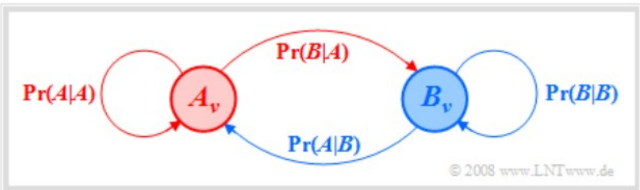
Teil 2 der 5. VL: Markov

Im zweiten Teil der Vorlesung 5 beschäftigen wir uns mit Markov. Um genauer zu sein, werden wir uns mit den Markov Modellen, Markov Ketten und dem hidden Markov Modell beschäftigen. Dieser Teil ist also in 3 Teile aufgeteilt.

Teil 1: Das Markov Modell

- behandelt Prognosen über zukünftige Entwicklung
- die zukünftige Entwicklung hängt nur vom aktuellen Zustand (Gegenwart) ab und nicht von den vorherigen Zuständen (Vergangenheit)
- ist nach dem russischen Mathematiker Andrey Markov benannt



- stochastische Form des Automatenmodells
- stochastische Methode für sich zufällig ändernde Systeme
- Zustandsübergänge werden mit Übergangswahrscheinlichkeiten und Zustandswahrscheinlichkeiten modelliert

Frage: Wofür wird das Markov Modell benutzt – Was wird damit vorhergesagt?

Das Markov-Modell, benannt nach dem russischen Mathematiker Andrey Markov, ist eine mathematische Methode zur **Vorhersage zukünftiger Zustände basierend auf gegenwärtigen Zuständen**.

Frage: Welche Annahme macht das Markov-Modell? Wovon hängen zukünftige Ereignisse ab und wovon eben nicht?

Es basiert auf der Idee, dass die zukünftige Entwicklung eines Systems oder Prozesses **nur von seinem aktuellen Zustand abhängt und nicht von den vergangenen Ereignissen**. Wir schauen uns also nicht die Vergangenheit an, sondern *nur die Wahrscheinlichkeit, mit der aktuelle Zustände auftreten, und die Wahrscheinlichkeit, mit der einzelne Zustände in andere Zustände überführt werden*.

Markov-Modelle sind stochastisch, weil sie auf probabilistischen Annahmen und Zufallswahrscheinlichkeiten basieren. Dies bedeutet, dass die Entwicklung eines Systems in der Zukunft nicht deterministisch vorhergesagt wird, sondern vielmehr auf Wahrscheinlichkeiten beruht.
Ein Markov-Modell geht davon aus, dass die zukünftige Entwicklung eines Systems oder Prozesses von seinem gegenwärtigen Zustand abhängt, aber es gibt Unsicherheit darüber, welcher Zustand als nächstes erreicht wird. Die Übergänge zwischen den Zuständen sind stochastisch, da sie mit Wahrscheinlichkeiten verknüpft sind. Diese Wahrscheinlichkeiten geben an, mit welcher Wahrscheinlichkeit das System in einen bestimmten nächsten Zustand übergeht.
Markov-Modell ist stochastisch.

Das Automatenmodell ist eine mathematische Abstraktion, die zur Beschreibung von Zustandsübergängen in einem System oder Prozess verwendet wird. Ein Automat besteht aus einer endlichen Menge von Zuständen und Übergangsregeln, die bestimmen, wie der Automat von einem Zustand in einen anderen wechselt. Es gibt verschiedene Arten von Automaten, darunter endliche Automaten (finite state machines), Moore-Automaten, Mealy-Automaten und mehr. Diese Modelle werden oft in der Informatik, Automatentheorie und im Bereich der formalen Sprachen verwendet, um Algorithmen und Verhaltensweisen von Systemen zu beschreiben und zu analysieren.
Markov-Modelle sind eine spezielle Art von Automaten, die als Markov-Ketten bezeichnet werden. Sie sind stochastisch, da sie auf Wahrscheinlichkeiten und Zustandsübergänge basieren. In einem Markov-Modell gibt es eine endliche Menge von Zuständen, und die Übergänge zwischen diesen Zuständen werden mit Wahrscheinlichkeiten verknüpft. Anders als in deterministischen Automaten, bei denen der Übergang von einem Zustand in den nächsten eindeutig festgelegt ist, erfolgen die Übergänge in Markov-Ketten auf probabilistischer Basis.

