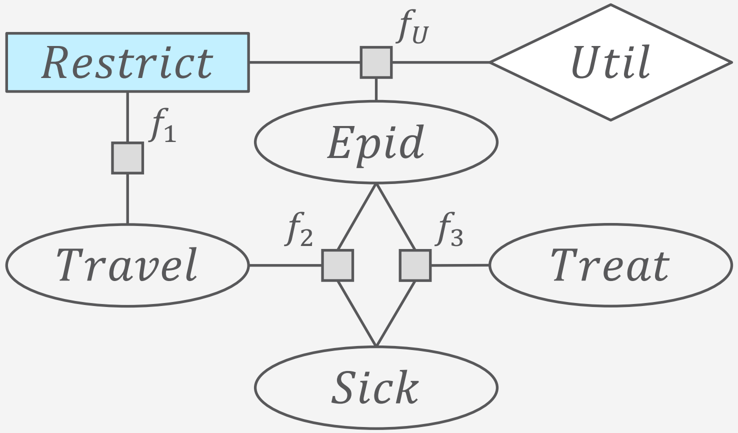


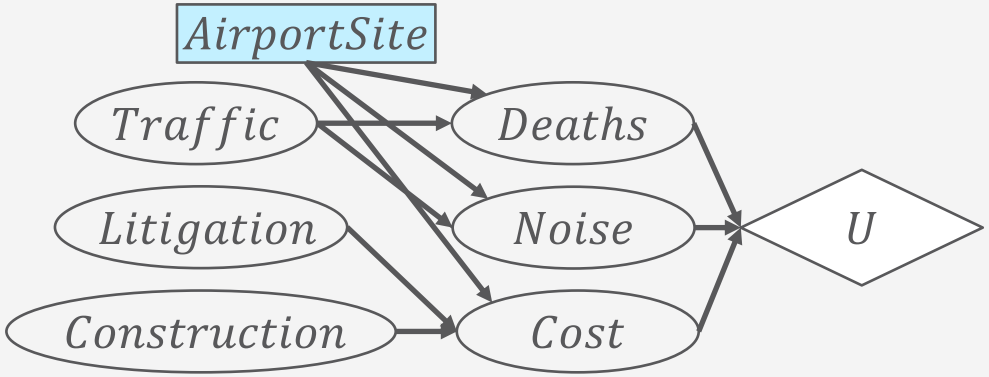
Ronald A. Howard: Information Value Theory. In: *IEEE Transactions on Systems Science and Cybernetics*, 1966.
Ronald A. Howard, James E. Matheson: Influence Diagrams. In: *Readings on the Principles and Applications of Decision Analysis*, 1984.

# Entscheidungsvariablen

- Entscheidungsvariable $D$
  - Zufallsvariable, die immer eine Zuweisung erhält
    - Man kann sich nicht nicht entscheiden!
- Menge von Entscheidungsvariablen $\boldsymbol{D}$ im Modell, i.e., $\boldsymbol{R} = \boldsymbol{D} \cup \boldsymbol{V}$
  - Im Faktormodell: als Argument eines Faktors $\phi(\mathcal{R})$: $D \in \mathcal{R}$
    - Beispiel: $\phi_1(Restrict, Travel)$
  - Im BN: als Knoten mit CPD $P(D|\mathrm{Pa}(D))$
    - Kommt als Elternknoten in weiteren CPDs $P(R_i|\mathrm{Pa}(R_i))$, $D \in \mathrm{Pa}(R_i)$ vor

| $Restrict$ | $Travel$ | $\phi_1$ |
|---|---|---|
| $free$ | $false$ | 1 |
| $free$ | $true$ | 1 |
| $ban$ | $false$ | 1 |
| $ban$ | $true$ | 0 |



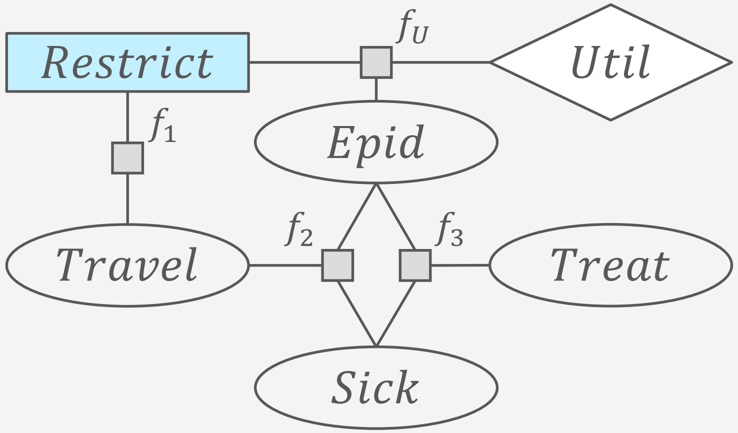


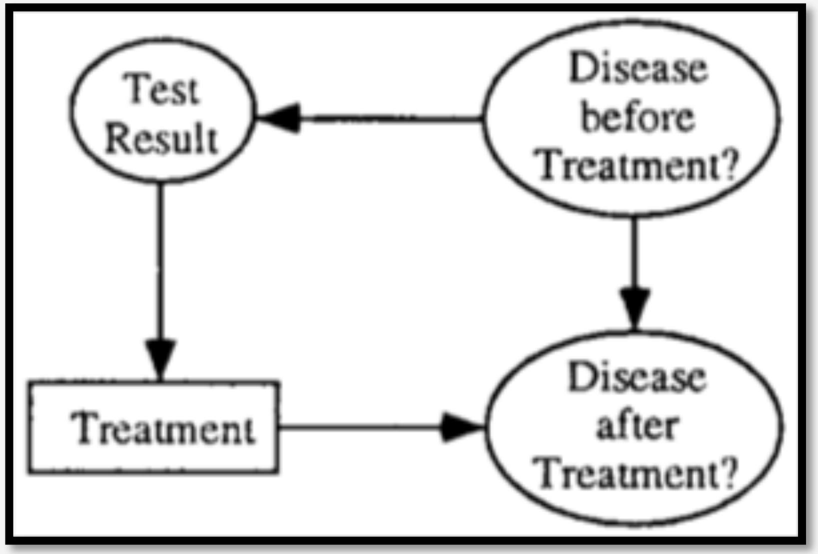

Ronald A. Howard, James E. Matheson: Influence Diagrams. In: *Readings on the Principles and Applications of Decision Analysis*, 1984.

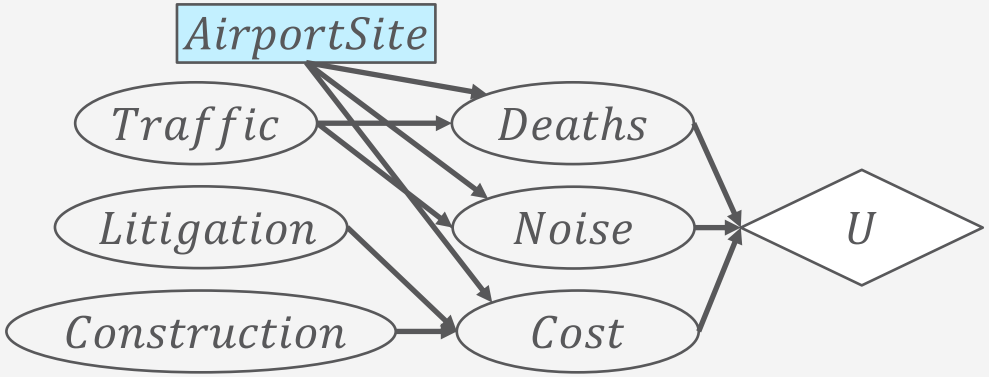
# Entscheidungsvariablen

Nach Howard und Matheson (1984) bedeutet es für eingehende Kanten der Entscheidungsvariable, dass die Werte diese Variablen zur Entscheidungsfindung vorliegen (Evidenz).

| $Restrict$ | $Travel$ | $\phi_1$ |
|---|---|---|
| $free$ | $false$ | 1 |
| $free$ | $true$ | 1 |
| $ban$ | $false$ | 1 |
| $ban$ | $true$ | 0 |



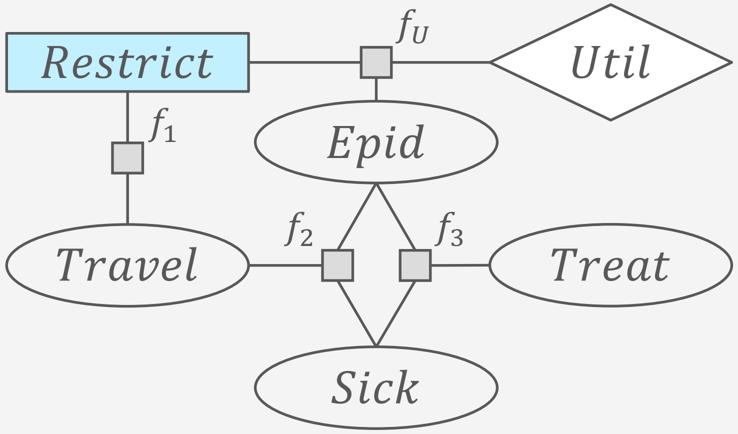


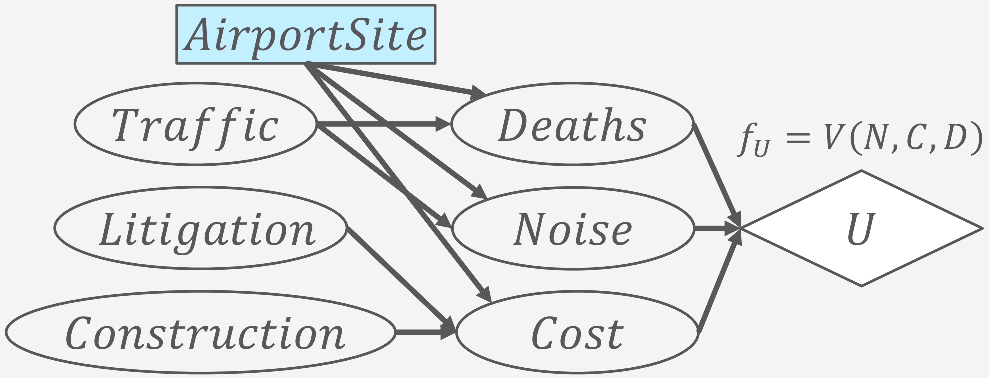

Ronald A. Howard, James E. Matheson: Influence Diagrams. In: *Readings on the Principles and Applications of Decision Analysis*, 1984.
Abbildung: Robert F. Nease, Jr., Douglas K. Owens: Use of Influence Diagrams to Structure Medical Decisions . In: Medical Decision Making, 1997.

# Nutzenvariablen in Faktoren

- Nutzenvariable $U$: erhält Wert von der Nutzen- bzw. Wertefunktion zugewiesen
- In Faktormodellen und BNs: Nutzenfaktor $\phi_U(\mathcal{R})$
  - Argumente $\mathcal{R} = (R_1, \dots, R_m)$ eine Sequenz von Zufalls- bzw. Entscheidungsvariablen
    - $\mathrm{rv}\big(\phi_U(\mathcal{R})\big) = \{R_1, \dots, R_m\}$
  - $U$ eine Nutzenvariable
  - Funktion $\phi: \times_{i=1}^{m} \mathrm{Val}(R_i) \mapsto \mathrm{Val}(U)$
    - Tabellarisch oder additive / multiplikative Kombination
      - Beispiel tabellarisch: $\phi_{Util}(Restrict, Epid)$
      - Beispiel von Folie 27 additiv: $V(N, C, D) = -Noise \cdot 10^4 - Cost - Deaths \cdot 10^{12}$

| $Restrict$ | $Epid$ | $Util$ |
|---|---|---|
| $free$ | $false$ | 10 |
| $free$ | $true$ | −10 |
| $ban$ | $false$ | −20 |
| $ban$ | $true$ | 5 |



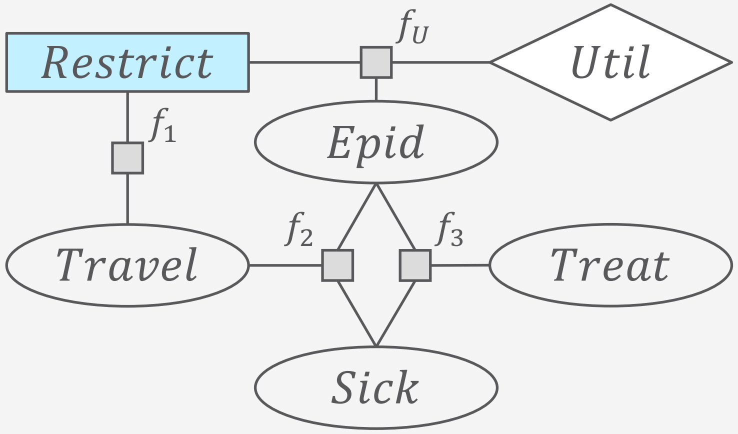
# Entscheidungsmodell

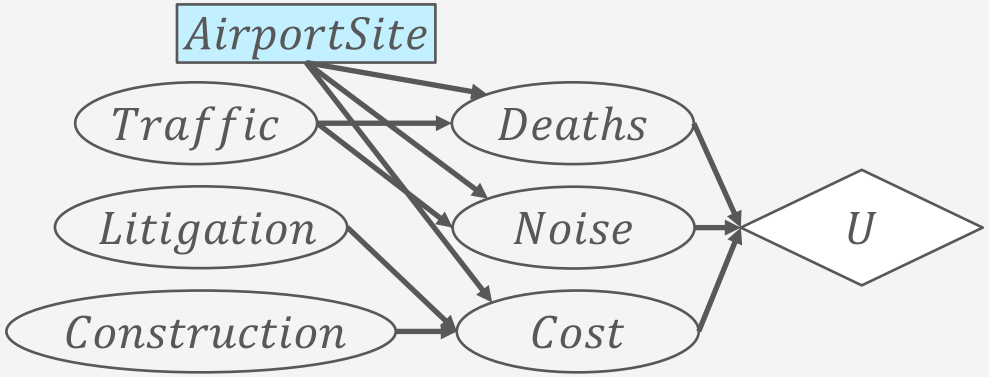
- Entscheidungsmodell = Modell, welches Entscheidungsvariablen in den Argumenten seiner Faktoren und Nutzenfaktoren zulässt
  - Der Einfachheit halber konzentrieren wir uns erstmal auf Modelle mit einem Nutzenfaktor, der auf eine Nutzenvariable abbildet
  - Formal

$$M = \{m_i\}_{i=1}^n \cup \{m_U\}$$

    - Bei Faktormodellen: $m_i$ Faktoren $f_i$, $m_U$ Nutzenfaktor $f_U$
    - Bei BNs: $m_i$ CPDs $P(R_i|\mathrm{Pa}(R_i))$, $m_U$ Nutzenfaktor $f_U$
  - Beispiel:
    - $F = \{f_1, f_2, f_3, f_U\}$
    - $B = \{P(T), P(L), P(Cn), P(A), P(D|T,A), P(N|T,A), P(C|Cn,A), V(D,N,C)\}$

Apriori-Informationen zu den Aktionen, oft aber gleichverteilt → weglassen möglich

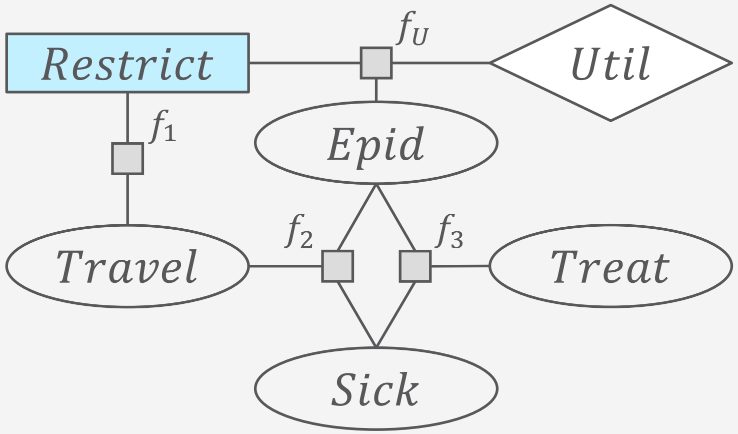



$T \triangleq Traffic, L \triangleq Litigation, Cn \triangleq Construction, A \triangleq AirportLocation, D \triangleq Deaths, N \triangleq Noise, C \triangleq Cost$

# Aktionszuweisungen

- $\boldsymbol{D} = \{D_1, \dots, D_k\}$ die Menge der Entscheidungsvariablen in Entscheidungsmodell $M$
- $\boldsymbol{d}$ ein zusammengesetztes Event für $\boldsymbol{D}$, welches jedem $D_i$ einen Wert (Aktion) $d_i$ zuweist
  - Wir nennen $\boldsymbol{d}$ eine Aktionszuweisung
- Beispiel: Mögliche Aktionszuweisungen
  - Entscheidungsvariable $Restrict$ mit Domäne $\{ban, free\}$
    - $\boldsymbol{d}_1 = \{ban\}$
    - $\boldsymbol{d}_2 = \{free\}$
  - Gegeben eine weitere Entscheidungsvariable $D$ mit Domäne $\{d', d'', d'''\}$
    - $\boldsymbol{d}_1 = \{ban, d'\}$
    - $\boldsymbol{d}_2 = \{ban, d''\}$
    - $\boldsymbol{d}_3 = \{ban, d'''\}$
    - $\boldsymbol{d}_4 = \{free, d'\}$
    - $\boldsymbol{d}_5 = \{free, d''\}$
    - $\boldsymbol{d}_6 = \{free, d'''\}$



# Entscheidungen machen: Aktionen zuweisen

- Gegeben Modell $M$ und Aktionszuweisung $\boldsymbol{d}$
- Setze $\boldsymbol{d}$ in $M$, i.e.,

$$M_{\boldsymbol{d}} = \bigcup_{d \in \boldsymbol{d}} \bigcup_{m \in M} \text{absorb}\big(m, D, \phi_d(D)\big)$$

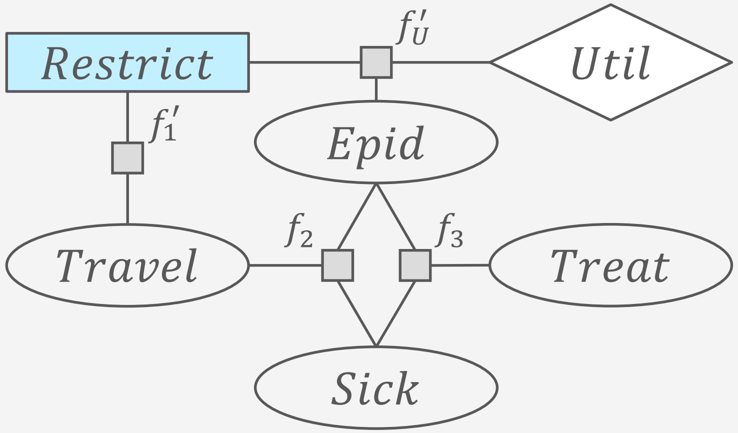
  - $\phi_d(D)$ eine $\langle 1{,}0 \rangle$-Verteilung gemäß $d \in \boldsymbol{d}$
  - In Faktor / CPD $m$ mit Entscheidungsvariable $D$,
    - Lasse die Abbildungen / Einträge $D \neq d$ und danach $D$ fallen
  - Beispiel: Setze $\boldsymbol{d}_1 = \{ban\}$ in $F = \{f_1, f_2, f_3, f_U\}$
    - Absorbiere $\boldsymbol{d}_1$ in $f_1$ und $f_U$
    - $F_{\boldsymbol{d}_1} = \{f_1', f_2, f_3, f_U'\}$
      - $f_1' = \phi_1'(Travel)$
      - $f_U' = \phi_{Util}'(Epid)$

| $Restrict$ | $Travel$ | $\phi_1$ |
|---|---|---|
| ~~$free$~~ | ~~$false$~~ | ~~1~~ |
| ~~$free$~~ | ~~$true$~~ | ~~1~~ |
| $ban$ | $false$ | 1 |
| $ban$ | $true$ | 0 |

| $Restrict$ | $Epid$ | $Util$ |
|---|---|---|
| ~~$free$~~ | ~~$false$~~ | ~~10~~ |
| ~~$free$~~ | ~~$true$~~ | ~~$-10$~~ |
| $ban$ | $false$ | $-20$ |
| $ban$ | $true$ | 5 |

| $Travel$ | $\phi_1'$ |
|---|---|
| $false$ | 1 |
| $true$ | 0 |

| $Epid$ | $Util$ |
|---|---|
| $false$ | $-20$ |
| $true$ | 5 |



# Semantik des Entscheidungsmodells

- Semantik eines Entscheidungsmodelle $M = \{m_i\}_{i=1}^{n} \cup \{m_U\}$ über Entscheidungsvariablen $\boldsymbol{D} = \{D_1, \dots, D_k\}$ und verbleibende Zufallsvariablen $\boldsymbol{V} = \{V_1, \dots, V_l\}$
  - Gegeben Aktionszuweisung $\boldsymbol{d}$ für $\boldsymbol{D}$ → Müssen immer gesetzt sein (keine Nicht-Entscheidung!)
  - Vollständige gemeinsame Wahrscheinlichkeitsverteilung der Nicht-Nutzenfaktoren

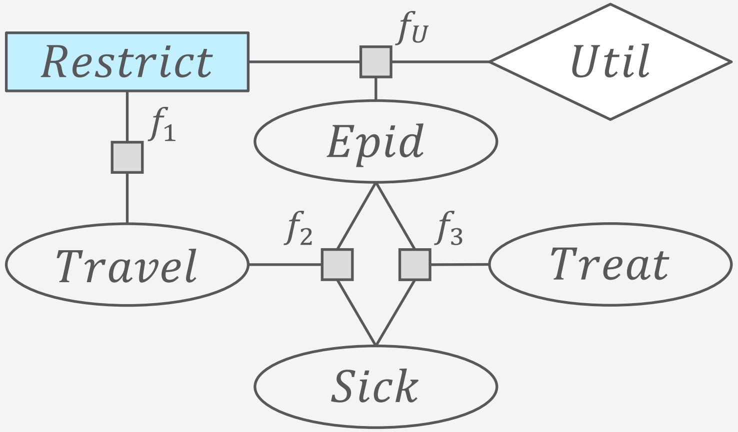
$$P_M[\boldsymbol{d}] = \frac{1}{Z} \prod_{m \in M_{\boldsymbol{d}} \setminus \{m_U\}} m = \boldsymbol{P}_{M_{\boldsymbol{d}}}$$