- Markov-Entscheidungsprozess:
  - bei kontrollierten Systemen mit vollständig beobachtbaren Zuständen
  - ist gedächtnislos; die Vorgeschichte des Prozesses spielt keine Rolle

- Markov-Annahme:
  - der Folgezustand hängt nur vom aktuellen Zustand ab
  - dadurch lassen sich vergleichbar gute Ergebnisse erzielen

- teilweise beobachtbare Markov-Entscheidungsprozesse:
  - bei kontrollierten Systemen mit teilweise beobachtbaren Zuständen

Frage: Welche Merkmale und Voraussetzungen charakterisieren den Markov-Entscheidungsprozess (MDP) in Bezug auf die beobachteten Zustände, und wie unterscheidet sich ein MDP in Bezug auf sein Gedächtnis von anderen Modellierungsmethoden in der Entscheidungstheorie?

Der Markov-Entscheidungsprozess wird bei kontrollierten Systemen *mit* ***vollständig beobachteten Zuständen*** eingesetzt. Das heißt, wir kennen alle Zustände. Es ist weiterhin ***Gedächtnislos***. Das heißt – wie bereits erwähnt, dass wir uns nicht die Vergangenheit anschauen, sondern nur die aktuellen Zustände und ihre Übergangswahrscheinlichkeiten.

Beispiel:

Hier ist ein einfaches Beispiel: Angenommen, Sie haben ein Schachbrett, und Sie können immer sehen, welche Schachfigur sich auf welchem Feld befindet. Das Schachbrett ist der Zustand des Systems. Wenn Sie am Zug sind, können Sie basierend auf der aktuellen Position des Spiels und den Regeln eine Entscheidung treffen, wie Sie Ihre nächste Schachfigur bewegen möchten. Sie brauchen nicht die gesamte Historie des Spiels zu kennen, sondern nur den aktuellen Zustand des Schachbretts, um Ihre Entscheidung zu treffen. Das wäre eine Art von MDP, da Sie Entscheidungen auf der Grundlage des vollständig beobachteten aktuellen Zustands treffen.

Frage: Welche Annahme macht das Markov-Modell? **Welche Annahme resultiert aus dem MDP**?

Das Markov-Modell macht die Annahme, dass die zukünftige Entwicklung eines Systems nur von seinem gegenwärtigen Zustand abhängt und nicht von den vergangenen Ereignissen. Dies wird oft als die "Markov'sche Eigenschaft" oder das "Markov'sche Prinzip" bezeichnet. In anderen Worten, das Modell geht davon aus, dass die Zukunft eines Systems oder Prozesses nur durch den aktuellen Zustand und die zugehörigen Wahrscheinlichkeiten für Zustandsübergänge beschrieben werden kann, ohne die Notwendigkeit, die gesamte Vergangenheit des Systems zu kennen. Dadurch lassen sich vergleichbar gute Ergebnisse erzielen.

F**rage:Was sind teilweise beobachtbare Markov-Entscheidungsprozesse?**

Es gibt auch teilweise beobachtbare Markov-Entscheidungsprozesse bei kontrollierten Systemen mit teilweise beobachtbaren Zuständen. Das heißt, nicht alle Zustände sind bekannt. Dazu kommen wir später.

Ein gutes Beispiel für einen teilweise beobachtbaren Markov-Entscheidungsprozess (POMDP) ist die Navigation eines autonomen Roboters in einem unbekannten Gebäude. Der Roboter bewegt sich in einem Gebäude, in dem er sich nicht auskennt und bei dem er nicht genau weiß, in welchem Raum er sich gerade befindet oder welchen Zustand das Gebäude hat.

Die Situation ist teilweise beobachtbar, da der Roboter zwar Sensoren hat, die ihm Informationen über seine Umgebung liefern, aber diese Informationen können unvollständig oder ungenau sein. Zum Beispiel kann der Roboter Entfernungssensoren und Kameras haben, die ihm Hinweise auf Objekte oder Hindernisse in seiner Umgebung geben, aber diese Informationen sind nicht ausreichend, um genau zu bestimmen, in welchem Raum er sich befindet oder welches Hindernis genau vor ihm ist.

Der Roboter muss in dieser unsicheren Umgebung Entscheidungen treffen, wie er sich bewegt, um ein Ziel zu erreichen. Dies kann bedeuten, dass er auf Basis der unsicheren Informationen Schätzungen über seinen aktuellen Zustand erstellen und dann Entscheidungen darüber treffen muss, in welche Richtung er gehen sollte, um sein Ziel zu erreichen. Dies ist ein klassisches Beispiel für die Anwendung eines POMDPs, da der Roboter Entscheidungen auf der Grundlage unvollständiger Beobachtungen in einer unsicheren Umgebung treffen muss.