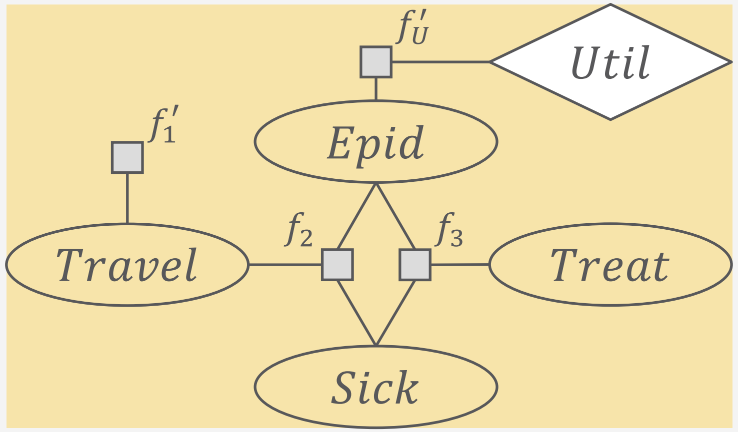
$$Z = \sum_{r_1 \in \mathrm{Val}(V_1)} \dots \sum_{r_N \in \mathrm{Val}(V_l)} \prod_{m \in M_{\boldsymbol{d}} \setminus \{m_U\}} m$$

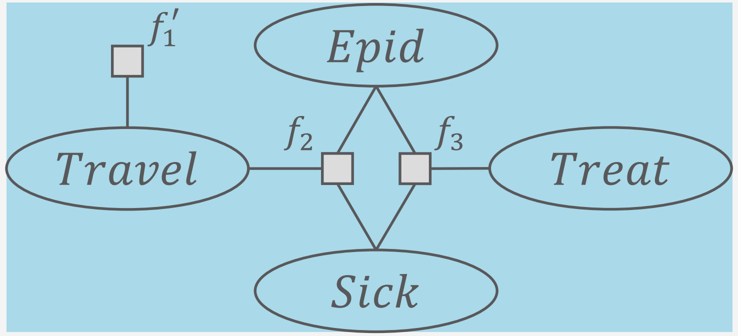
*Multiplikative* Semantik mit einem inneren Produkt und äußerer Summe: Faktoren multiplizieren, dann Zufallsvariablen aussummieren.
→ Summen-Produkt-Algorithmus (*sum-product algorithm*)

  - Nutzenfaktoren irrelevant für probabilistisches Verhalten

# Beispiel




- Links: Entscheidungsmodell $F = \{f_1, f_2, f_3, f_U\}$
- Mitte: $F$ mit $\boldsymbol{d}_1 = \{ban\}$ gesetzt $F_{\boldsymbol{d}_1} = \{f'_1, f_2, f_3, f'_U\}$
  - $f'_1 = \phi'_1(Travel)$
  - $f'_U = \phi'_{Util}(Epid)$
- Rechts Modell relevant für bisherige Anfragetypen: $F = \{f'_1, f_2, f_3\}$

# Anfragentypen in Entscheidungsmodellen

- Gegeben Entscheidungsmodell $M = \{m_i\}_{i=1}^{n} \cup \{m_U\}$ über Zufallsvariablen $\boldsymbol{R}$
  - Anfragen an (bedingte) Wahrscheinlichkeitsverteilungen gegeben eine Aktionszuweisung $\boldsymbol{d}$ für die Entscheidungsvariablen $\boldsymbol{D} \subseteq \boldsymbol{R}$ gegeben die Semantik für den probabilistischen Teil, $P_M[\boldsymbol{d}]$
  - Neuer Anfragetyp: Anfrage zu einem erwarteten Nutzen (*expected utility*, EU)
    - Was ist der erwartete Nutzen der Entscheidungen $\boldsymbol{d}$ in $M$ gegeben Evidenz $\boldsymbol{e}$?

$$eu(\boldsymbol{e}, \boldsymbol{d}) = \sum_{\boldsymbol{s} \in \mathrm{Val}(\mathrm{rv}(m_U))} P(\boldsymbol{s}|\boldsymbol{e}, \boldsymbol{d}) \cdot \phi_U(\boldsymbol{s})$$

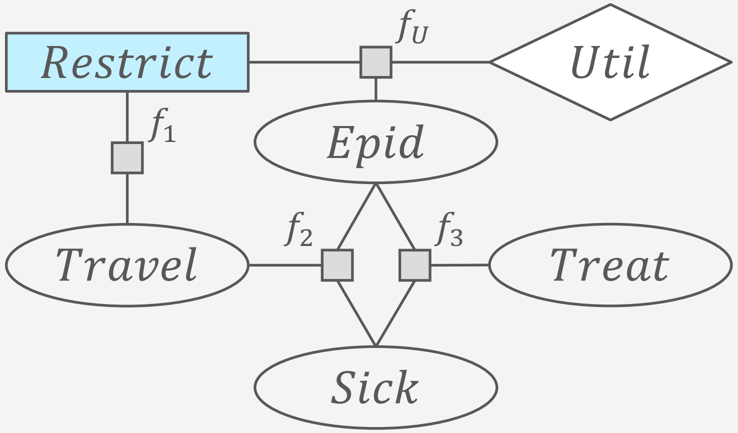
    - Summe über alle möglichen Ausgänge (in Form von Zuständen) der Entscheidungen
      - Produkt der Wahrscheinlichkeit einen Zustand $\boldsymbol{s}$ zu erreichen und des Nutzens in dem Zustand
    - $P(\boldsymbol{s}|\boldsymbol{e}, \boldsymbol{d})$ beinhaltet alle weiteren Zufallsvariablen in $M$, i.e., $\boldsymbol{R} \setminus \mathrm{rv}(\boldsymbol{e}) \setminus \boldsymbol{D} \setminus \underbrace{\mathrm{rv}(m_u)}_{= \boldsymbol{S}}$, zu eliminieren

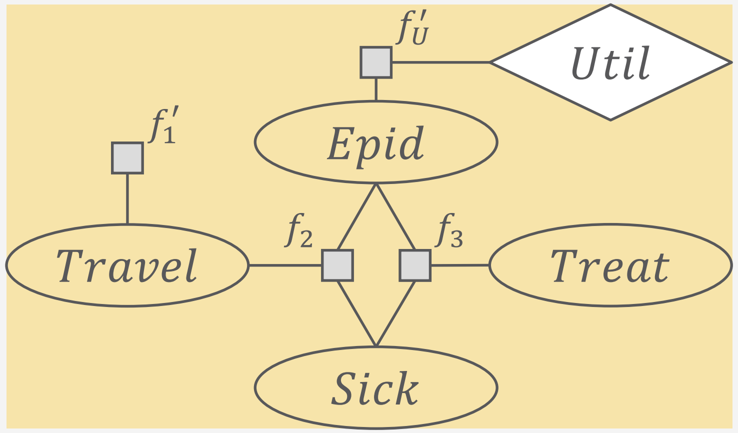
# EU Anfrage: Beispiel

- Erwarteter Nutzen von Aktionszuweisung $\boldsymbol{d}_1 = \{ban\}$ in $F = \{f_1, f_2, f_3, f_U\}$ ohne Evidenz ($\boldsymbol{e} = \emptyset$)

$$eu(\boldsymbol{d}_1) = \sum_{v \in \mathrm{Val}(Epid)} P(Epid = v | \boldsymbol{d}_1) \cdot \phi_U(Epid = v)$$

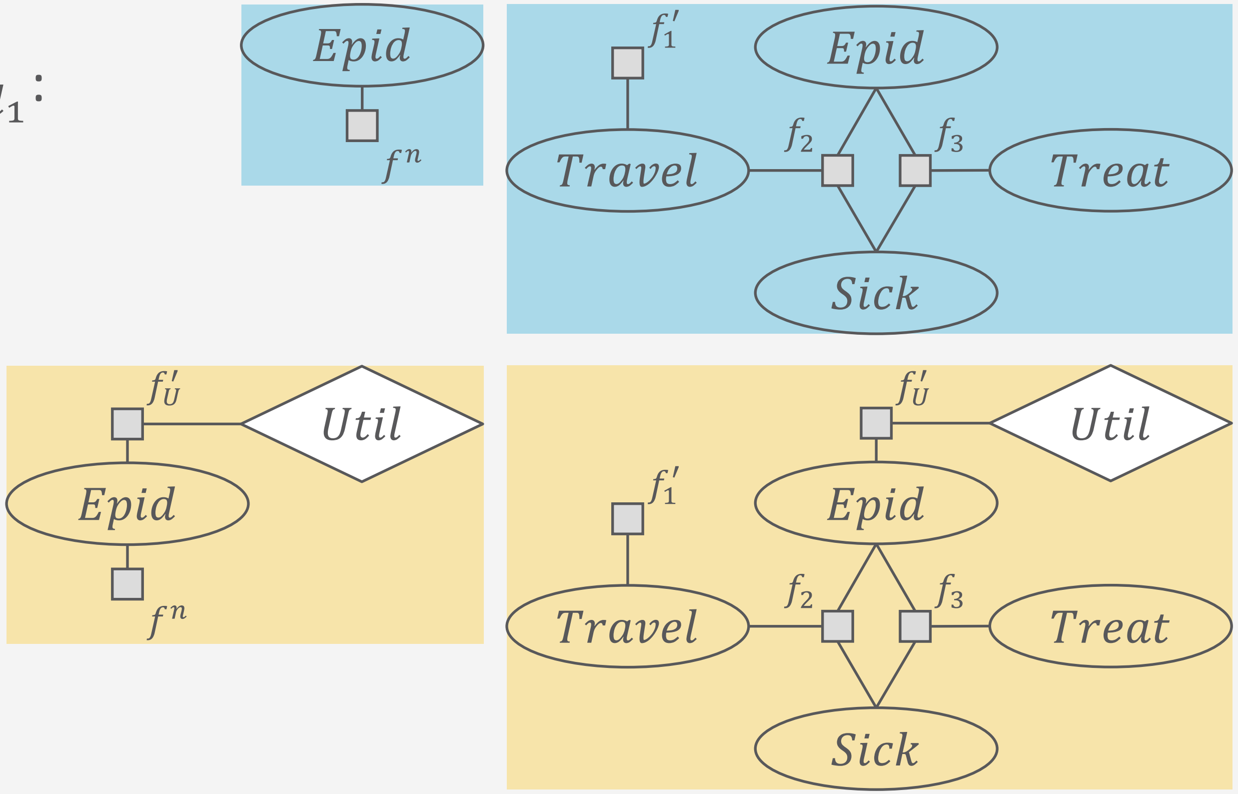
1. Berechne $P(Epid = v | \boldsymbol{d}_1)$ in $F$
   - durch Berechnung von $P(Epid = v)$ in $F_{\boldsymbol{d}_1} = \{f_1', f_2, f_3, f_U'\}$ (ohne Berücksichtigung der Nutzenfaktoren)
2. Berechne $eu(\boldsymbol{d}_1)$ mit $P(Epid = v | \boldsymbol{d}_1)$



# EU Anfrage: Beispiel

1. Berechne $P(Epid = v)$ in $F_{\boldsymbol{d}_1} = \{f_1', f_2, f_3, f_U'\}$
   - Eliminiere mittels VE alle Nicht-Anfragen-Zufallsvariablen (ohne Nutzenvariablen) in $F_{\boldsymbol{d}_1}$:
     - Eliminiere $Treat, Travel, Sick$
     - Ergebnis: $f = \phi(Epid)$
     - Normalisiere $f$ um Wahrscheinlichkeit $P(Epid = v)$ in $F_{\boldsymbol{d}_1}$ zu erhalten: $f^n = \phi^n(Epid)$
       - Entspricht $P(Epid = v|\boldsymbol{d}_1)$ in $F$

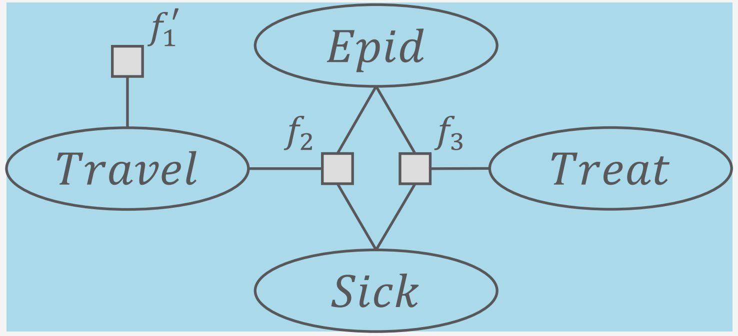


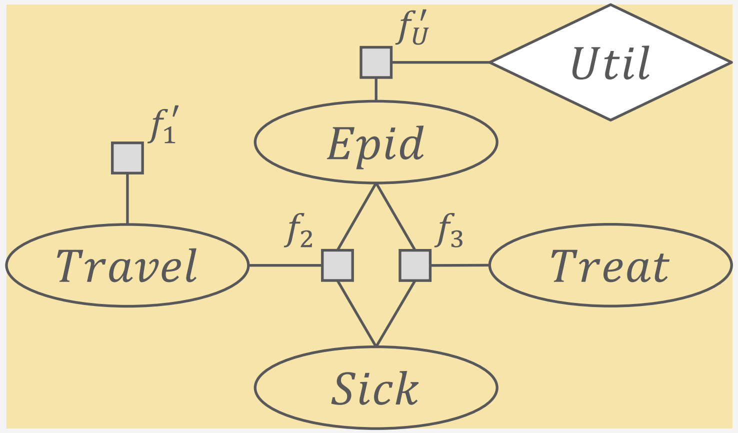
# EU Anfrage: Beispiel

- Eliminiere $Treat$: Aus $f_3$ aussummieren

| $Epid$ | $Sick$ | $Treat$ | $\phi_3$ |
|---|---|---|---|
| $false$ | $false$ | $false$ | 5 |
| $false$ | $false$ | $true$ | 1 |
| $false$ | $true$ | $false$ | 3 |
| $false$ | $true$ | $true$ | 2 |
| $true$ | $false$ | $false$ | 5 |
| $true$ | $false$ | $true$ | 4 |
| $true$ | $true$ | $false$ | 1 |
| $true$ | $true$ | $true$ | 7 |

| $Epid$ | $Sick$ | $\phi'_3$ |
|---|---|---|
| $false$ | $false$ | $5+1=6$ |
| $false$ | $true$ | $3+2=5$ |
| $true$ | $false$ | $5+4=9$ |
| $true$ | $true$ | $1+7=8$ |



# EU Anfrage: Beispiel

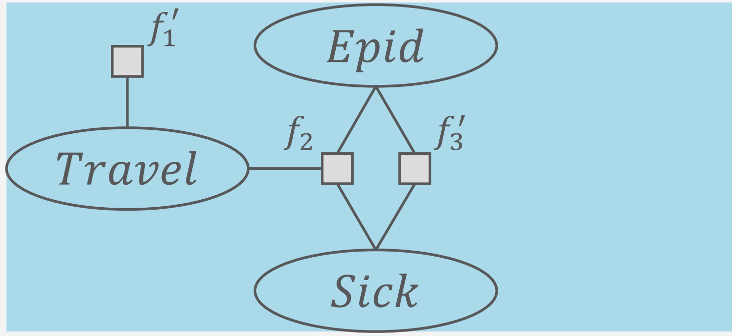
| $Travel$ | $\phi'_1$ |
|---|---|
| $false$ | 1 |
| $true$ | 0 |

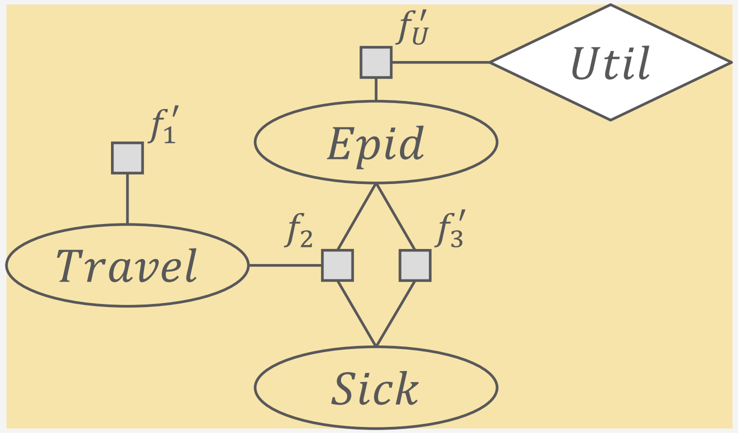
- Eliminiere $Travel$: Multipliziere $f'_1, f_2$, summiere $Travel$ aus Ergebnis aus

| $Travel$ | $Epid$ | $Sick$ | $\phi_2 \cdot \phi'_1$ |
|---|---|---|---|
| $false$ | $false$ | $false$ | $20 \cdot 1 = 20$ |
| $false$ | $false$ | $true$ | $24 \cdot 1 = 24$ |
| $false$ | $true$ | $false$ | $5 \cdot 1 = 5$ |
| $false$ | $true$ | $true$ | $6 \cdot 1 = 6$ |
| $true$ | $false$ | $false$ | $28 \cdot 0 = 0$ |
| $true$ | $false$ | $true$ | $8 \cdot 0 = 0$ |
| $true$ | $true$ | $false$ | $7 \cdot 0 = 0$ |
| $true$ | $true$ | $true$ | $2 \cdot 0 = 0$ |

+
+
+
+

| $Epid$ | $Sick$ | $\phi'_{12}$ |
|---|---|---|
| $false$ | $false$ | $20 + 0 = 20$ |
| $false$ | $true$ | $24 + 0 = 24$ |
| $true$ | $false$ | $5 + 0 = 5$ |
| $true$ | $true$ | $6 + 0 = 6$ |



# EU Anfrage: Beispiel

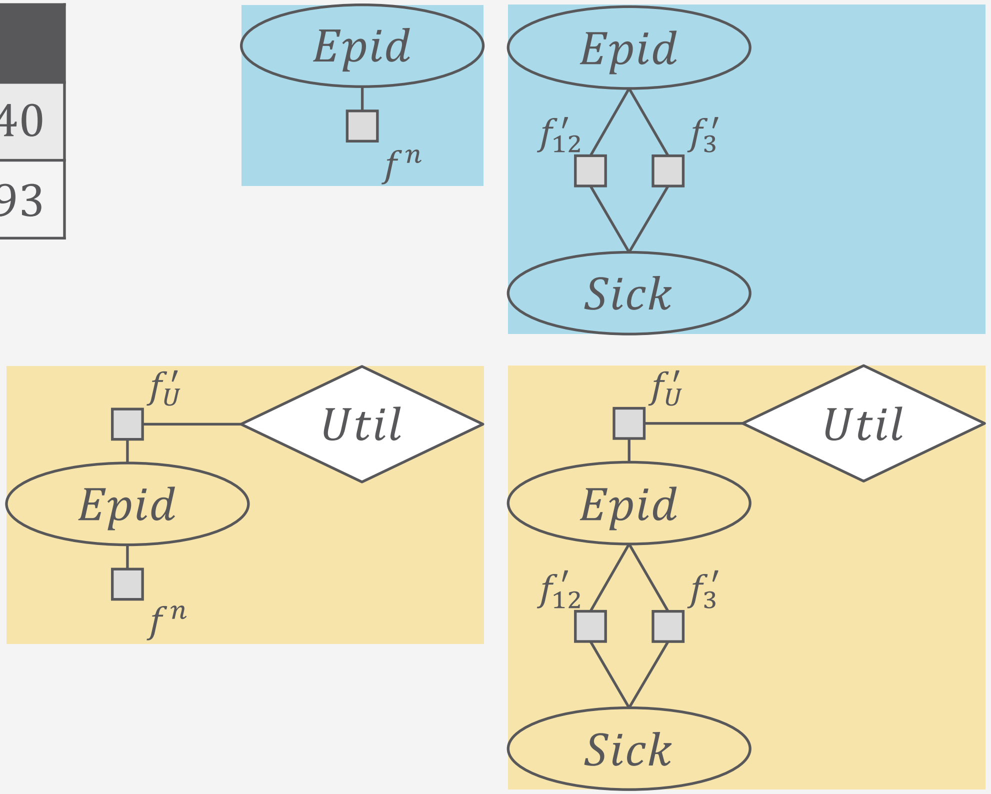
- Eliminiere $Sick$: Multipliziere $f'_{12}, f'_3$, summiere $Sick$ aus Ergebnis aus, normalisiere

| $Epid$ | $Sick$ | $\phi'_{12} \cdot \phi'_3$ |
|---|---|---|
| $false$ | $false$ | $20 \cdot 6 = 120$ |
| $false$ | $true$ | $24 \cdot 5 = 120$ |
| $true$ | $false$ | $5 \cdot 9 = 045$ |
| $true$ | $true$ | $6 \cdot 8 = 048$ |

| $Epid$ | $\phi$ |
|---|---|
| $false$ | $120 + 120 = 240$ |
| $true$ | $45 + 48 = 93$ |

Ergebnis nach Normalisierung: $f^n = \phi^n(Epid)$

| $Epid$ | $\phi^n$ |
|---|---|
| $false$ | $0.721$ |
| $true$ | $0.279$ |



# EU Anfrage: Beispiel

2. Berechne $eu(\boldsymbol{d}_1)$ mit $P(Epid = v|\boldsymbol{d}_1)$
   - Ergebnis $\phi^n(Epid)$ für $P(Epid|\boldsymbol{d}_1)$ mit $\phi'_U(Epid)$ multiplizieren, $Epid$ eliminieren

$$eu(\boldsymbol{d}_1)$$

$$= \sum_{v \in \mathrm{Val}(Epid)} P(Epid = v|\boldsymbol{d}_1) \cdot \phi'_U(Epid = v)$$

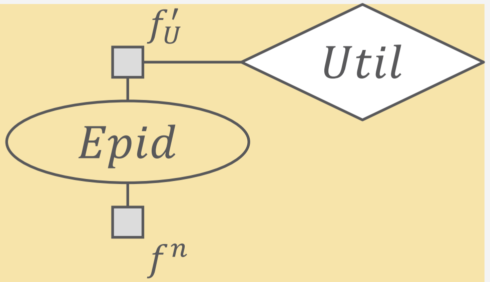
$$= \sum_{v \in \mathrm{Val}(Epid)} \phi^n(Epid = v) \cdot \phi'_U(Epid = v)$$

$$= \sum_{v \in \mathrm{Val}(Epid)} \phi''_U(Epid = v) = \phi'''_U(.)$$

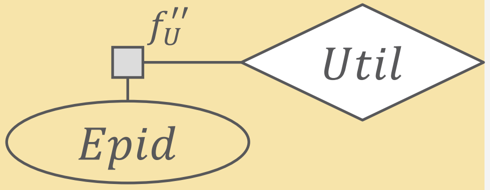
Erwarteter Nutzen von $\boldsymbol{d}_1 = \{ban\}$ in $F = \{f_1, f_2, f_3, f_U\}$

| $Epid$ | $\phi^n$ |
|---|---|
| $false$ | 0.72 |
| $true$ | 0.28 |

| $Epid$ | $Util$ |
|---|---|
| $false$ | $-20$ |
| $true$ | 5 |




| $Epid$ | $Util$ |
|---|---|
| $false$ | $0.72 \cdot (-20) = -14.4$ |
| $true$ | $0.28 \cdot 5 = 1.4$ |




| . | $Util$ |
|---|---|
| . | $-14.4 + 1.4 = -13.0$ |

# EU-Anfragen beantworten (mittels VE)

- Gegeben ein Entscheidungsmodell $M = \{m_i\}_{i=1}^{n} \cup \{m_U\}$, Evidenz $\boldsymbol{e}$ und eine Aktionszuweisung $\boldsymbol{d}$
  - Berechne Wahrscheinlichkeit $P(\boldsymbol{U}|\boldsymbol{e},\boldsymbol{d})$ des *Markov Blankets des Nutzenknotens* $U$, $\mathrm{rv}(m_U)$
    - I.e., $\boldsymbol{U} = \mathrm{rv}(m_U) \setminus \mathrm{rv}(\boldsymbol{d}) \setminus \mathrm{rv}(\boldsymbol{e})$ (verbleibende Zufallsvariablen in $m_U$)
    - Mittels VE: Mit $\boldsymbol{U}$ als Anfragevariablen, eliminiere alle Nicht-Anfragevariablen in $M$, i.e., rufe auf: $\mathrm{VE}(M \setminus \{m_U\}, \boldsymbol{U}, \boldsymbol{d} \cup \boldsymbol{e})$
      - Mit $\boldsymbol{d} \cup \boldsymbol{e}$ als Evidenz setzt VE Aktionszuweisung $\boldsymbol{d}$ mittels Absorption (stellt $M_{\boldsymbol{d}}$ her) und absorbiert $\boldsymbol{e}$
  - Berechne den erwarteten Nutzen durch Eliminierung von $\boldsymbol{U}$:

$$eu(\boldsymbol{e},\boldsymbol{d}) = \sum_{\boldsymbol{u} \in \mathrm{Val}(\boldsymbol{U})} P(\boldsymbol{u}|\boldsymbol{e},\boldsymbol{d}) \cdot \phi_U(\boldsymbol{u})$$

    - Mittels VE: Eliminiere verbleibende Zufallsvariablen in $M$
      - Ergebnis: Leerer Faktor, der auf einen einzigen Wert abbildet ($U$), welcher den erwarteten Nutzen darstellt

# MEU Problem

- Gegeben ein Entscheidungsmodell $M$ und Evidenz $\boldsymbol{e}$ über Zufallsvariablen $\boldsymbol{R} = \boldsymbol{V} \cup \boldsymbol{D}$
- Problem des maximalen erwarteten Nutzens (*maximum expected utility*, MEU, *problem*)
  - Finde die Aktionszuweisung, welche den höchsten erwarteten Nutzen in $M$ hat
  - Formal:

$$\mathrm{meu}(M|\boldsymbol{e}) = \left(\boldsymbol{d}^*, eu(\boldsymbol{e}, \boldsymbol{d}^*)\right)$$

$$\boldsymbol{d}^* = \underset{\boldsymbol{d} \in \mathrm{Val}(\boldsymbol{D})}{\arg\max}\, eu(\boldsymbol{e}, \boldsymbol{d})$$

*Additive* Semantik mit innerer Summe und äußerem Max:

$$\boldsymbol{d}^* = \underset{\boldsymbol{d} \in \mathrm{Val}(\boldsymbol{D})}{\arg\max} \sum_{\boldsymbol{u} \in \mathrm{Val}(\mathrm{rv}(m_U))} P(\boldsymbol{u}|\boldsymbol{e}, \boldsymbol{d}) \cdot \phi_U(\boldsymbol{u})$$

Addiere erwarteten Nutzen, wähle maximum

→ Max-Summen-Algorithmus (*max-sum algorithm*)

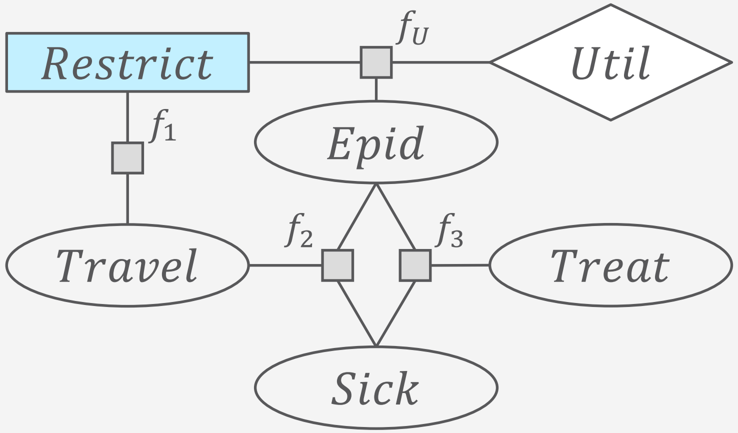
  - Alternative Schreibweise:

$$\mathrm{meu}(M|\boldsymbol{e}) = \left(\underset{\boldsymbol{d} \in \mathrm{Val}(\boldsymbol{D})}{\mathrm{argmax}}\, eu(\boldsymbol{e}, \boldsymbol{d}), \max_{\boldsymbol{d} \in \mathrm{Val}(\boldsymbol{D})} eu(\boldsymbol{e}, \boldsymbol{d})\right)$$

- Für eine exakte Lösung von $\mathrm{meu}(M|\boldsymbol{e})$ muss ein Algorithmus durch alle $\boldsymbol{d} \in \mathrm{Val}(\boldsymbol{D})$
  - Größe von $\mathrm{Val}(\boldsymbol{D})$ exponentiell in $|\boldsymbol{D}|$

# MEU Problem: Beispiel

- Probleminstanz mit $F = \{f_1, f_2, f_3, f_U\}$, $\boldsymbol{e} = \emptyset$ :
$$\mathrm{meu}(G) = \left(\boldsymbol{d}^*, eu(\boldsymbol{d}^*)\right) \quad \boldsymbol{d}^* = \underset{\boldsymbol{d} \in \{\boldsymbol{d}_1, \boldsymbol{d}_2\}}{\mathrm{argmax}}\, eu(\boldsymbol{d})$$
  - $\boldsymbol{d}_1 = \{ban\}, \boldsymbol{d}_2 = \{free\}$
  - Erwarteter Nutzen von $\boldsymbol{d}_1 = \{ban\}$: $eu(\boldsymbol{d}_1) = -13.0$ (auf vorherigen Folien berechnet)
  - Erwarteter Nutzen von $\boldsymbol{d}_2 = \{free\}$: $eu(\boldsymbol{d}_2) = 4.46$
- Lösung
  - $\boldsymbol{d}^* = \underset{\boldsymbol{d} \in \{\boldsymbol{d}_1, \boldsymbol{d}_2\}}{\mathrm{argmax}}\, eu(\boldsymbol{d}) = \boldsymbol{d}_2$
  - $\mathrm{meu}(G) = \left(\boldsymbol{d}_2, eu(\boldsymbol{d}_2)\right) = (\boldsymbol{d}_2, 4.46)$
  - Entscheidung, die zum maximalen erwarteten Nutzen im Modell führt: $Restrict = free$



# VE für MEU Probleme

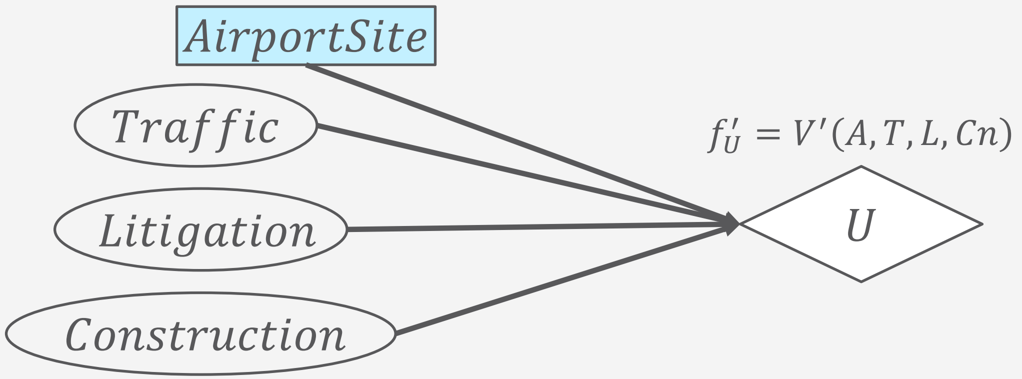
**MEU–VE**$(M = \{m_i\}_{i=1}^{n} \cup \{m_U\}, \boldsymbol{e})$
  Absorbiere $\boldsymbol{e}$ in $M$
  $\boldsymbol{d}^* \leftarrow \emptyset$
  $eu_{max} \leftarrow -\infty$
  **for** jede Aktionszuweisung $\boldsymbol{d} \in \mathrm{Val}(\boldsymbol{D})$ **do**
    $m \leftarrow \mathrm{VE}(M \setminus \{m_U\}, \mathrm{rv}(m_U), \boldsymbol{d})$ ▹ $m$ normalisiert
    $eu \leftarrow \mathrm{VE}(\{m_U, m\}, \emptyset, \emptyset)$ ▹ ohne Normalisierung
    **if** $eu > eu_{max}$ **then**
      $\boldsymbol{d}^* \leftarrow \boldsymbol{d}$
      $eu_{max} \leftarrow eu$
  **return** $\boldsymbol{a}^*$

MEU-VE

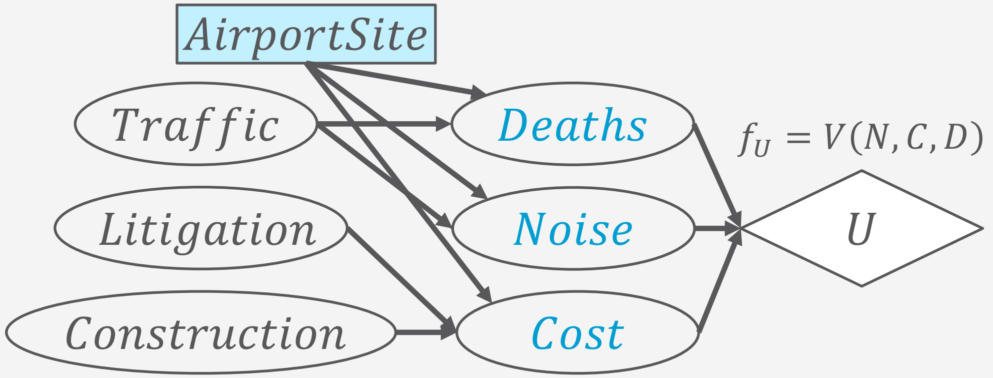
- Anpassung möglich um alle Zuweisungen zu speichern, die innerhalb einer $\varepsilon$-Umgebung liegen

# Ein Wort (okay, mehrere) zu Zufallsvariablen

- Menge von Zufallsvariablen $\boldsymbol{V}$ zur Modellierung der Umgebung inklusive Zufallsvariablen, die *Attribute darstellen, die Präferenzen kodieren*
  - Repräsentieren Ergebniszustände nach Entscheidung
    - Beispiel BN: $Deaths, Noise, Cost$
  - Werden nie mit Evidenz belegt sein
    - Könnte man im Vorhinein eliminieren (keine Wahrscheinlichkeitsanfragen dann möglich) → *vereinfachtes Modell*
  - Ermöglichen mehr Flexibilität in der Modellierung
    - Beispiel:
      - CPD von $Noise$ bei Änderung der Lärmpegel anpassen
      - Gewichtungsänderung der Lärmemissionen in $V$ anpassen

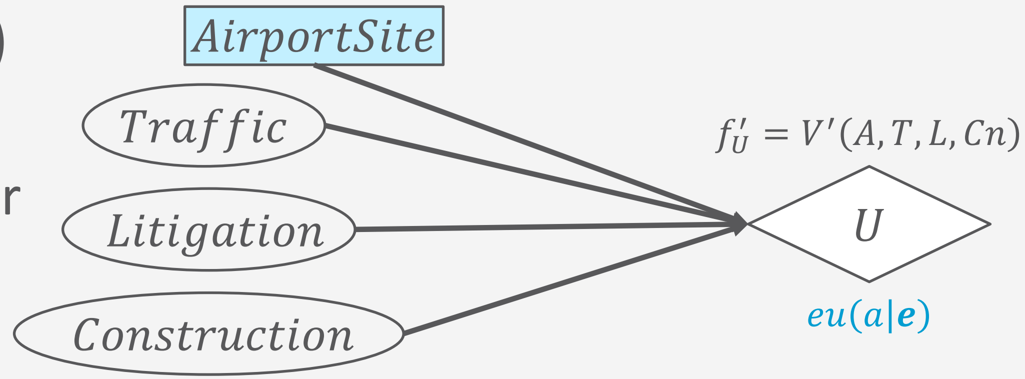


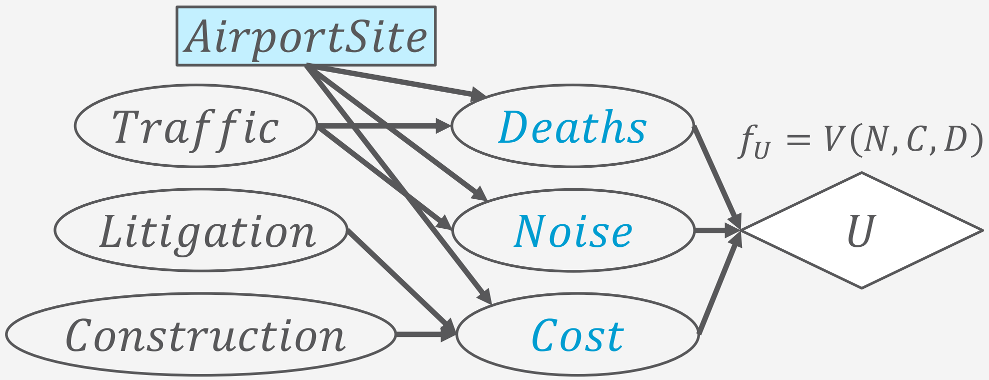

$$V'(A,T,L,Cn) = \sum_{n,c,d \in \mathrm{Val}(N,C,D)} P(d|A,T)P(n|A,T)P(cn|A,L,Cn)V(n,c,d)$$



# Ein Wort (okay, mehrere) zu Zufallsvariablen

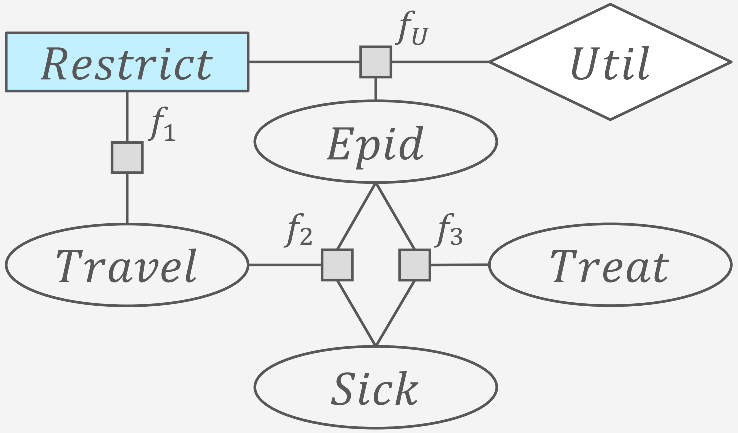
- ACHTUNG: Definition von Nutzenknoten in AIMA(de)
  - Stellen Nutzenfunktion des Agenten dar
  - Im vereinfachten Modell: Stellen erwarteten Nutzen dar
    - Nutzenknoten ist mit einer Aktion-Nutzen-Funktion verbunden, definiert nach $eu(\boldsymbol{e}, \boldsymbol{a})$
- Um im Stile der Knoten in BNs zu bleiben, weichen wir von dieser Definition ab
  - Nutzenknoten: Knoten, die anstatt eine CPD eine Nutzenfunktion verbunden haben und denen der (erwartete) Nutzen zugewiesen wird
  - $eu(\boldsymbol{e}, \boldsymbol{a})$ dann eine Anfrage ans Modell



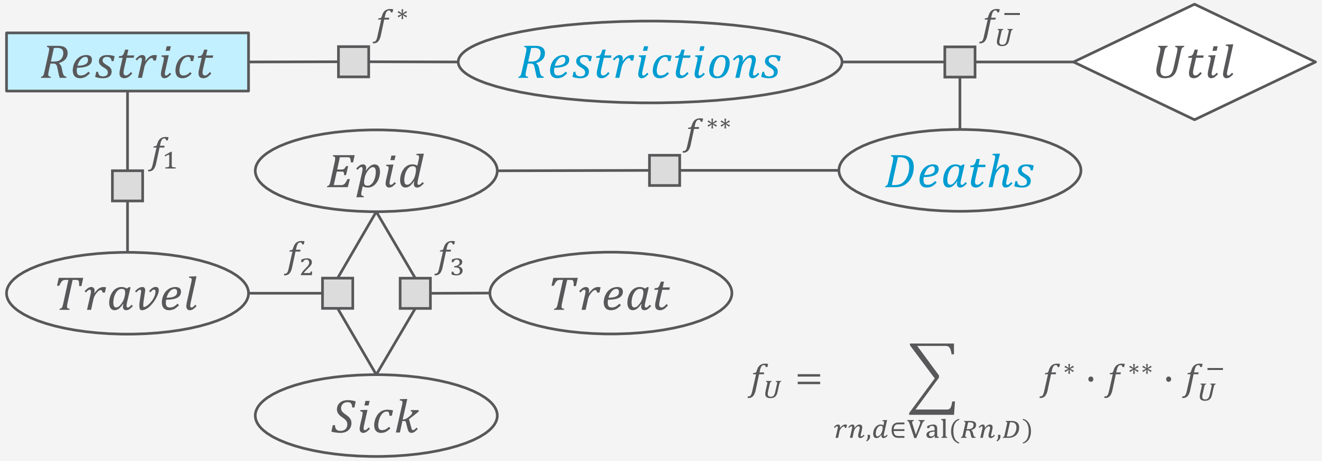



# Ein Wort (okay, mehrere) zu Zufallsvariablen

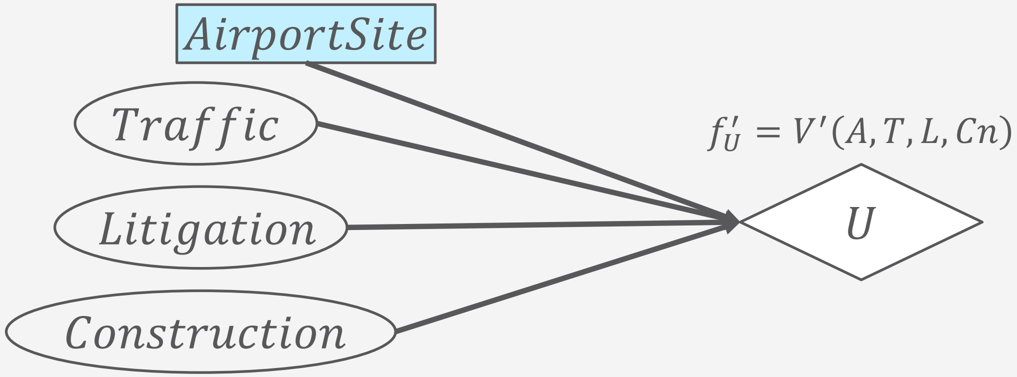
- Beispiel Faktormodell
  - Ohne explizite Zufallsvariablen für Ergebniszustände (vereinfachtes Modell)

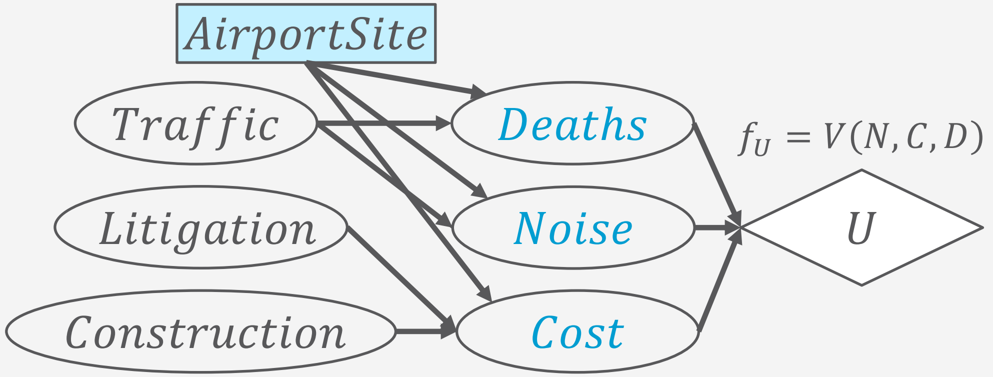



  - Erweitertes Modell:



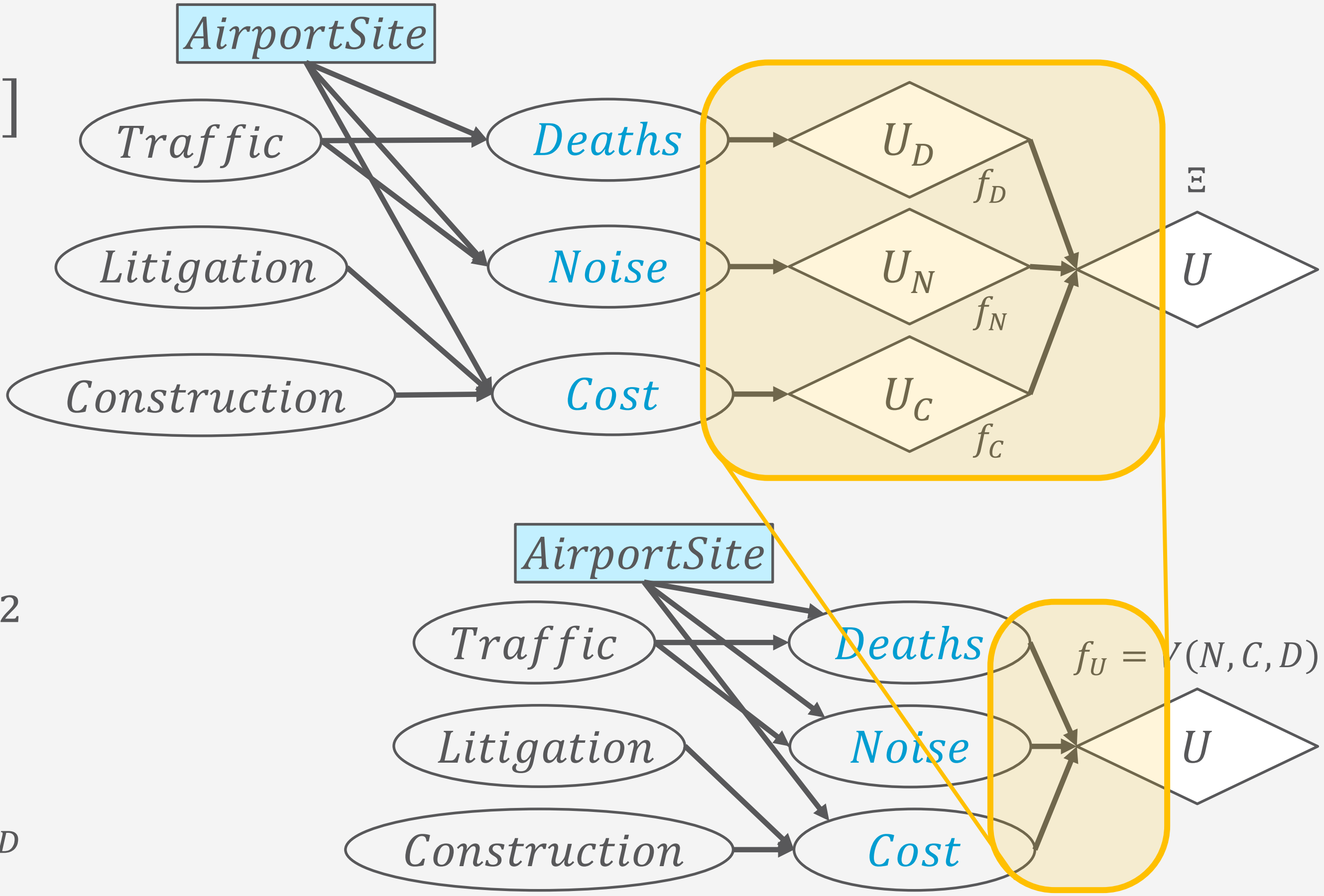

$$f_U = \sum_{rn,d \in \mathrm{Val}(Rn,D)} f^* \cdot f^{**} \cdot f_U^-$$



# Nutzenfaktoren und Mengen von Nutzenvariablen

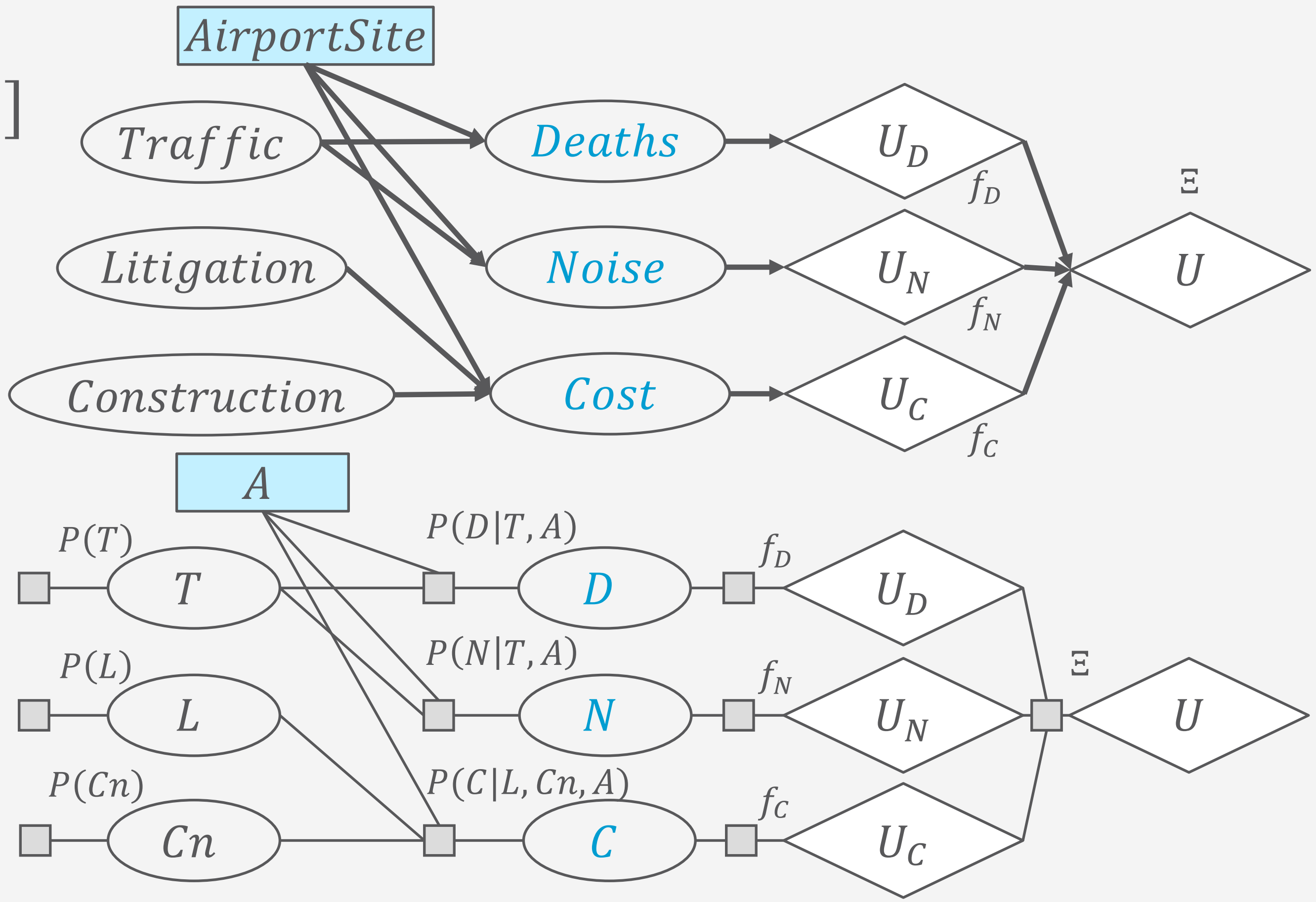
- Nutzenfunktion möglicherweise
  $$U(r_1, \ldots, r_m) = F[f_1(r_1), \ldots, f_m(r_m)]$$
  - Einzelne Funktionen $f_u(r_u)$
  - Kombinationsfunktion $F = \Xi$ (*)
- Im PGM: Nutzenvariablen $U_u$ für $f_u(r_u)$
  - Ein $U$ für Gesamtnutzen
- Beispiel BN mit Nutzenfunktion
  $$V(Noise, Cost, Deaths)$$
  $$= -Noise \cdot 10^4 - Cost - Deaths \cdot 10^{12}$$
  - $f_N = V_N(Noise) = -Noise \cdot 10^4$ auf $U_N$
  - $f_C = V_C(Cost) = -Cost$ auf $U_C$
  - $f_D = V_D(Deaths) = -Deaths \cdot 10^{12}$ auf $U_D$
  - Kombinationsfunktion $\Xi \triangleq \Sigma$ (Addition)




(*) Um Verwechslung mit dem Symbol $F$ für Faktormodelle zu vermeiden, wechseln wir zu $\Xi$ als Symbol für die Kombinationsfunktion

# Nutzenfaktoren und Mengen von Nutzenvariablen

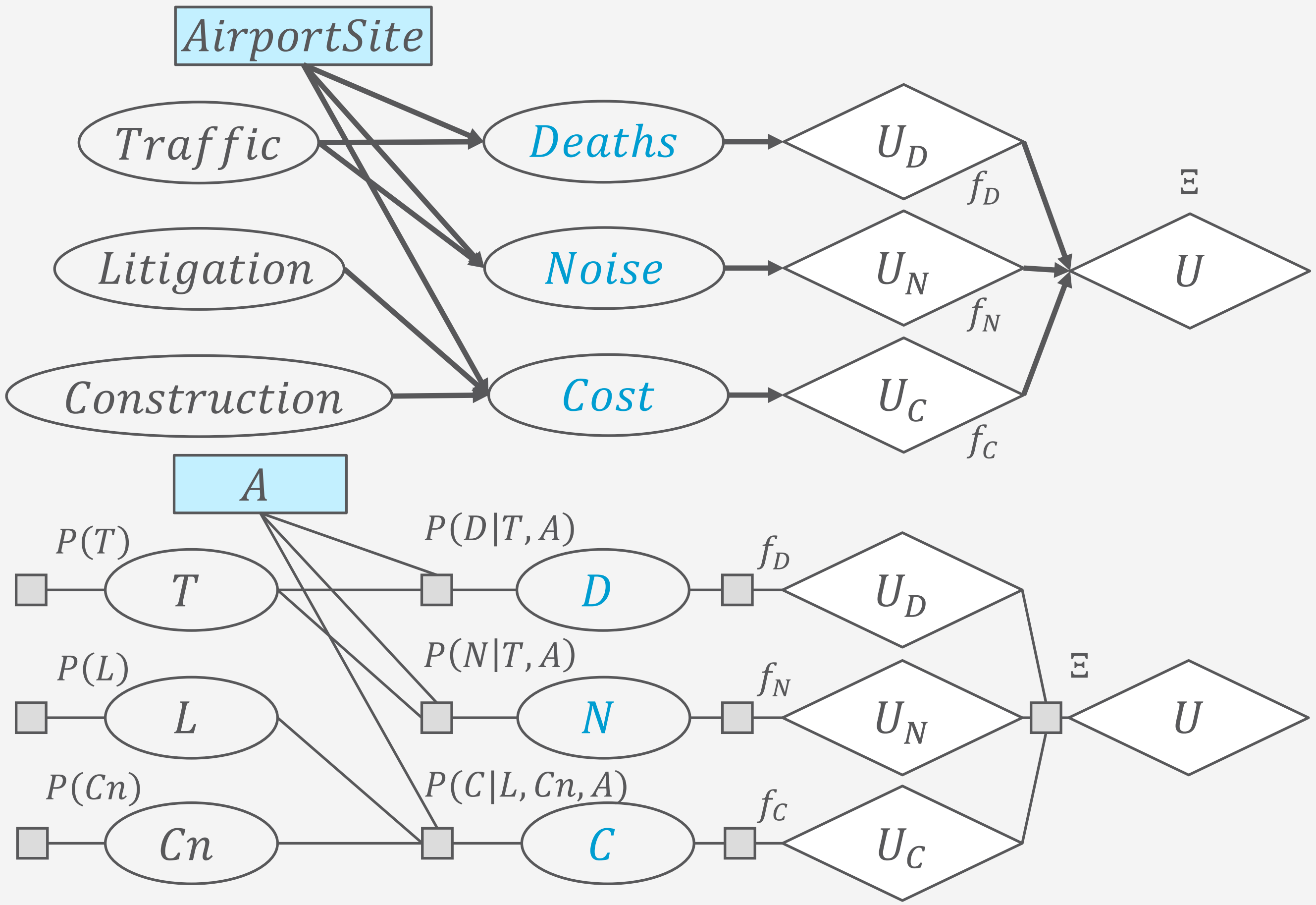
- Nutzenfunktion möglicherweise
  $$U(r_1, \dots, r_m) = F[f_1(r_1), \dots, f_m(r_m)]$$
  - Einzelne Funktionen $f_u(r_u)$
  - Kombinationsfunktion $F = \Xi$ (*)
- Im PGM: Nutzenvariablen $U_u$ für $f_u(r_u)$
  - Ein $U$ für Gesamtnutzen
  - Im Faktorgraph Faktorknoten für die einzelnen $f_u(r_u)$
    - Ursprüngliches $f_U$ dann die Kombinationsfunktion $\Xi$, i.e., $f_u = \Xi$
  - Menge von $f_u(r_u)$ über $\Xi$ immer in ein $f_U$ überführbar




(*) Um Verwechslung mit dem Symbol $F$ für Faktormodelle zu vermeiden, wechseln wir zu $\Xi$ als Symbol für die Kombinationsfunktion

# Nutzenfaktoren und Mengen von Nutzenvariablen

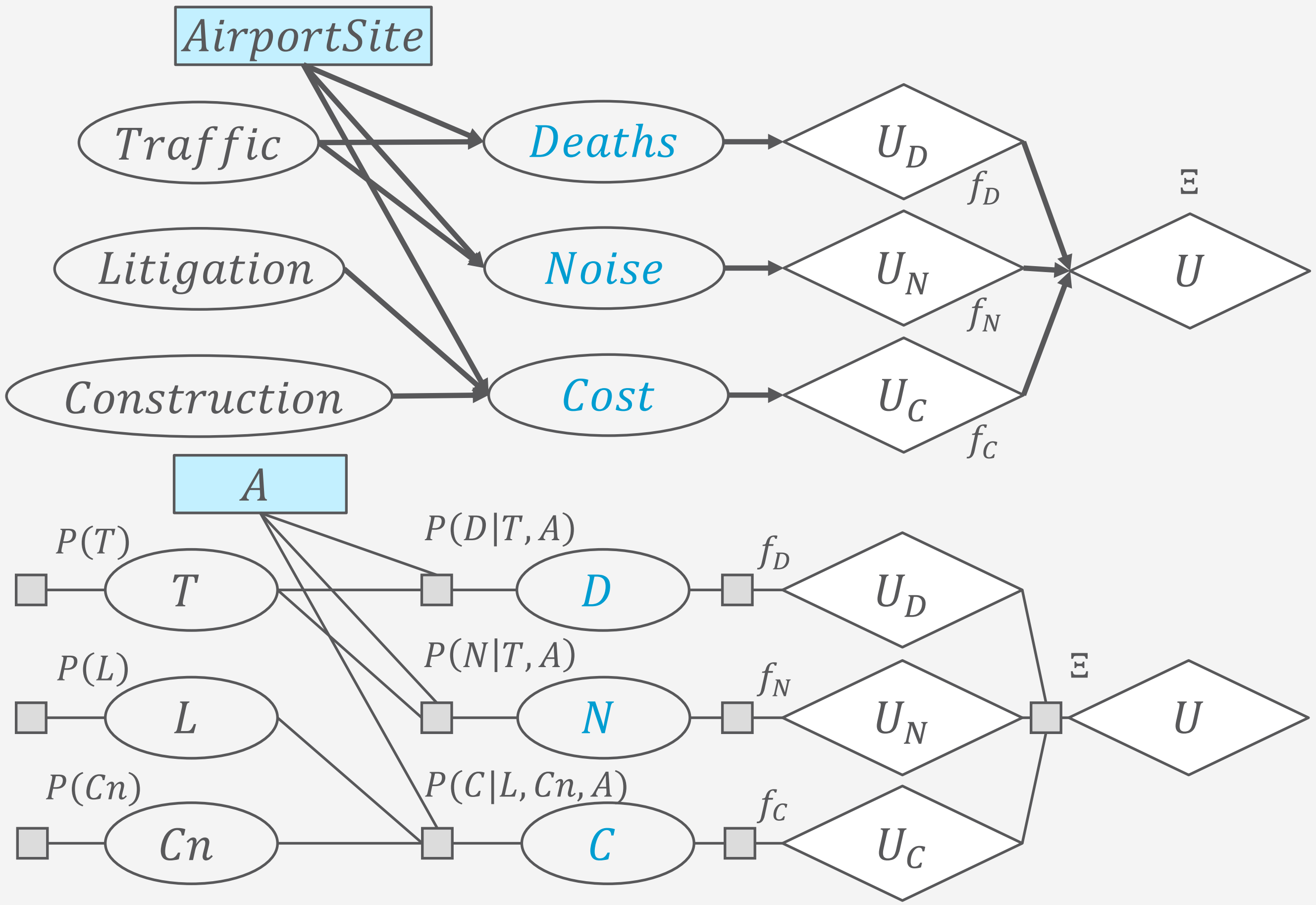
- Entscheidungsmodell mit expliziter Kombinationsfunktion $\Xi$
$$M = \{m_i\}_{i=1}^{n} \cup \{f_u\}_{u=1}^{m} \cup \{\Xi\}$$
  - $f_u = \phi_u(R_u)$ Nutzenfaktoren, bilden auf Nutzenvariablen $U_u$ ab, die mittels $\Xi$ zu Gesamtnutzen $U$ kombiniert werden
    - Setzt voraus, dass Nutzenfunktion durch $\Xi[f_1(R_1), \ldots, f_m(R_m)]$ repräsentiert werden kann
    - Verbindung zu vorheriger Definition:
      - $f_U = \Xi[f_1(R_1), \ldots, f_m(R_m)]$ direkt auf $U$ abgebildet



# Nutzenfaktoren und Mengen von Nutzenvariablen

- Entscheidungsmodell mit expliziter Kombinationsfunktion $\Xi$
$$M = \{m_i\}_{i=1}^{n} \cup \{f_u\}_{u=1}^{m} \cup \{\Xi\}$$
  - EU-Anfrage: Wahrscheinlichkeitsanfrage über $\boldsymbol{U} = \mathrm{rv}(\{f_u\}_{u=1}^{m}) \setminus \mathrm{rv}(\boldsymbol{d}) \setminus \mathrm{rv}(\boldsymbol{e})$
    - Markov Blanket aller $U_u$ ohne $U$ : $\boldsymbol{U} = \bigcup_{u=1}^{m} \mathrm{MB}(U_u) \setminus \{U\}$
    - Äquivalent zum vorherigen Ausdruck, da $f_U$ über $\mathrm{rv}(\{f_u\}_{u=1}^{m})$ ging

  $$eu(\boldsymbol{e}, \boldsymbol{d}) = \sum_{\boldsymbol{u} \in \mathrm{Val}(\boldsymbol{U})} P(\boldsymbol{u}|\boldsymbol{e}, \boldsymbol{d}) \cdot \Xi[\underbrace{f_1(r_1), \ldots, f_m(r_m)}_{\substack{r_u = \pi_{R_u}(\boldsymbol{u}) \\ u = 1, \ldots, m}}]$$



# Nutzenfaktoren und Mengen von Nutzenvariablen

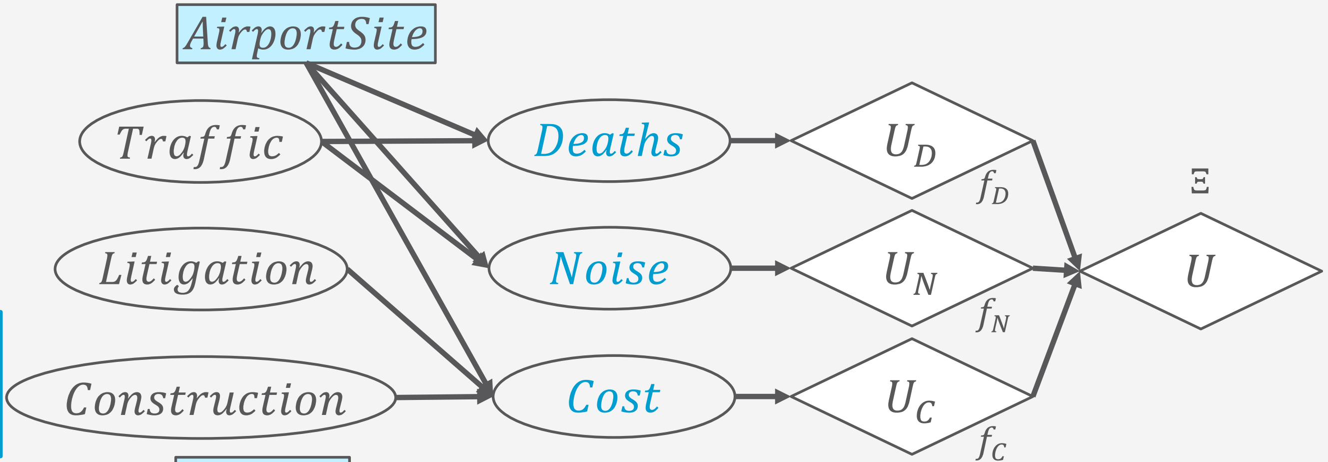
- Entscheidungsmodell mit expliziter Kombinationsfunktion $\Xi$

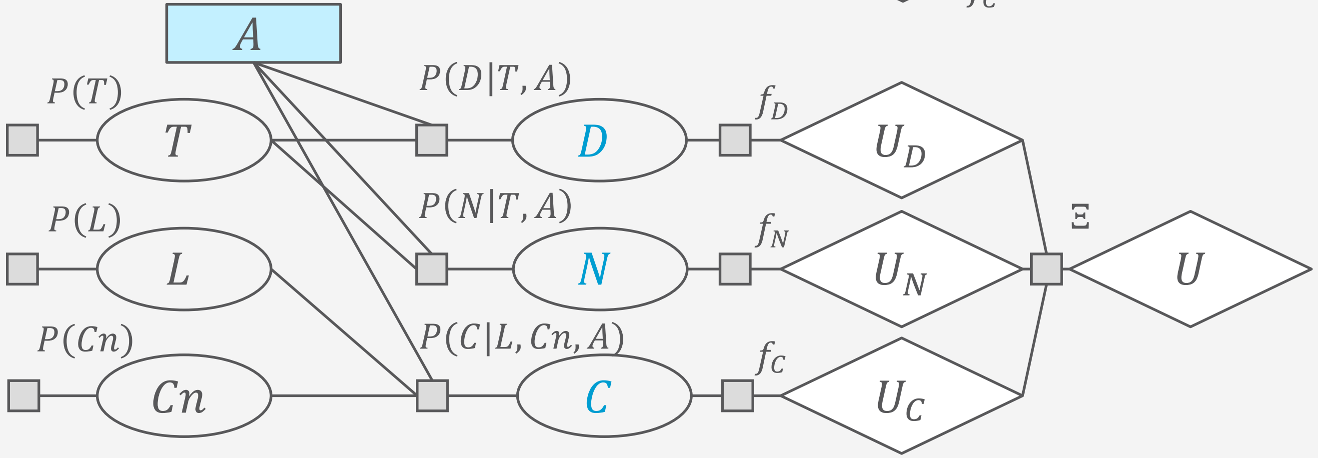
$$M = \{m_i\}_{i=1}^{n} \cup \{f_u\}_{u=1}^{m} \cup \{\Xi\}$$

  - Änderung in **MEU−VE**:

$m \leftarrow \mathrm{VE}(M \setminus \{f_u\}_{u=1}^{m}, \mathrm{rv}(\{f_u\}_{u=1}^{m}), \emptyset)$ ▹ $m$ normalisiert
$eu \leftarrow \mathrm{VE}(\{\{f_u\}_{u=1}^{m}, m\}, \emptyset, \emptyset)$

    - Bei additivem $\Xi \triangleq \Sigma$
    - Sonst: $\Xi$ implementieren



# Additive Kombinationsfunktion

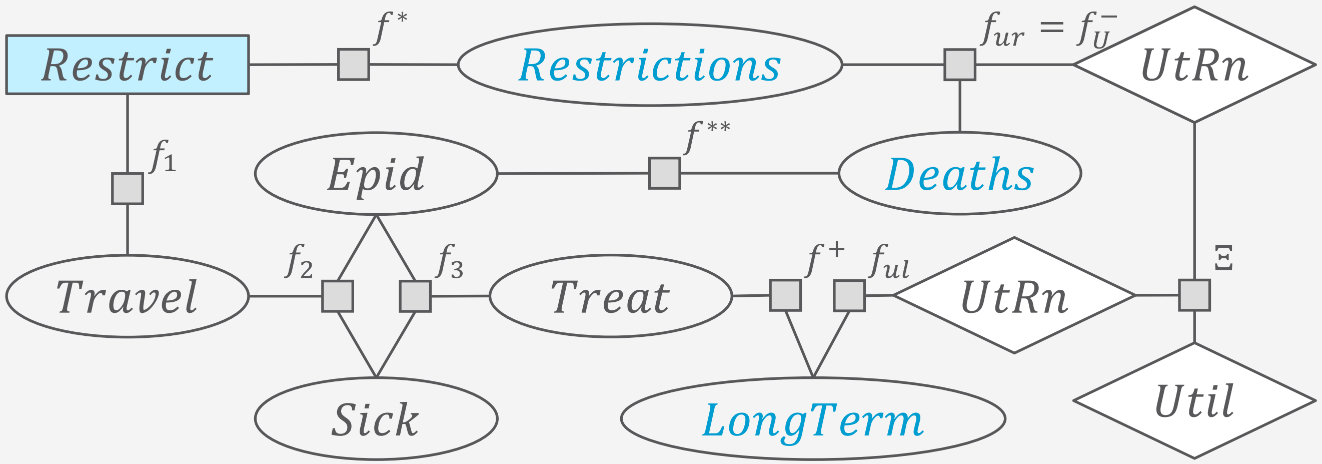
- Additive Kombinationsfunktion $\Xi$ setzt *präferentielle Unabhängigkeit* zwischen den Attributen $R_u$ voraus
  - Falls nur zwischen Untermengen (größer 1) gegeben, $f_u$ über Menge von $R_{u_1}, \dots, R_{u_k}$ möglich mit anschließender Kombination über $\Xi$
    - $f_u = \phi_u(R_{u_1}, \dots, R_{u_k})$
  - Beispiel:
    - $Restrictions, Death$ nicht PI
    - $f_{ur} = \phi_{ur}(Restrictions, Death)$
    - $f_{ul} = \phi_{ul}(LongTerm)$
    - $\Xi \triangleq \Sigma$

Nicht zerlegbar in $\phi(Rn) + \phi(D)$
→ nicht PI

| $Rn$ | $D$ | $UtD$ |
|---|---|---|
| $false$ | $false$ | $10$ |
| $false$ | $true$ | $-10$ |
| $true$ | $false$ | $-20$ |
| $true$ | $true$ | $5$ |

| $LT$ | $UtRn$ |
|---|---|
| $false$ | $0$ |
| $true$ | $10$ |

| $Rn$ | $D$ | $LT$ | $Util$ |
|---|---|---|---|
| $false$ | $false$ | $false$ | $10 + 0 = 10$ |
| $false$ | $false$ | $true$ | $10 + 10 = 20$ |
| $false$ | $true$ | $false$ | $-10 + 0 = -10$ |
| $false$ | $true$ | $true$ | $-10 + 10 = 0$ |
| $true$ | $false$ | $false$ | $-20 + 0 = -20$ |
| $true$ | $false$ | $true$ | $-20 + 10 = -10$ |
| $true$ | $true$ | $false$ | $5 + 0 = 50$ |
| $true$ | $true$ | $true$ | $5 + 10 = 15$ |



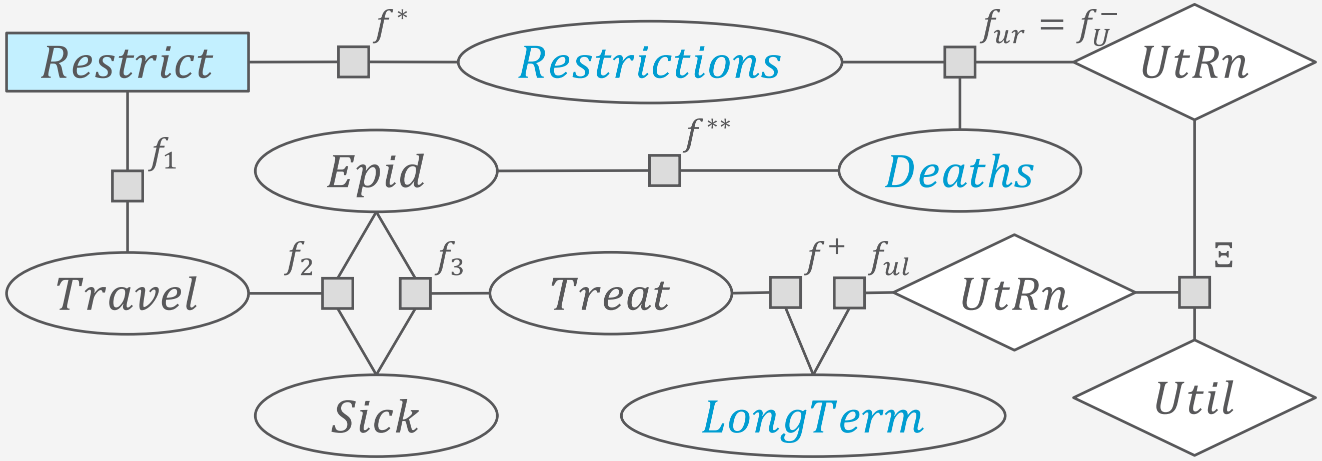
# Additive Kombinationsfunktion

- Entscheidungsmodell $M = \{m_i\}_{i=1}^n \cup \{f_u\}_{u=1}^m$ mit additiver Kombinationsfunktion $\Xi$
  - $f_u = \phi_u(R_{u_1}, \dots, R_{u_k})$ Nutzenfaktoren, bilden auf Nutzenvariablen $U_u$ ab, die mittels $\Xi$ zu Gesamtnutzen $U$ addiert werden
  - EU-Anfrage
    - $\boldsymbol{U} = \mathrm{rv}(\{f_u\}_{u=1}^m) \setminus \mathrm{rv}(\boldsymbol{d}) \setminus \mathrm{rv}(\boldsymbol{e})$

$$eu(\boldsymbol{e}, \boldsymbol{d}) = \sum_{\boldsymbol{u} \in \mathrm{Val}(\boldsymbol{U})} P(\boldsymbol{u}|\boldsymbol{e}, \boldsymbol{d}) \cdot F[f_1(\boldsymbol{r}_1), \dots, f_m(\boldsymbol{r}_m)]$$

$$= \sum_{\boldsymbol{u} \in \mathrm{Val}(\boldsymbol{U})} \boxed{P(\boldsymbol{u}|\boldsymbol{e}, \boldsymbol{d})} \cdot \sum_{u=1}^{m} \phi_u(\underbrace{\boldsymbol{r}_{u_1}, \dots, \boldsymbol{r}_{u_k}}_{\boldsymbol{r}_{u_i} = \pi_{\boldsymbol{R}_{u_i}}(\boldsymbol{u}),\ i = 1, \dots, k})$$

Kombiniert alle Zufallsvariablen aus Nutzenfaktoren



# Additive Kombinationsfunktion: Vereinfachung

- EU-Anfrage
  - $\boldsymbol{U} = \mathrm{rv}(\{f_u\}_{u=1}^{m}) \setminus \mathrm{rv}(\boldsymbol{d}) \setminus \mathrm{rv}(\boldsymbol{e})$

$$eu(\boldsymbol{e},\boldsymbol{d}) = \sum_{\boldsymbol{u}\in\mathrm{Val}(\boldsymbol{U})} P(\boldsymbol{u}|\boldsymbol{e},\boldsymbol{d}) \cdot F[f_1(\boldsymbol{r}_1), \dots, f_m(\boldsymbol{r}_m)]$$

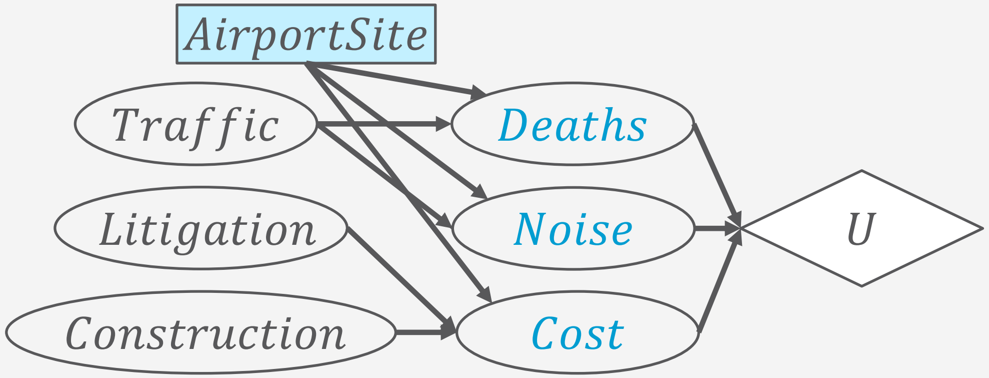
$$= \sum_{\boldsymbol{u}\in\mathrm{Val}(\boldsymbol{U})} P(\boldsymbol{u}|\boldsymbol{e},\boldsymbol{d}) \cdot \sum_{u=1}^{m} \phi_u(\boldsymbol{r}_u)$$

$$= \sum_{\boldsymbol{u}\in\mathrm{Val}(\boldsymbol{U})} \sum_{u=1}^{m} \underbrace{P(\boldsymbol{r}_u|\boldsymbol{e},\boldsymbol{d})}_{\text{Wahrscheinlichkeitsanfrage an rv}(f_u)\text{ je Nutzenfaktor}} \phi_u(\boldsymbol{r}_u)$$

Vereinfachung liefert korrektes Ergebnis, wenn zwischen den $\mathrm{rv}(f_u)$ der unterschiedlichen $f_u$ tatsächlich Unabhängigkeit herrscht

- Ähnlich zu Boyen-Koller-Algorithmus bei sequentieller Inferenz
- Präferentielle und stochastische Unabhängigkeit folgen nicht auseinander!

$$P(D, N, C|a) \neq P(D|a)P(N|a)P(C|a)$$



# MEU-JT

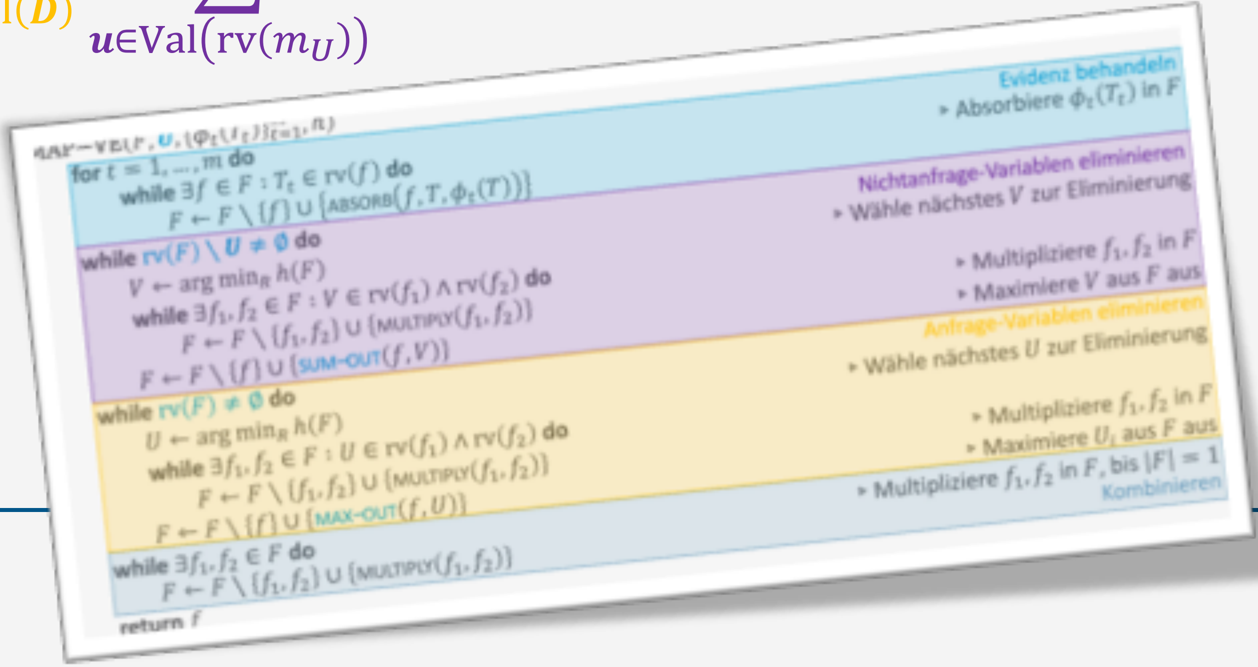
- Beantworte EU Anfragen, Wahrscheinlichkeits- und Zustandsanfragen
  - Nehme *additive* Kombinierungsfunktion an
  - Baue Jtree für Entscheidungsmodell
    - Über einen Hilfsfaktor $f$ über $\boldsymbol{U} = \mathrm{rv}(\{f_u\}_{u=1}^m)$ dafür sorgen, dass alle $\boldsymbol{U}$ für Berechnung von $P(\boldsymbol{u}|\boldsymbol{e}, \boldsymbol{d})$ in einem Cluster
    - Ohne $f$, Subgraph über $\boldsymbol{U}$ für jede $P(\boldsymbol{u}|\boldsymbol{e}, \boldsymbol{d})$
      - Cluster könnten so kleiner sein → besser für andere Anfragen
    - Mit $f$ ist $\boldsymbol{U}$ in einem Cluster enthalten → einfacher für $P(\boldsymbol{u}|\boldsymbol{e}, \boldsymbol{d})$
      - Cluster können dann aber größer sein → schlechter für andere (kleine) Anfragen
    - Bei Verwendung der Vereinfachung: kein $f$ nötig
  - Berechnung EU-Anfrage wie in MEU-VE als zwei Teile ($P(\boldsymbol{u}|\boldsymbol{e}, \boldsymbol{d})$, damit dann $eu(\boldsymbol{e}, \boldsymbol{d})$)
    - Bei Vereinfachung: Menge von Teil-EU-Anfragen berechnen, Ergebnisse *addieren*

# MEU Problem als MAP Anfrage

- Gegeben ein Entscheidungsmodell $M$ und Evidenz $\boldsymbol{e}$ über Zufallsvariablen $\boldsymbol{R} = \boldsymbol{V} \cup \boldsymbol{D}$
- MEU Problem interpretiert als MAP Anfrage
  - Finde die Aktionszuweisung, welche den höchsten erwarteten Nutzen in $M$ hat
  - Formal: $\mathrm{meu}(M|\boldsymbol{e}) = \left(\boldsymbol{d}^*, eu(\boldsymbol{e}, \boldsymbol{d}^*)\right)$ mit

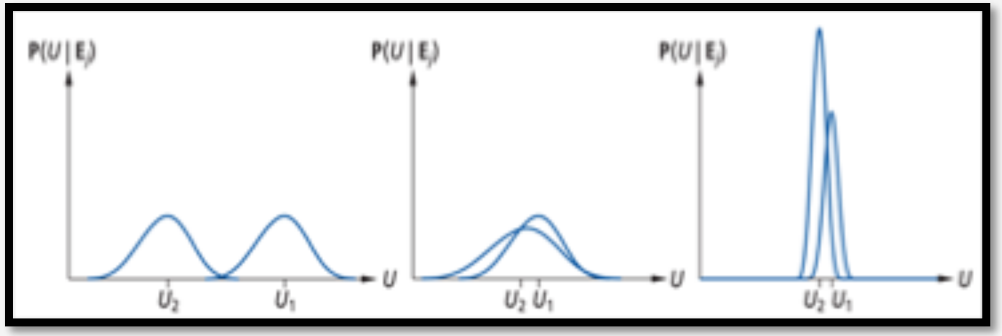
$$\boldsymbol{d}^* = \underset{\boldsymbol{d}\in \mathrm{Val}(\boldsymbol{D})}{\arg\max}\, eu(\boldsymbol{e}, \boldsymbol{d}) = \underset{\boldsymbol{d}\in \mathrm{Val}(\boldsymbol{D})}{\arg\max} \sum_{\boldsymbol{u}\in \mathrm{Val}(\mathrm{rv}(m_U))} P(\boldsymbol{u}|\boldsymbol{e}, \boldsymbol{a}) \cdot \phi_U(\boldsymbol{u})$$

    - Nicht-Entscheidungs- und Nicht-Evidenzvariablen aussummieren
    - Entscheidungsvariablen ausmaximieren
- Komplexität von MEU folgt hier der von MAP mit den gleichen Einschränkungen bzgl. Baumweite



# Zwischenzusammenfassung

- Entscheidungsmodell
  - Probabilistische graphische Modelle erweitert mit Entscheidungs- und Nutzenvariablen
    - Nutzenfaktoren
    - Zufallsvariablen kann man unterscheiden in normale Zufallsvariablen und die, welche einen Ergebniszustand der Aktionen repräsentieren, die dann den Nutzen bestimmen
      - Vereinfachtes Modell: Zufallsvariablen für Ergebniszustände eliminieren → kleiner, aber weniger ausdrucksstark
  - Neue Inferenzaufgaben: EU Anfragen, MEU Problem
    - Finde eine Menge von Aktionen (Entscheidungen), welche zum maximalen erwarteten Nutzen führen
    - MEU-VE nutzt Aufrufe von VE um EU Anfragen zu beantworten
  - Bei präferentieller Unabhängigkeit: additive Kombinationsfunktion annehmen
    - Approximation über Annahme, dass neben präferentieller Unabhängigkeit auch stochastische Unabhängigkeit zwischen den Attributen (Ergebniszuständen) herrscht




# Informationswerttheorie

Wert perfekter Information

Informationen sammelnder Agent



Abbildung: AIMAde, S. 735

# Entscheidungsfindung in Entscheidungsnetzwerken

- Annahme, dass alle zur Verfügung stehenden Informationen dem Agenten bereitgestellt werden, bevor der Agent eine Entscheidung tätigt
  - Fast nie der Fall
  - Wichtig zu wissen, wann Fragen zu stellen (Informationen zu sammeln)
- Informationswerttheorie
  - Wähle, welche Informationen zu beschaffen sind
  - Annahme, dass der Agent vor der Wahl einer Aktion die Belegung jeder der potentiell beobachtbaren Zufallsvariablen erhalten kann
    - Vereinfachte Version der sequentiellen Entscheidungsfindung
      - Beobachtungsaktionen haben nur Auswirkungen auf den internen Zustand des Agenten, nicht den externen (physischen) Zustand der Umgebung
      - Sequentielle Entscheidungsfindung: Aktionen haben Auswirkungen auf die Umgebung

# Informationswert

- Idee: Berechne Wert davon, jede mögliche Evidenz zu beschaffen
  - Direkt im Entscheidungsnetzwerk möglich
  - Beispiel: Kauf von Rechten zu Ölbohrungen
    - Zwei Blöcke $A$ und $B$, genau eins hat Öl, mit dem Wert $k$
    - Apriori Wahrscheinlichkeit Öl zu haben je $0.5$, gegenseitig ausschließend
    - Momentaner Preis von jedem Block: $^k/_2$
    - Seismologe bietet Ergebnisse einer Untersuchung von $A$ an (Information) → Ordentlicher Preis dafür?
    - Lösung: Berechne erwarteten Informationswert
      - = erwarteter Wert der besten Aktion gegeben die Information
        minus erwarteter Wert der besten Aktion ohne Information
      - Untersuchung sagt „Öl in A" oder „kein Öl in A", Wahrscheinlichkeit je $0.5$ (gegeben)
        $= [0.5 \cdot \text{Wert von „kaufe A" gegeben „Öl in A"}$
        $+ 0.5 \cdot \text{Wert von „kaufe B" gegeben „kein Öl in A"}] - 0$
        $= (0.5 \cdot {}^k/_2) + (0.5 \cdot {}^k/_2) - 0 = {}^k/_2$

# Allgemeine Formel

- Aktuelle Evidenz $\boldsymbol{e}$, aktuelle beste Aktion $d^*$, mögliche Ausgänge der Aktion $s_i$, potentielle neue Evidenzvariable $E_j$

$$MEU(\boldsymbol{e}) = \arg\max_d \sum_i P(s_i|\boldsymbol{e}, d)U(s_i) = d^*$$

- Nehme an, wir wüssten $E_j = e_{jk}$; dann würden wir $d^*_{jk}$ wählen, so dass

$$MEU\left(\boldsymbol{e}, E_j = e_{jk}\right) = \arg\max_d \sum_i P\left(s_i\middle|\boldsymbol{e}, E_j = e_{jk}, d\right)U(s_i) = d^*_{jk}$$

- $E_j$ ist eine Zufallsvariable, deren Belegung momentan nicht bekannt ist
  → erwarteten Gewinn über alle möglichen Werte berechnen:

$$VPI_{\boldsymbol{e}}(E_j) = \left(\sum_k P\left(e_{jk}\middle|\boldsymbol{e}\right) eu\left(d^*_{jk}\middle|\boldsymbol{e}, E_j = e_{jk}\right)\right) - eu(d^*|\boldsymbol{e})$$

  - Wert perfekter Information (*value of perfect information*, VPI)

# Eigenschaften von VPI

- Nicht-negativ – in der Erwartung

$$\forall j, E : VPI_E(E_j) \geq 0$$

- Nicht-additiv – überlege z.B. $E_j$ zweimal zu beschaffen

$$VPI_{\boldsymbol{e}}(E_j, E_k) \neq VPI_E(E_j) + VPI_E(E_k)$$

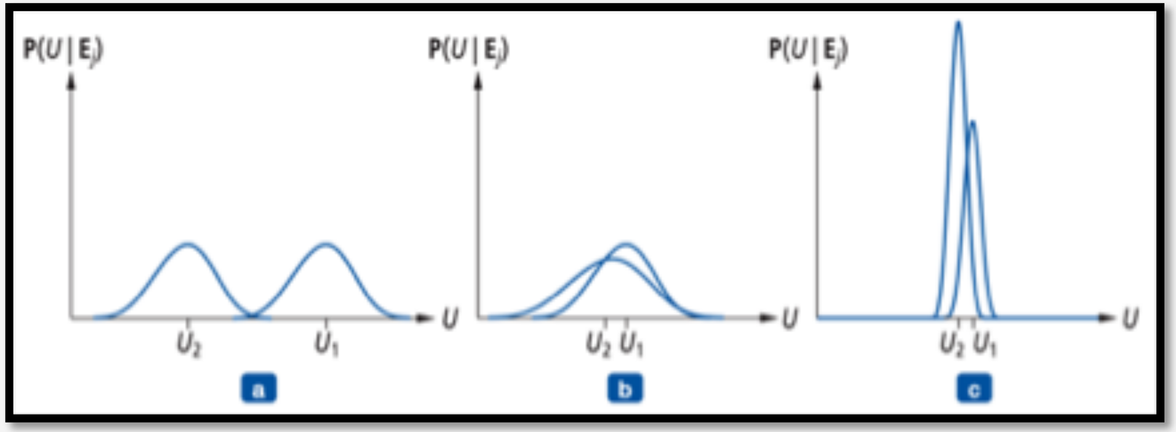
- Reihenfolge-unabhängig

$$VPI_{\boldsymbol{e}}(E_j, E_k) = VPI_{\boldsymbol{e}}(E_j) + VPI_{\boldsymbol{e},E_j}(E_k) = VPI_{\boldsymbol{e}}(E_k) + VPI_{\boldsymbol{e},E_k}(E_j)$$

- Bemerkung: Wenn mehr als ein Evidenzbestandteil gesammelt werden kann, ist die Maximierung von VPI pro Bestandteil nicht immer optimal
  - Kurzsichtiges / gieriges Vorgehen

→ Evidenz sammeln wird zum sequentiellen Entscheidungsproblem

# Qualitatives Verhalten




a. Wahl ist offensichtlich, Information wenig wert
b. Wahl ist nicht offensichtlich, Information viel wert
c. Wahl ist nicht offensichtlich, Information wenig wert

- *Information hat Wert in dem Umfang, wie es wahrscheinlich ist, dass sie eine Änderung des Plans auslöst und der neue Plan signifikant besser als der alte Plan ist*


Abbildung: AIMAde, S. 735

# Informationen sammelnder Agent

```
function INFORMATION-GATHERING-AGENT(percept)
returns: an action
persistent: M, a decision model

   integrate percept into M
   j ← the value that maximises VPI(E_j)/Cost(E_j)
   if VPI(E_j) > Cost(E_j) then
      return Request(E_j)
   else
      return the best action from M
```

- Mit Menschen im Verbund wichtig: $Request(E_j)$ in einer vernünftigen Reihenfolge fragen
- Irrelevante Fragen vermeiden
- Wichtigkeit eines Stücks Information $j$ relativ zu den Kosten $Cost(E_j)$ berücksichtigen

# Zwischenzusammenfassung

- Informationswert
  - Wie viel kostet es ein neues Stück Informationen zu beschaffen?
  - Wird das Stück Information den aktuell besten Plan ändern?
  - Wie viel mehr Nutzen kann man vom neuen besten Plan erwarten?
  - Ist es das wert?

- Informationen sammelnder Agent
  - Frage eine Evidenzbelegung an, wenn der erwartete Nutzen der Aktion / des Plans mit dieser Information die Kosten überwiegt

# Überblick: 7. Entscheidungstheoretische PGMs und Inferenz

A. *Nutzentheorie*
- Präferenzen, Nutzen, Maximum-Expected-Utility Prinzip, Dominanz und Unabhängigkeiten

B. *Entscheidungstheoretische episodische PGMs und Inferenz*
- Syntax: Knoten für Nutzen / Entscheidungen, Einflussdiagram, Entscheidungsmodell; Semantik
- Inferenzaufgaben: Erwarteter Nutzen, Aktionen mit höchstem erwarteten Nutzen
- Entscheidungstheoretische VE-Variante
- Informationswerttheorie

C. ***Entscheidungstheoretische sequentielle PGMs und Inferenz***
- Zeitliche Inferenzaufgaben
- VE-basierter Inferenzalgorithmus
- Automatisiertes Handeln

# Über die Zeit Entscheidungen treffen

Bisher:

- Berechne den erwarteten Nutzen einer Entscheidung und deren Auswirkung auf den (aktuellen) Zustand

Mit Zeit

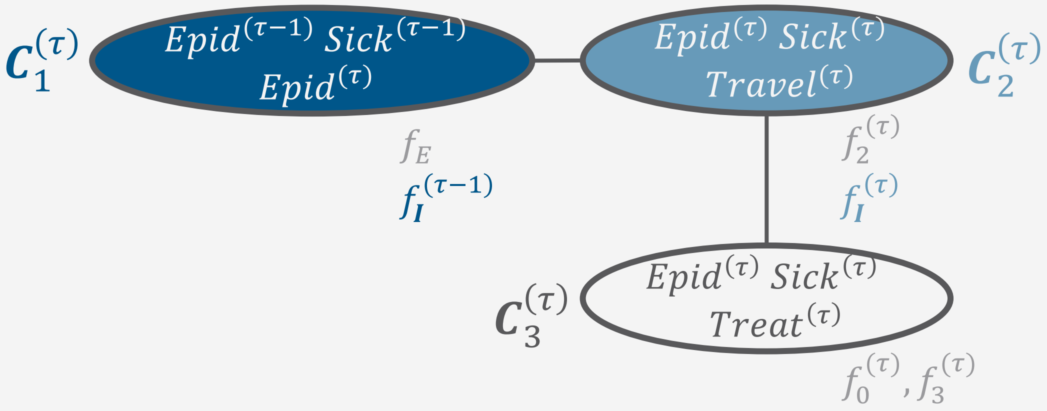
- Entscheidungen müssen jeden Zeitschritt getroffen werden
- Jeder Zeitschritt ergibt einen Nutzen
- Entscheidungen/Aktionen haben einen indirekten Effekt nicht nur auf den aktuellen Zustand / Nutzen, sondern auch auf die Zukunft und damit auf zukünftige Entscheidungen

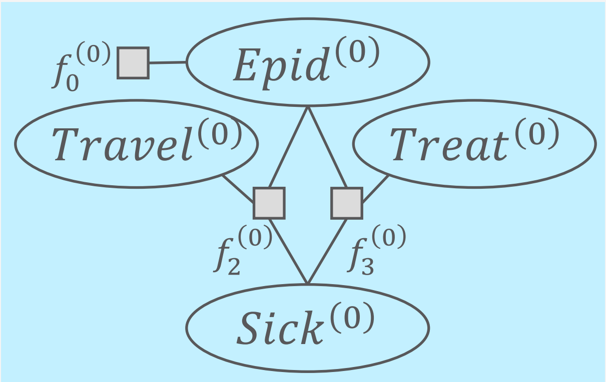
→ Sequenz von zukünftigen Entscheidungen betrachten und ihren Effekt in die Zukunft projizieren

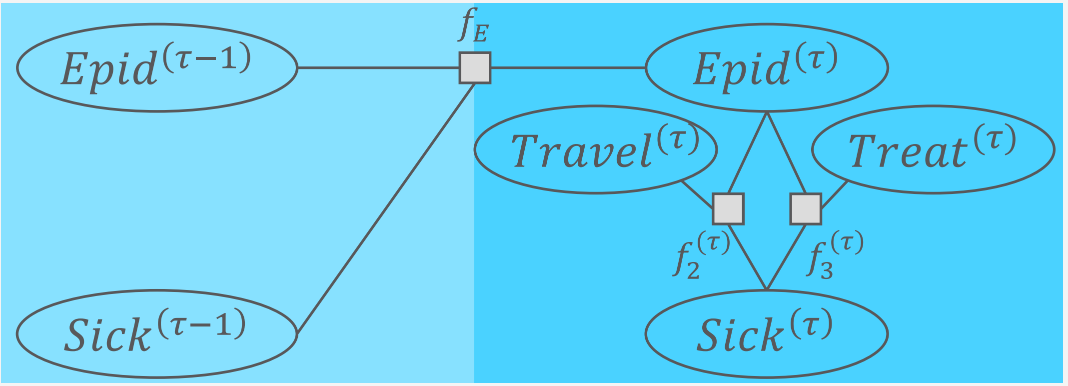
→ Erfordert Berechnung des erwarteten Nutzens über eine Sequenz

# Wiederholung: Sequentielle (dynamische) Modelle

- Dynamische Modelle sind Tupel $(M^0, M^{\rightarrow})$
  - $M^0$ episodisches Modell für $t = 0$
  - $M^{\rightarrow}$ ein 2-Zeitscheiben-Modell über $\tau, \tau - 1$
    - Template-Modell, das für $\tau = t$ instanziiert wird
  - Semantik über Ausrollen für $T$ Zeitschritte
- Anfragebeantwortung:
  - Anfragen: Filterung, Vorhersage, Rückschau
  - Algorithmus: DJT
    - Interfaces
    - Vorwärts- / Rückwärtsnachrichten, um sich in der Zeit zu bewegen



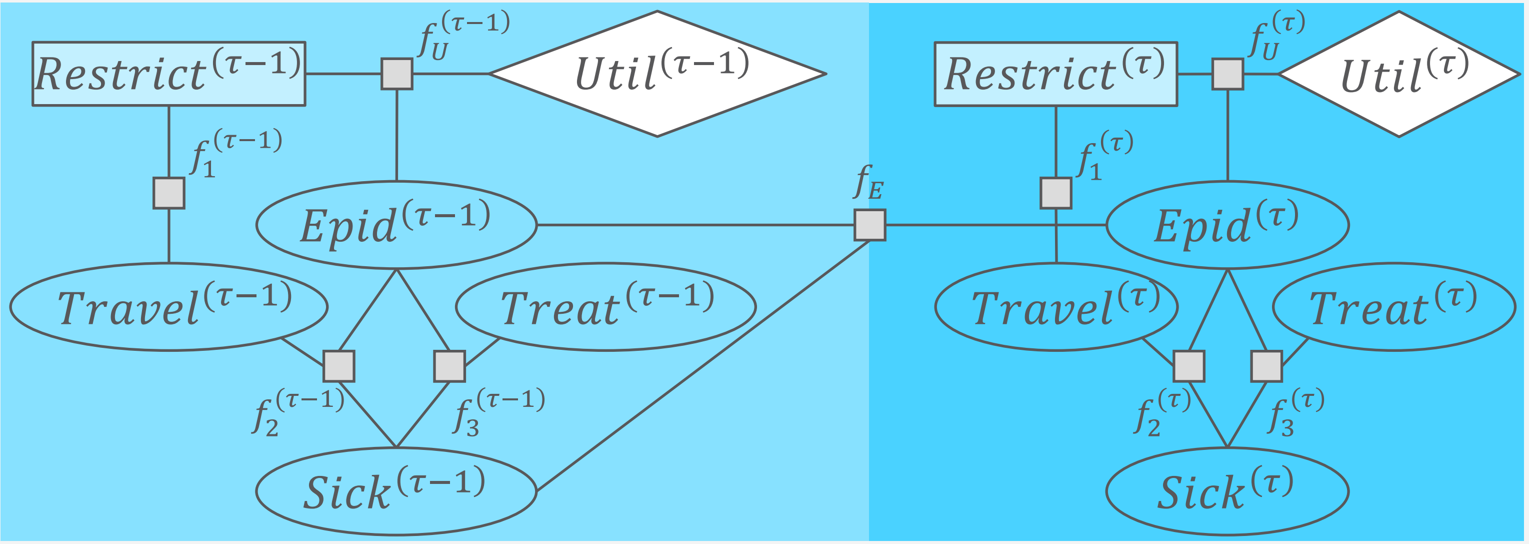



# Über die Zeit Entscheidungen treffen

- Basis: dynamisches Modell $(M^0, M^{\rightarrow})$
  - Inter-Zeitscheiben-Verhalten über die Zeit
  - Intra-Zeitscheiben-Verhalten (für $t = 0$ in $M^0$, sonst in $M^{\rightarrow}$)

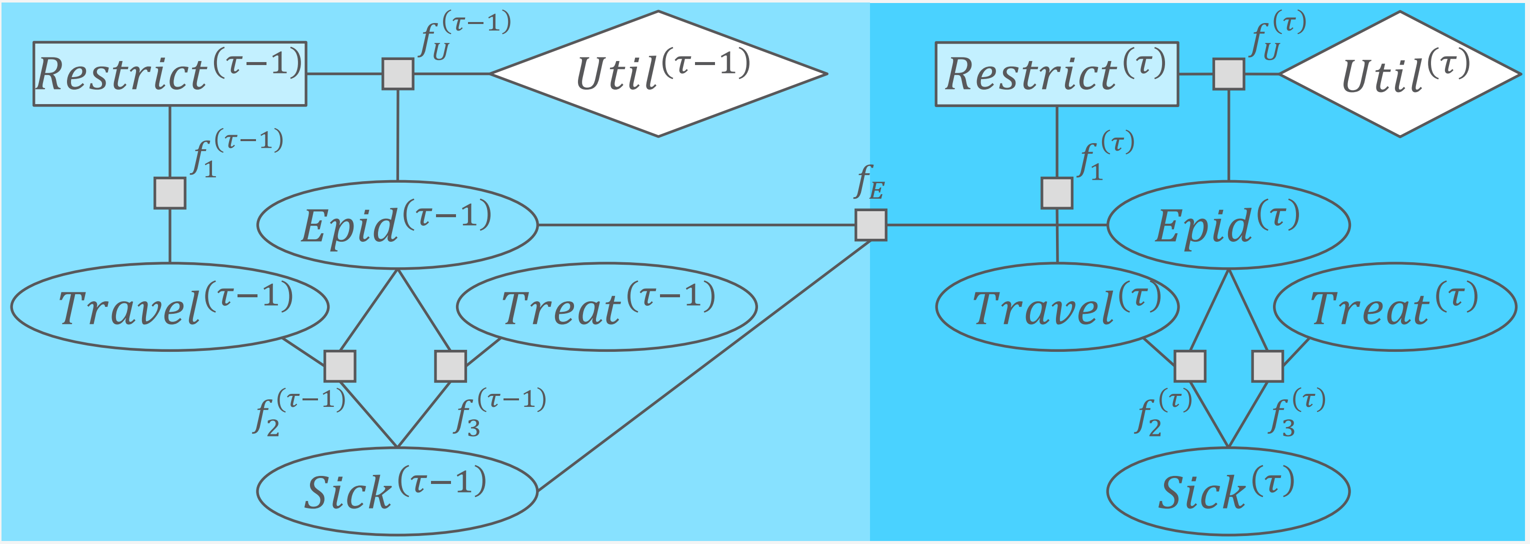
→ Intra-Zeitscheiben-Teil mit Entscheidungs- und Nutzenvariablen erweitern

  - Anstatt klassische episodische Modelle für $M^0, M^{\rightarrow}$ *Entscheidungsmodelle* nutzen
  - Effekte von Aktionen und deren Nutzen innerhalb einer Zeitscheibe modellieren
- Was passiert mit Aktionen und Nutzen über die Zeit?



# Annahmen über Aktionen

- Annahmen mit Bezug zur Zeit in sequentiellen Modellen
  - Diskrete Zeitschritte, Markov-1, Stationarität
- Führen zu Annahmen / Einschränkungen für Entscheidungen
  - Aktionen können von einem Zeitschritt zum nächsten ausgeführt werden (keine Dauer)
  - Effekt / Ergebnis der Aktion im aktuellen oder unmittelbar im nächsten Zeitschritt zu erfassen
  - Aktionen wirken sich nicht auf Stationarität aus



# Aktionen über die Zeit

- Gegeben eine Menge von Template-Entscheidungsvariablen $\boldsymbol{D} = \{D_1, \dots, D_n\}$
- Temporale Sequenz von Entscheidungen = sequentielle (temporale) Aktionszuweisung
  - Zusammengesetztes Event für Sequenz von Zeitschritten $t_1 : t_2$
  - I.e.,

$$\boldsymbol{d}^{(t_1:t_2)} = \left(\boldsymbol{d}^{(t_1)}, \boldsymbol{d}^{(t_1+1)}, \dots, \boldsymbol{d}^{(t_2-1)}, \boldsymbol{d}^{(t_2)}\right)$$

    - $\boldsymbol{d}^{(t)} = \left\{D_1^{(t)} = d_1, \dots, D_n^{(t)} = d_n\right\}, t \in \{t_1, \dots, t_2\}$
- Entscheidungsfindung über die Zeit beinhaltet für gewöhnlich die *beste* Sequenz zu berechnen, beginnend beim aktuellen Zeitschritt $t$ für die nächsten $k$ Schritte
  - I.e., die beste sequentielle Aktionszuweisung $\boldsymbol{d}^{(t:t+k)}$

Bisher haben wir Nutzen verwendet, um „das beste“ zu bestimmen. Wie können wir die im temporalen Kontext verwenden?

# Nutzen über die Zeit

- Über sequentielle Aktionszuweisungen iterieren und die mit dem maximalen erwarteten Nutzen wählen
  - Jede Zeitscheibe hat einen Nutzen (oder Reward)
    → Automatisch Mehrfachattribute in der Nutzentheorie bei Betrachtung als ausgerolltes Modell
    - Nutzen pro Zeitscheibe in einen erwarteten Nutzen der gesamten Sequenz zusammenfassen
  - Annahme: Präferenz einer Sequenz über eine andere Sequenz hängt nicht von der Zeit ab
    → Präferenzen sind stationär
    - Formal: Wenn zwei Zustandssequenzen $(\boldsymbol{r}_0, \boldsymbol{r}_1, \boldsymbol{r}_2, \dots)$ und $(\boldsymbol{s}_0, \boldsymbol{s}_1, \boldsymbol{s}_2, \dots)$ den gleichen Startzustand haben $(\boldsymbol{r}_0 = \boldsymbol{s}_0)$, dann sollten die zwei Sequenzen $(\boldsymbol{r}_1, \boldsymbol{r}_2, \dots)$ und $(\boldsymbol{s}_1, \boldsymbol{s}_2, \dots)$ die gleiche Ordnung haben bezogen auf die Präferenz wie $(\boldsymbol{r}_0, \boldsymbol{r}_1, \boldsymbol{r}_2, \dots)$ und $(\boldsymbol{s}_0, \boldsymbol{s}_1, \boldsymbol{s}_2, \dots)$

    → Präferenzunabhängigkeit zwischen Zeitscheiben

    → Additive Kombinationsfunktion, um Nutzen pro Zeitscheibe zusammenzufassen

# Temporale Kombinationsfunktion

- Rein additiv
  - Summe über individuellen Nutzen
$$U(\boldsymbol{r}_0, \boldsymbol{r}_1, \boldsymbol{r}_2, \dots) = U(\boldsymbol{r}_0) + U(\boldsymbol{r}_1) + U(\boldsymbol{r}_2) + \cdots$$
- Vermindert additiv (*discounted*)
  - Verallgemeinerung der rein additiven temporalen Kombinationsfunktion
  - Mittels Verminderungsfaktor (*discount factor*) $\gamma \in (0,1]$
    - Nutzen jetzt wichtiger als der in $k$ Schritten
    - Wenn $\gamma = 1$: additiv
  - Nutzen einer Sequenz = Summe über verminderten individuellen Nutzen
$$U(\boldsymbol{r}_0, \boldsymbol{r}_1, \boldsymbol{r}_2, \dots) = \gamma^0 \cdot U(\boldsymbol{r}_0) + \gamma^1 \cdot U(\boldsymbol{r}_1) + \gamma^2 \cdot U(\boldsymbol{r}_2) \dots$$
- Modell für sequentielle Entscheidungsfindung muss so eine temporale Kombinationsfunktion $\Psi$ beinhalten, um Nutzen über die Zeit zusammenzufassen

# Sequentielles Entscheidungsmodell

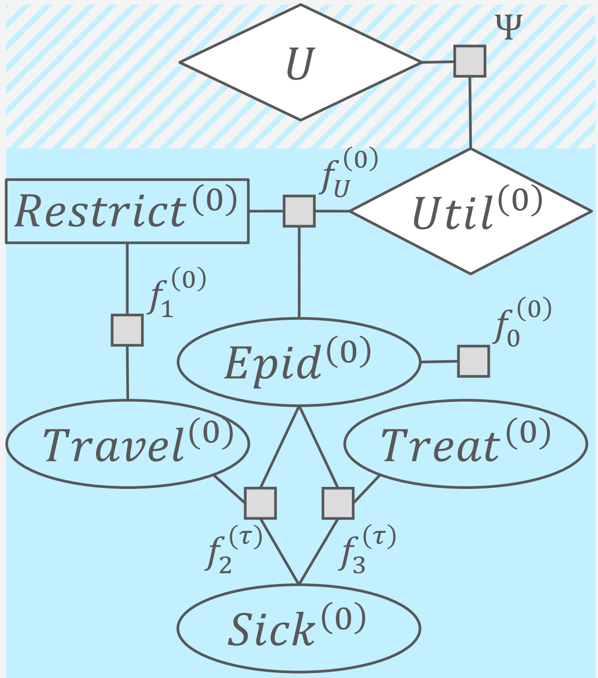
- Sequentielles Entscheidungsmodell $M$ ist ein dynamisches Modell $(M^0, M^\rightarrow)$, in dem die zwei episodischen Modelle $M^0, M^\rightarrow$ zeitindizierte Entscheidungsmodelle sind, erweitert um eine temporale Kombinationsfunktion $\Psi$
  - Zeit-indiziertes Entscheidungsmodell, optional mit expliziter Kombinationsfunktion $\Xi$

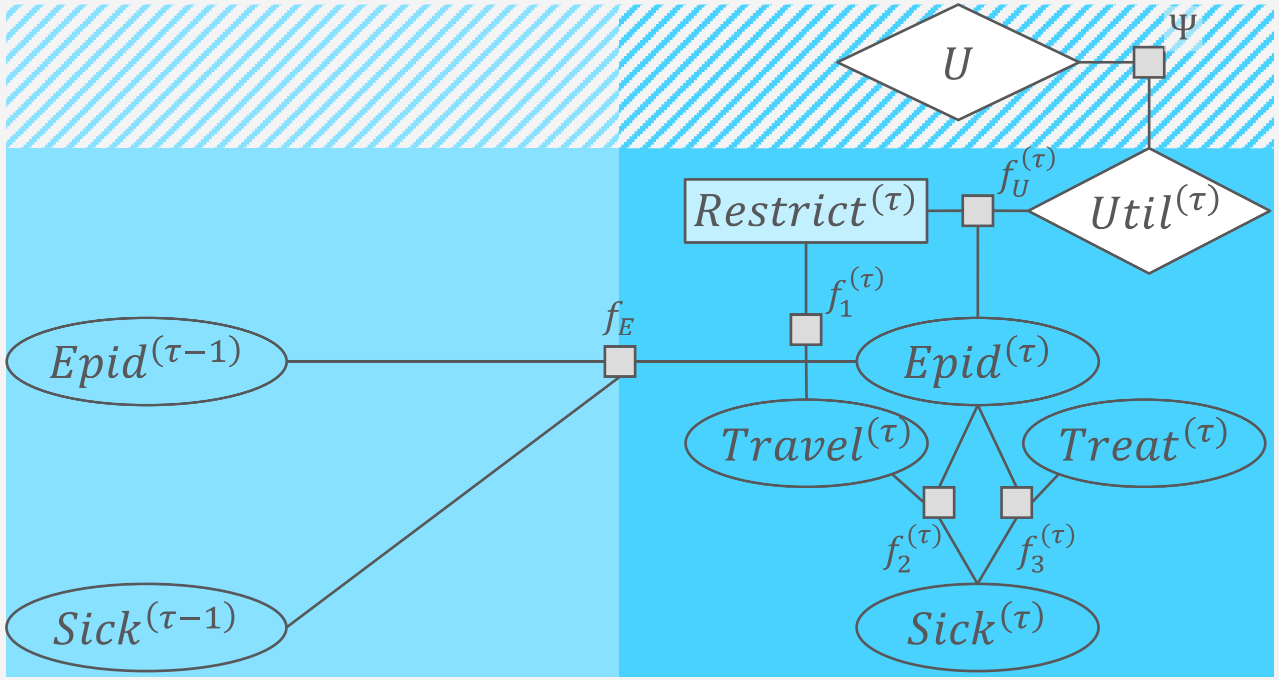
$$M^{(\tau)} = \left\{m_i^{(\tau)}\right\}_{i=1}^{n} \cup \left\{m_j\right\}_{j=1}^{l} \cup \left\{f_U^{(\tau)}\right\} \quad \text{oder} \quad M^{(\tau)} = \left\{m_i^{(\tau)}\right\}_{i=1}^{n} \cup \left\{m_j\right\}_{j=1}^{l} \cup \left\{f_u^{(\tau)}\right\}_{u=1}^{m} \cup \left\{\Xi^{(\tau)}\right\}$$

    - $f_U^{(\tau)} = \phi_U\left(\mathcal{R}^{(\tau)}\right)$ Nutzenfaktor, bildet auf Nutzenvariable $U^{(\tau)}$ ab
    - $f_u^{(\tau)} = \phi_u\left(R_u^{(\tau)}\right)$ Nutzenfaktoren, bilden auf Nutzenvariablen $U_u^{(\tau)}$ ab, die mittels $\Xi$ zu Gesamtnutzen $U^{(\tau)}$ einer Zeitscheibe $\tau$ kombiniert werden
    - Syntax gilt so für $M^\rightarrow$ und bei $m = 0$ für $M^0$
  - Temporale Kombinationsfunktion $\Psi$ additiv, um die $U^{(t)}$ zu einem Gesamtnutzen $U$ über die betrachteten Zeitschritte zu kombinieren

# Sequentielles Entscheidungsmodell: Beispiel

- Faktormodell mit $f_U^{(t)}$
  - Global: Gesamtnutzen $U$ als Ergebnis der temporalen Kombinationsfunktion $\Psi$



# EU-Anfrage über die Zeit

- EU-Anfrage zum Zeitpunkt $t$ über die temporale Aktionszuweisung $\boldsymbol{d}^{(t:t+k)}$ für einen Horizont von $k \in \mathbb{N}$ mit temporaler Kombinationsfunktion $\Psi$, Evidenz $\boldsymbol{e}^{(0:t)}$ und ausgeführten Aktionen $\boldsymbol{d}^{(0:t-1)}$
  - Für $M^{(\tau)} = \left\{m_i^{(\tau)}\right\}_{i=1}^{n} \cup \left\{m_j\right\}_{j=1}^{l} \cup \left\{f_U^{(\tau)}\right\}$

$$eu\left(\boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right) = \sum_{\boldsymbol{u} \in \mathrm{Val}\left(\mathrm{rv}\left(f_U^{(t:t+k)}\right)\right)} P\left(\boldsymbol{u} \middle| \boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right) \cdot \Psi\left(\phi_U^{(t:t+k)}(\boldsymbol{u})\right)$$

  - Für additives $\Psi$ mit $\gamma \in (0,1]$

$$eu\left(\boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right) = \sum_{\boldsymbol{u} \in \mathrm{Val}\left(\mathrm{rv}\left(f_U^{(t:t+k)}\right)\right)} P\left(\boldsymbol{u} \middle| \boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right) \cdot \sum_{k'=0}^{k} \gamma^{k'} \cdot \phi_U^{(t+k')}(\boldsymbol{u})$$

# EU-Anfrage über die Zeit

- EU-Anfrage zum Zeitpunkt $t$ über die temporale Aktionszuweisung $\boldsymbol{d}^{(t:t+k)}$ für einen Horizont von $k \in \mathbb{N}$ mit temporaler Kombinationsfunktion $\Psi$, Evidenz $\boldsymbol{e}^{(0:t)}$ und ausgeführten Aktionen $\boldsymbol{d}^{(0:t-1)}$
  - Für $M^{(\tau)} = \left\{m_i^{(\tau)}\right\}_{i=1}^{n} \cup \left\{m_j\right\}_{j=1}^{l} \cup \left\{f_u^{(\tau)}\right\}_{u=1}^{m} \cup \left\{\Xi^{(\tau)}\right\}$, $\boldsymbol{U} = \mathrm{rv}\left(\left\{f_u^{(t:t+k)}\right\}_{u=1}^{m}\right)$

$$eu\left(\boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right) = \sum_{\boldsymbol{u} \in \mathrm{Val}(\boldsymbol{U})} P\left(\boldsymbol{u} \middle| \boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right) \cdot \Psi\left( \times_{k'=0}^{k} \Xi^{(t+k')}\left( \times_{u=1}^{m} \phi_u^{(t+k')}\underbrace{(\boldsymbol{u}_u)}_{\boldsymbol{u}_u = \pi_{\boldsymbol{R}_u}(\boldsymbol{u})} \right)\right)$$

  - Für additives $\Psi$ mit $\gamma \in (0,1]$ und additives $\Xi$

$$eu\left(\boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right) = \sum_{\boldsymbol{u} \in \mathrm{Val}(\boldsymbol{U})} P\left(\boldsymbol{u} \middle| \boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right) \cdot \sum_{k'=0}^{k} \gamma^{k'} \cdot \sum_{u=1}^{m} \phi_u^{(t+k')}(\boldsymbol{u}_u)$$

# EU-Anfrage über die Zeit: Exakt → Problem

Realistisch **nicht** berechenbar!

- EU-Anfrage beinhaltet Wahrscheinlichkeitsanfrage an Zufallsvariablen aus den Nutzenfunktionen über $k$ Zeitschritte

$$eu\left(\boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right) = \sum_{\boxed{\boldsymbol{u} \in \mathrm{Val}\left(\mathrm{rv}\left(f_U^{(t:t+k)}\right)\right)}} \boxed{P\left(\boldsymbol{u} \middle| \boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right)} \cdot \Psi\left(\phi_U^{(t:t+k)}(\boldsymbol{u})\right)$$

  - Bedeutet Ausrollen des Modells über $k$ Zeitschritte
  - Unabhängig von der Ausprägung der Nutzen- und Kombinationsfunktionen
    - Für additives $\Psi$ mit $\gamma \in (0,1]$ und additives $\Xi$ für Menge von Nutzenfaktoren

$$eu\left(\boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right) = \sum_{\boxed{\boldsymbol{u} \in \mathrm{Val}(\boldsymbol{U})}} \boxed{P\left(\boldsymbol{u} \middle| \boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+\kappa)}\right)} \cdot \sum_{k'=0}^{k} \gamma^{k'} \cdot \sum_{u=1}^{m} \phi_u^{(t+k')}(\boldsymbol{u}_u)$$

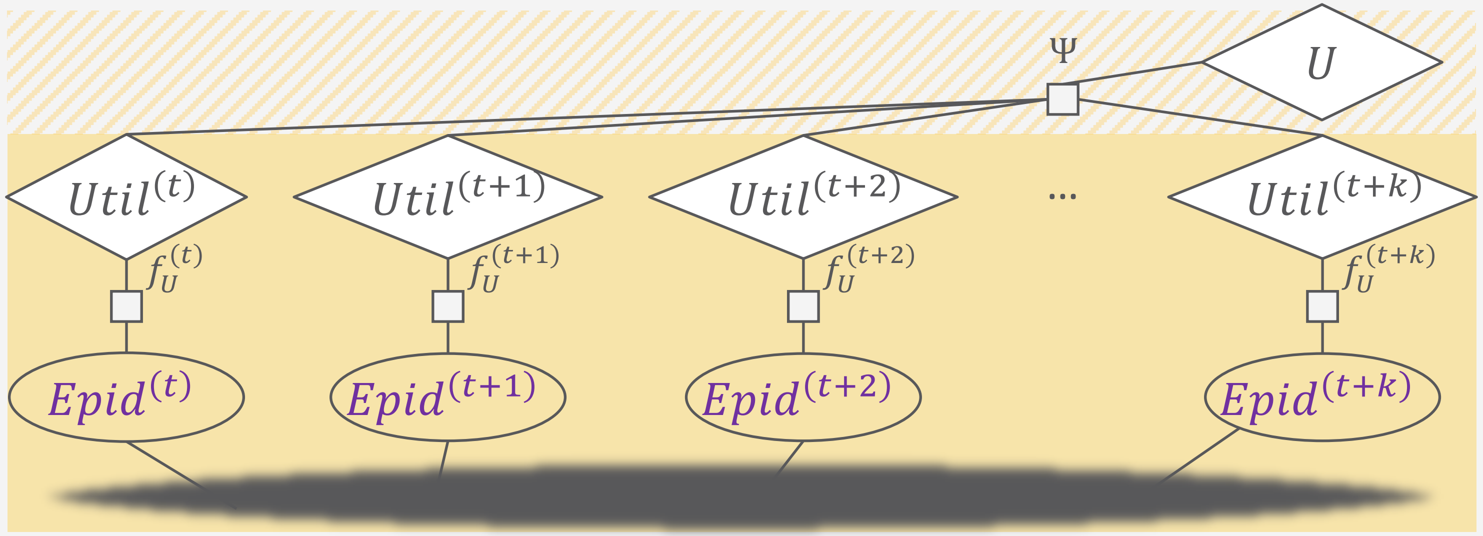
  - $\boxed{\boldsymbol{U} = \mathrm{rv}\left(\left\{f_u^{(t:t+k)}\right\}_{u=1}^{m}\right)}$

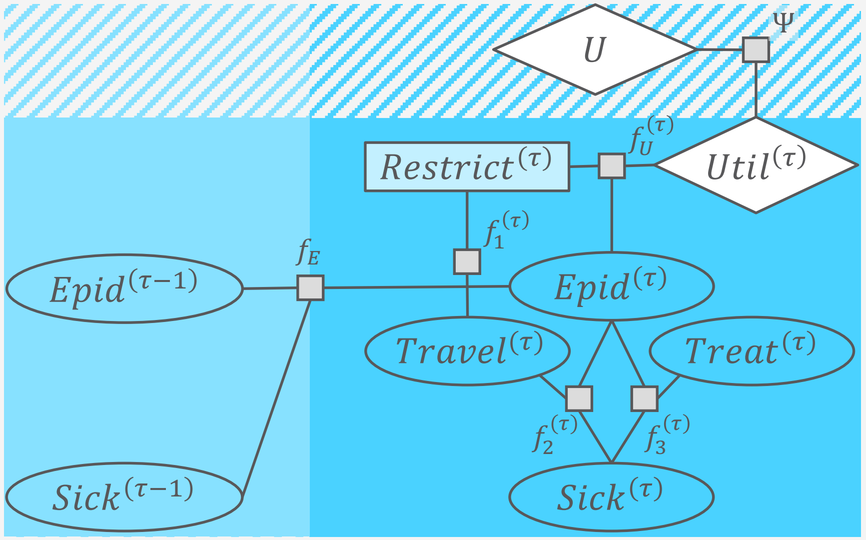
# EU-Anfrage über die Zeit: Exakt → Problem

Realistisch **nicht** berechenbar!

- EU-Anfrage beinhaltet Wahrscheinlichkeitsanfrage an Zufallsvariablen aus den Nutzenfunktionen über $k$ Zeitschritte

$$eu\left(\boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right) = \sum_{\boxed{\boldsymbol{u} \in \mathrm{Val}\left(\mathrm{rv}\left(f_U^{(t:t+k)}\right)\right)}} \boxed{P\left(\boldsymbol{u} \middle| \boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right)} \cdot \Psi\left(\phi_U^{(t:t+k)}(\boldsymbol{u})\right)$$



# EU-Anfrage über die Zeit: Solution?

- Anstatt:

$$eu\left(\boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right) = \sum_{\boxed{\boldsymbol{u} \in \mathrm{Val}\left(\mathrm{rv}\left(f_U^{(t:t+k)}\right)\right)}} \boxed{P\left(\boldsymbol{u} \middle| \boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right)} \cdot \Psi\left(\phi_U^{(t:t+k)}(\boldsymbol{u})\right)$$

- Wahrscheinlichkeitsanfrage pro Zeitschritt berechnen
  - Für additives $\Psi$ mit $\gamma \in (0,1]$

$$eu\left(\boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right) = \sum_{\boldsymbol{u} \in \mathrm{Val}\left(\mathrm{rv}\left(f_U^{(t:t+k)}\right)\right)} \sum_{k'=0}^{k} \gamma^{k'} \cdot P\left(\boldsymbol{u}^{(t+k')} \middle| \boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right) \cdot \phi_U^{(t+k')}(\boldsymbol{u})$$

$$= \sum_{k'=0}^{k} \gamma^{k'} \cdot \underbrace{\sum_{\boldsymbol{u} \in \mathrm{Val}\left(\mathrm{rv}\left(f_U^{(t:t+k)}\right)\right)} P\left(\boldsymbol{u}^{(t+k')} \middle| \boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right) \cdot \phi_U^{(t+k')}(\boldsymbol{u})}_{eu\left(\boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right)}$$

Summe über erwarteten Nutzen pro $t$ nur exakt bei *Unabhängigkeit zwischen den Zeitscheiben* gegeben $\boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}$

- *Unrealistische Annahme* → Approximative Lösung!

# EU-Anfrage über die Zeit: Approximation bis ins Innere der EU-Anfrage

- Anstatt:

$$eu\left(\boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right) = \sum_{\boxed{\boldsymbol{u} \in \mathrm{Val}\left(\mathrm{rv}\left(\boldsymbol{f}_U^{(t:t+k)}\right)\right)}} \boxed{P\left(\boldsymbol{u} \middle| \boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right)} \cdot \Psi\left(\phi_U^{(t:t+k)}(\boldsymbol{u})\right)$$

- Wahrscheinlichkeitsanfrage pro Zeitschritt berechnen
  - Für additives $\Psi$ mit $\gamma \in (0,1]$

$$eu\left(\boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right) = \sum_{\boldsymbol{u} \in \mathrm{Val}\left(\mathrm{rv}\left(\boldsymbol{f}_U^{(t:t+k)}\right)\right)} \sum_{k'=0}^{k} \gamma^{k'} \cdot P\left(\boldsymbol{u}^{(t+k')} \middle| \boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right) \cdot \phi_U^{(t+k')}(\boldsymbol{u})$$

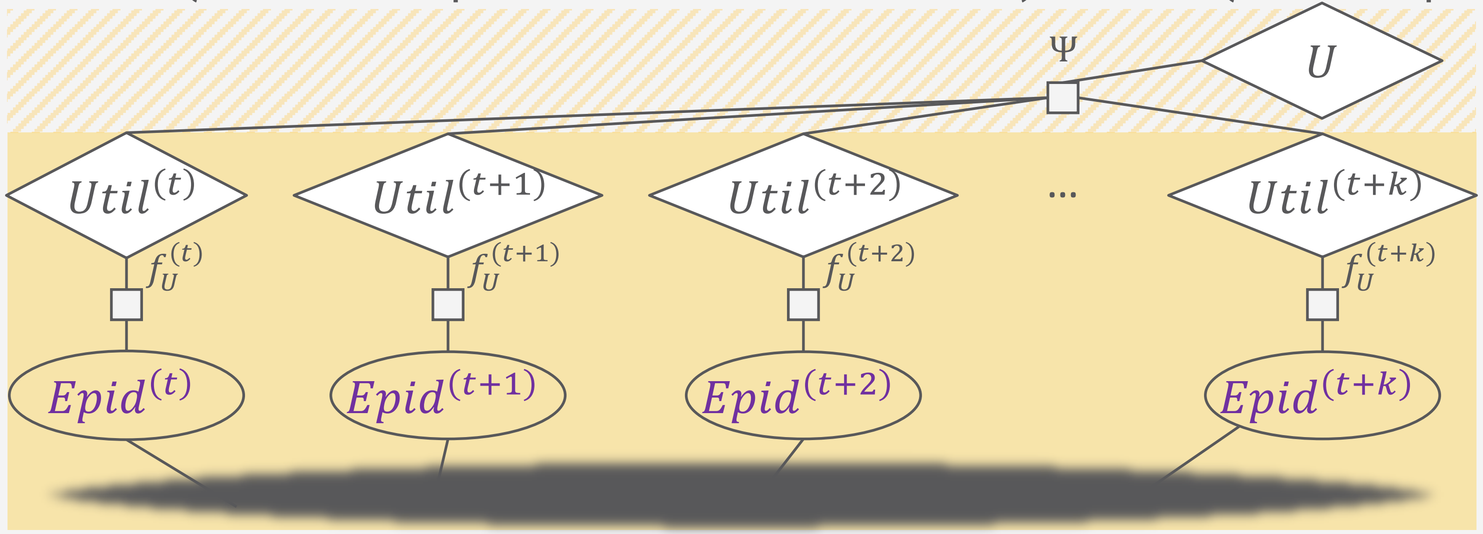
  - Für additives $\Psi$ mit $\gamma \in (0,1]$, plus Vereinfachung für additives $\Xi$ für Menge von Nutzenfaktoren

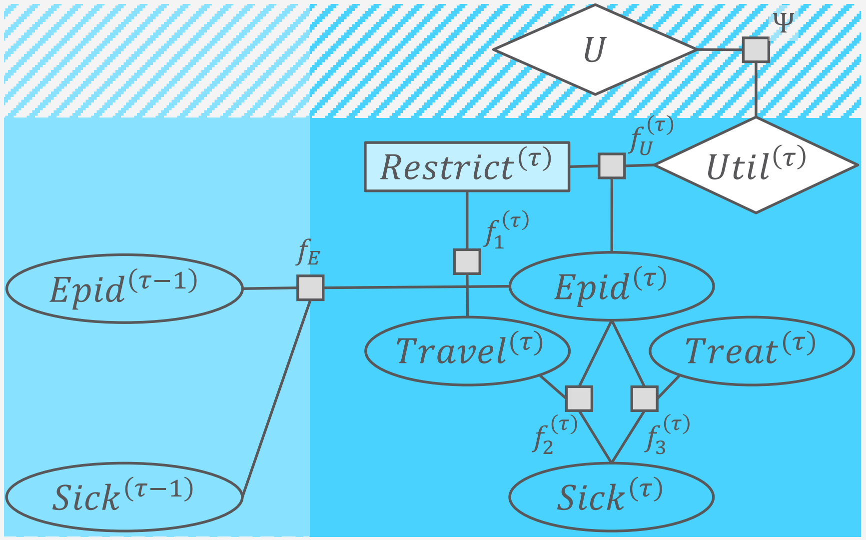
$$eu\left(\boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right) = \sum_{\boldsymbol{u} \in \mathrm{Val}(\boldsymbol{U})} \sum_{k'=0}^{k} \gamma^{k'} \cdot \sum_{u=1}^{m} P\left(\boldsymbol{u}_u \middle| \boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right) \cdot \phi_u^{(t+k')}(\boldsymbol{u}_u)$$

# EU-Anfrage über die Zeit: Beispiel

- EU-Anfrage: $eu\left(v_{Treat}^{(0:t)}, d_{Restrict}^{(0:t-1)}, d_{Restrict}^{(t:t+k)}\right)$ mit $r_R^{(t_1:t_2)}$ Werte $r \in \mathrm{Val}(R)$ für Variable $R$ über Intervall $t_1 : t_2$
  - Wahrscheinlichkeitsanfrage
    - $P\left(Epid^{(t:t+k)} \middle| v_{Treat}^{(0:t)}, d_{Restrict}^{(0:t-1)}, d_{Restrict}^{(t:t+k)}\right)$ wird $P\left(Epid^{(t')} \middle| v_{Treat}^{(0:t)}, d_{Restrict}^{(0:t-1)}, d_{Restrict}^{(t:t+k)}\right), t' \in [t, t+k]$

Exakt oder approximativ?



# Sequentielles MEU Problem

- Gegeben ein sequentielles Entscheidungsmodell $M$, Evidenz $\boldsymbol{e}^{(0:t)}$, umgesetzte Entscheidungen $\boldsymbol{d}^{(0:t-1)}$ (ausgeführte Aktionen) und ein Horizont $k$
- Sequentielles MEU Problem
  - Finde sequentielle Aktionszuweisung $\boldsymbol{d}^{(t:t+k)}$ mit maximalem erwarteten Nutzen in $M$

$$MEU\left(\boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, k\right) = \left(\boldsymbol{d}^*, eu\left(\boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^*\right)\right)$$

$$\boldsymbol{d}^* = \underset{\boldsymbol{d}^{(t:t+k)} \in \mathrm{Val}\left(\boldsymbol{D}^{(t:t+k)}\right)}{\arg\max} eu\left(\boldsymbol{e}^{(0:t)}, \boldsymbol{d}^{(0:t-1)}, \boldsymbol{d}^{(t:t+k)}\right)$$

  - Größe von $\mathrm{Val}\left(\boldsymbol{D}^{(t:t+k)}\right)$ exponentiell abhängig von der Anzahl an Entscheidungsvariablen und $k$

# Lösen eines sequentiellen MEU Problems: MEU-DJT

- Gegeben sequentielles Entscheidungsmodell $M$, anwachsende Evidenz $\boldsymbol{e}^{(0:t)}$, Horizont $k$
  - Ermöglicht auch Filterungs-, Vorhersage- und Rückschau-, MPE- und MAP-Anfragen $\boldsymbol{Q}$
- Vorgehen:
  - Dynamische Jtrees $J^0, J^{\rightarrow}$ bauen (evtl. Hilfsfaktor für einfachere EU-Anfragen nutzen)
  - Für $t = 0, \dots,$
    - $J^{(t)}$ instanziieren, $\alpha^{(t-1)}$ hinzufügen, $\boldsymbol{e}^{(t)}$ behandeln, Nachrichten versenden, $\boldsymbol{Q}$ beantworten
    - MEU Schleife: Für jede mögliche temporale Aktionszuweisung $\boldsymbol{d}^{(t:t+k)}$
      - Berechne EU-Anfragen, genau oder approximativ;
      - Gebe $\boldsymbol{d}^*$ als die temporale Aktionszuweisung $\boldsymbol{d}^{(t:t+k)}$ mit größtem erwarteten Nutzen zurück ($\rightarrow$ $\boldsymbol{d}^{(t)}$ ausführen)
    - Setze $\boldsymbol{d}^{(t)}$, berechne $\alpha^{(t)}$

Berechnet jeden Zeitschritt die MEU-Anfrage neu; adaptieren möglich, so dass $\boldsymbol{d}^*$ für $k$ Zeitschritte, oder bis die Ausführung zu unsicher erscheint, ausgeführt wird

- $\boldsymbol{d}^*$ speichern und, anstatt in die MEU-Schleife zu gehen, die nächsten Aktionen aus $\boldsymbol{d}^*$ setzen

# Lösen eines sequentiellen MEU Problems: MEU-DJT

- Berechnen einer EU-Anfrage für $\boldsymbol{d}^{(t:t+k)}$ in einem Jtree
  - Bei exaktem Vorgehen:
    - Jtree für $k$ Schritte ausrollen, $\boldsymbol{d}^{(t:t+k)}$ setzen, Wahrscheinlichkeitsanfrage berechnen, erwarteten Nutzen berechnen
  - Bei approximativem Vorgehen
    - Gesamtnutzen $U \leftarrow 0$
    - Für $t' = t, \dots, t + k$
      - Jtree $J^{(t')}$ instanziieren, $\alpha^{(t'-1)}$ hinzufügen (wenn beides nicht schon passiert)
      - $\boldsymbol{d}^{(t')}$ setzen, Nachrichten berechnen (aktualisieren), Wahrscheinlichkeitsanfrage berechnen, erwarteten Nutzen berechnen, Ergebnis zu $U$ addieren
      - $\alpha^{(t')}$ berechnen

# Sequentielles MEU Problem: Beispiel

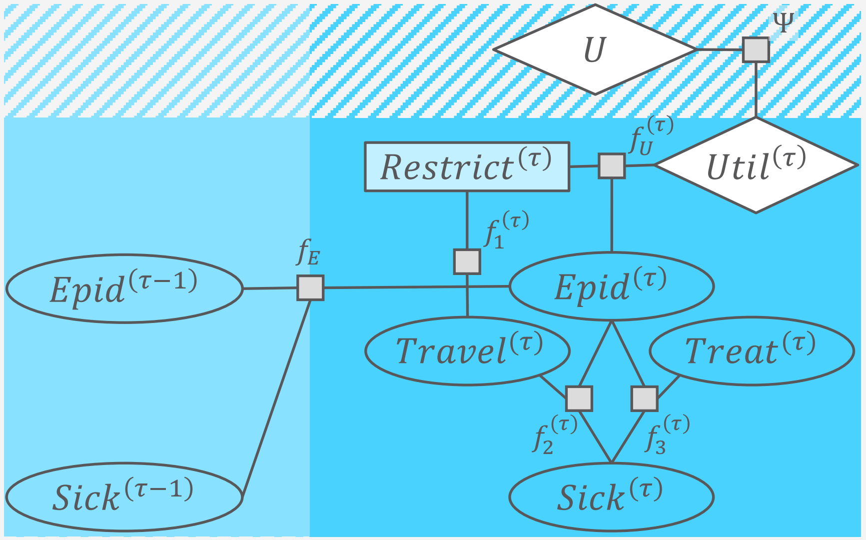
- Gegeben keine Evidenz, ausgeführte Aktionen $\{ban^{(0)}, ban^{(1)}\}$, $t = 2$ und $k = 2$

$$MEU\left(\emptyset^{(0:2)}, \{ban^{(0)}, ban^{(1)}\}, 2\right) = \left(\boldsymbol{d}^*, eu\left(\emptyset^{(0:2)}, \{ban^{(0)}, ban^{(1)}\}, \boldsymbol{d}^*\right)\right)$$

$$\boldsymbol{d}^* = \underset{\boldsymbol{d}^{(2:4)} \in \mathrm{Val}\left(\boldsymbol{D}^{(2:4)}\right)}{\arg\max} \; eu\left(\emptyset^{(0:2)}, \{ban^{(0)}, ban^{(1)}\}, \boldsymbol{d}^{(2:4)}\right)$$

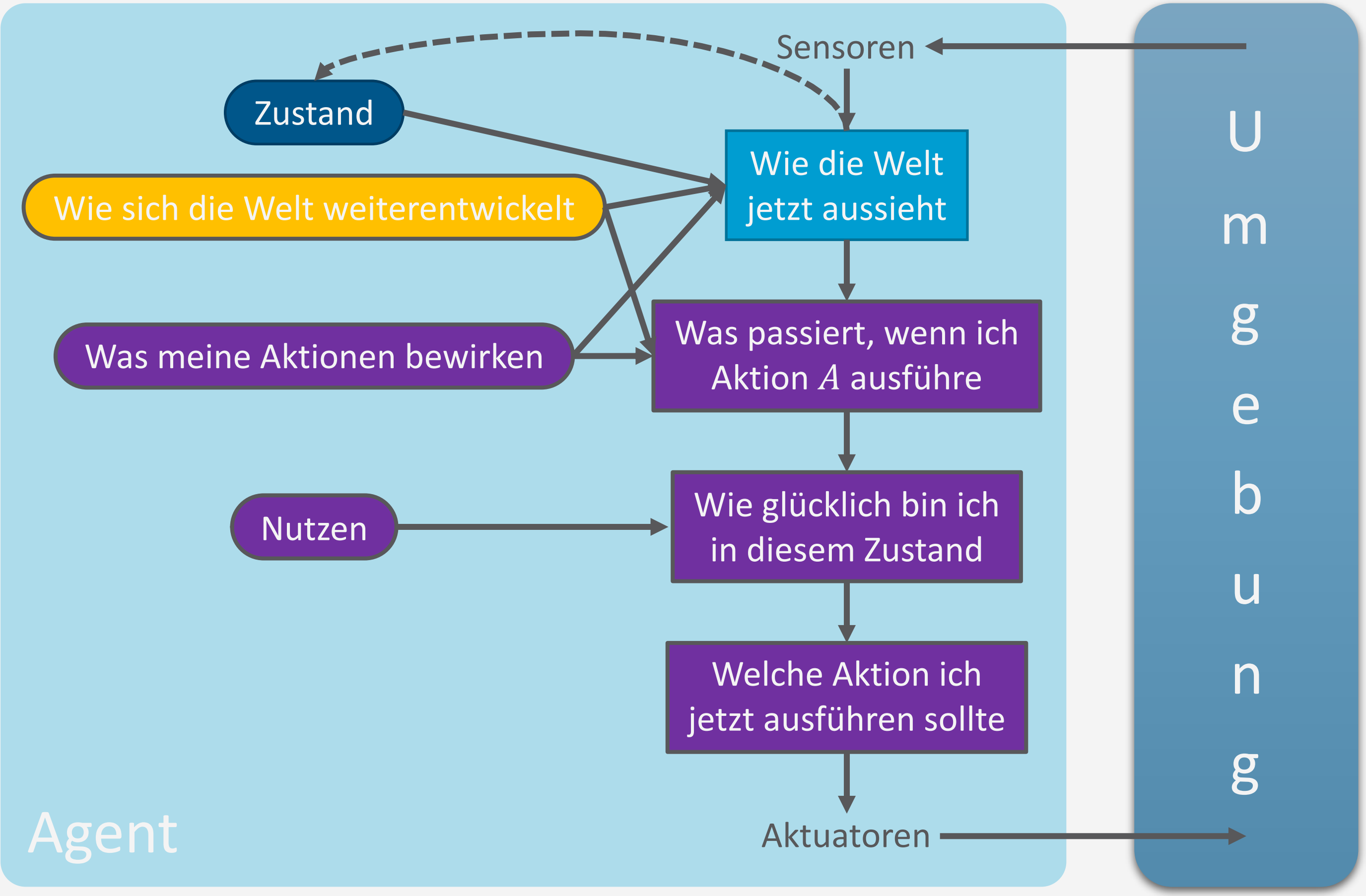
- Acht sequentielle Aktionszuweisungen testen
  - $\boldsymbol{d}_i^{(2:4)}$ im ausgerollten $\boldsymbol{J}^{(2:4)}$ setzen und EU-Anfrage beantworten
  - $\boldsymbol{d}_i^{(2:4)}$ nach und nach in $\boldsymbol{J}^{(t')}$ setzen und EU-Antworten addieren
  - Argmax $\boldsymbol{d}_i^{(2:4)}$ ausgeben

| | $d^{(2)}$ | $d^{(3)}$ | $d^{(4)}$ |
|---|---|---|---|
| $\boldsymbol{d}_1^{(2:4)}$ | $ban$ | $ban$ | $ban$ |
| $\boldsymbol{d}_2^{(2:4)}$ | $ban$ | $ban$ | $free$ |
| $\boldsymbol{d}_3^{(2:4)}$ | $ban$ | $free$ | $ban$ |
| $\boldsymbol{d}_4^{(2:4)}$ | $ban$ | $free$ | $free$ |
| $\boldsymbol{d}_5^{(2:4)}$ | $free$ | $ban$ | $ban$ |
| $\boldsymbol{d}_6^{(2:4)}$ | $free$ | $ban$ | $free$ |
| $\boldsymbol{d}_7^{(2:4)}$ | $free$ | $free$ | $ban$ |
| $\boldsymbol{d}_8^{(2:4)}$ | $free$ | $free$ | $free$ |



# Entscheidungen treffen und danach handeln

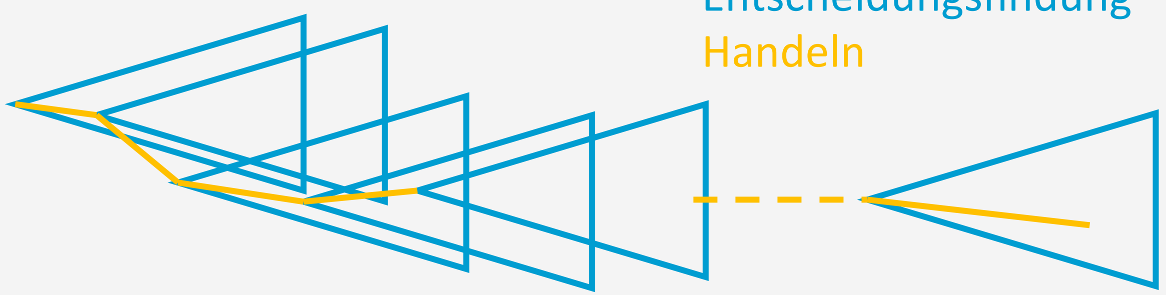
- Berechnete temporalen Aktionszuweisung umsetzen:
  - Aktuatoren bekommen Befehle zur Ausführung der aktuellen Aktion
  - Interner Zustand kann aktualisiert werden: Entscheidungsvariablen setzen, (Nachrichten versenden), Übergang zu nächster Zeitscheibe
- Agent kann dann weiter gemäß der temporalen Aktionszuweisung handeln oder eine neue temporale Aktionszuweisung berechnen gegeben der neuen Evidenz



# Entscheidungen treffen und danach handeln

- Zurückweichender Horizont:
  - Einen meu-Algorithmus aufrufen, um eine temporale Aktionszuweisung $\pi$ für die nächsten $k$ Schritte zu berechnen, zu erhalten, erste Aktion ausführen; meu aufrufen, …
- Nützlich, wenn unvorhersehbare Dinge häufig bzw. wahrscheinlich passieren
  - Unmittelbar neue Aktionszuweisung berechnen
- Mögliches Problem:
  - Kann wiederholt zu Unterbrechungen führen, während die Routine darauf wartet, dass meu die neue Aktionszuweisung zurückgibt

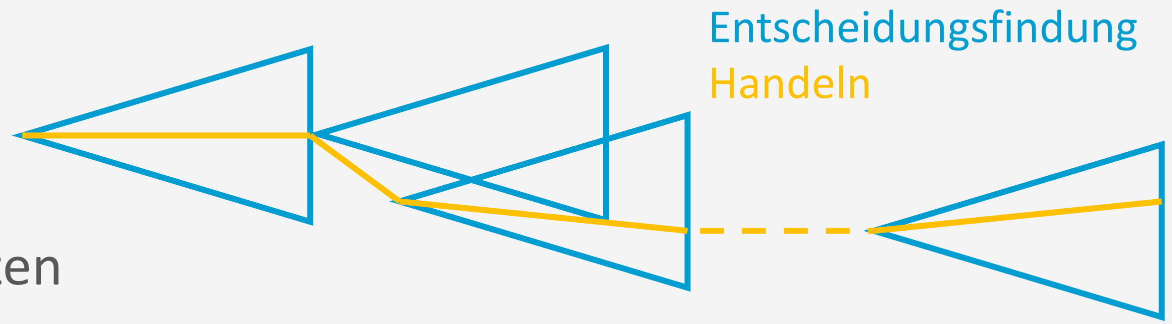
```
Run-MEU(M,k)
  while e ← new evidence do
    absorb e in M
    π ← meu(M,k)
    a ← pop-first(π)
    perform a
    update M with a
```



# Entscheidungen treffen und danach handeln

- meu aufrufen, temporale Aktionszuweisung $\pi$ soweit wie möglich ausführen, meu nur wieder aufrufen, wenn nötig
- Simulate testet, ob $\pi$ noch richtig ausführbar ist
  - Ohne Simulate möglich Probleme nicht zu bemerken, bis es zu spät ist
  - Durchführung z.B. über:
    - Aktionen verfeinern und Anwendbarkeit prüfen
    - Technische Simulation der Ausführung
- Mögliches Problem
  - Kann Möglichkeiten verpassen, $\pi$ mit einer besseren temporale Aktionszuweisung zu ersetzen

```
Run-Lazy-MEU(M,k)
   while e ← new evidence do
      absorb e in M
      π ← meu(M,k)
      while π ≠ () and Simulate(M,π)
               ≠ failure do
         a ← pop-first(π)
         perform a
         update M with a
```



# Entscheidungen treffen und danach handeln

- Thread 1 für laufende Neuberechnung einer temporalen Aktionssequenz $\pi$
- Thread 2 für Ausführung des aktuell vorhandenen $\pi$
- Kann Möglichkeiten früher bemerken $\pi$ durch etwas besseres zu ersetzen als `Run-Lazy-MEU`
  - Kann aber Dinge verpassen, die `Run-MEU` findet
- Auch hier: Ohne Simulate möglich Probleme nicht zu bemerken, bis es zu spät ist
  - Nicht so schlimm wie bei `Run-Lazy-MEU`, da ein Thread immer nach der nächst besseren Sequenz sucht

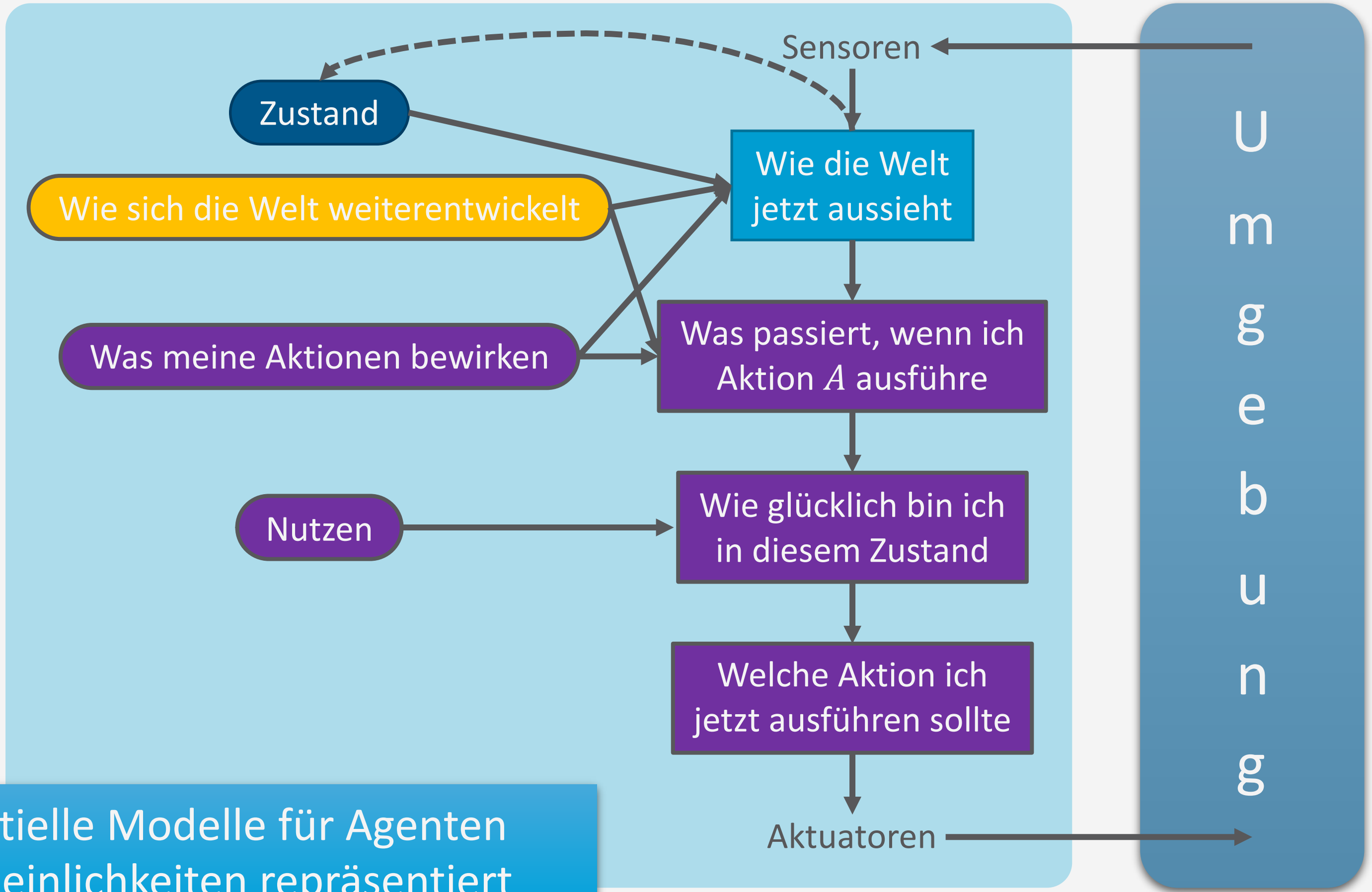
```
Run-Concurrent-MEU(M,k)
   π ← ⟨⟩
   // thread 1 + 2 run concurrently
   thread 1:
      while e ← new evidence do
         absorb e in M
         π ← meu(M,k)
   thread 2:
      while e ← new evidence do
         absorb e in M
         if π ≠ () and Simulate(G,π)
                  ≠ failure then
            a ← pop-first(π)
            perform a
            update M with a
```

# Zwischenzusammenfassung

- Sequentielle Entscheidungsmodelle als Kombination von
  - Sequentiellen Modellen für die zeitlichen Aspekte
  - Entscheidungsmodelle für Entscheidungsfindung
- Sequentieller erwarteter Nutzen
  - Reiner oder verminderter additiver Nutzen
  - Exakte Lösung so gut wie unmöglich zu berechnen
  - Annäherung durch Summe über die Ergebnisse von EU-Anfragen pro Zeitschritt
- Algorithmus basierend auf DJT
  - Lösung mittels Approximation kann die zeitlichen Unabhängigkeiten nutzen
- Handeln
  - Nach jedem Schritt neuberechnen; nur neu berechnen, wenn Simulation versagt; zwei Threads

# Einordnung der Vorlesung: *Modell- und nutzenbasierter Agent*

- Nachfolgende Themen der Vorlesung
  2. Episodische PGMs
  3. Exakte Inferenz in episodischen PGMs
  4. Approximative Inferenz in episodischen PGMs
  5. Lernalgorithmen für episodische PGMs
  6. Sequentielle PGMs und Inferenz
  7. Entscheidungstheoretische PGMs und Inferenz




Entscheidungstheoretische sequentielle Modelle für Agenten
- Unsicherheiten mittels Wahrscheinlichkeiten repräsentiert
- Zeitliche Einflüsse durch Zeitscheiben modelliert
- Entscheidungen und Effekte über Aktionen und deren Nutzen

# Überblick: 7. Entscheidungstheoretische PGMs und Inferenz

A. *Nutzentheorie*
   - Präferenzen, Nutzen, Maximum-Expected-Utility Prinzip, Dominanz und Unabhängigkeiten

B. *Entscheidungstheoretische episodische PGMs und Inferenz*
   - Syntax: Knoten für Nutzen / Entscheidungen, Einflussdiagram, Entscheidungsmodell; Semantik
   - Inferenzaufgaben: Erwarteter Nutzen, Aktionen mit höchstem erwarteten Nutzen
   - Entscheidungstheoretische VE-Variante
   - Informationswerttheorie

C. *Entscheidungstheoretische sequentielle PGMs und Inferenz*
   - Zeitliche Inferenzaufgaben
   - VE-basierter Inferenzalgorithmus
   - Automatisiertes Handeln

→ Abschlussbetrachtungen