Teil 2: Markov-Ketten

- Markovprozeß mit diskretem Zustandsraum (endlich oder abzählbar)
- bei autonomen Systemen mit vollständig beobachtbaren Zuständen
- Ziel: Wahrscheinlichkeiten für das Eintreten zukünftiger Ereignisse

- Markov Ketten 1. Ordnung:
  - der zukünftige Zustand ist nur vom aktuellen Zustand bedingt
  - diese Beziehung ist in der Praxis meistens auch am stärksten
  - homogen, wenn alle Übergangswahrscheinlichkeiten unabhängig vom betrachteten Zeitpunkt sind

- Markov Ketten n-ter Ordnung:
  - der zukünftige Zustand ist von n vorherigen Zuständen abhängig

Frage: Was ist die Markov-Kette und wofür wird Sie benutzt?

Eine Markov-Kette ist eine mathematische **Darstellung eines Markov-Prozesses**, bei dem die Zustände in einem **diskreten Zustandsraum modelliert werden**. Dieser Zustandsraum kann endlich sein, was bedeutet, dass es eine begrenzte Anzahl von Zuständen gibt, oder abzählbar, was bedeutet, dass es eine unendliche, aber zählbare Anzahl von Zuständen gibt.

Um das zu erklären, hier ist ein Beispiel:
Angenommen, Sie modellieren das Wetter mit einer Markov-Kette. In diesem Fall könnte der Zustandsraum endlich sein, da Sie das Wetter in "sonnig", "bewölkt" und "regnerisch" einteilen könnten. Oder Sie könnten den Zustandsraum als abzählbar betrachten, indem Sie das Wetter in "sonnig", "teilweise bewölkt", "stark bewölkt", "leichter Regen", "starker Regen" und so weiter unterteilen. Jeder Zustand repräsentiert eine bestimmte Art von Wetter.
In einer Markov-Kette basieren die Übergänge zwischen den Zuständen auf Wahrscheinlichkeiten, die angeben, mit welcher Wahrscheinlichkeit das Wetter von einem Zustand in einen anderen wechselt. Zum Beispiel könnte es eine hohe Wahrscheinlichkeit geben, dass ein sonniger Tag zu einem bewölkten Tag wird, während die Wahrscheinlichkeit eines plötzlichen Übergangs zu einem Regentag geringer ist.
Markov-Ketten werden in verschiedenen Anwendungen eingesetzt, um stochastische Prozesse zu modellieren, bei denen die zukünftige Entwicklung allein vom aktuellen Zustand abhängt.

Das Ziel ist die Vorhersage des Eintretens eines zukünftigen Ereignisses, anhand des aktuellen Zustandes und der Übergangswahrscheinlichkeiten für den jeweiligen Zustand.

Frage: Was sind Markov-Ketten erster Ordnung?

Wir schauen uns Markov-Ketten erster Ordnung an. Das bedeutet, dass der aktuelle Zustand alleine über den Zukünftigen Zustand entscheidet. Markov-Ketten erster Ordnung sind Markov-Ketten, bei denen die

Übergangswahrscheinlichkeiten zwischen den Zuständen **nur von ihrem direkten vorherigen Zustand abhängen**. Diese **Beziehung ist in der Praxis auch am stärksten**.

Angenommen, wir modellieren das Verhalten eines Studenten in einem Klassenzimmer, der entweder aufmerksam oder unaufmerksam sein kann. Der aktuelle Zustand des Schülers (aufmerksam oder unaufmerksam) beeinflusst, wie wahrscheinlich es ist, dass er im nächsten Augenblick aufmerksam oder unaufmerksam bleibt. Wenn der Schüler aufmerksam ist, könnte die Wahrscheinlichkeit, dass er im nächsten Moment auch aufmerksam ist, beispielsweise 80% betragen, und die Wahrscheinlichkeit, dass er unaufmerksam wird, könnte 20% betragen. Wenn der Schüler jedoch unaufmerksam ist, könnte die Wahrscheinlichkeit, dass er unaufmerksam bleibt, 70% betragen, während die Wahrscheinlichkeit, dass er aufmerksam wird, 30% beträgt.

In diesem Fall handelt es sich um eine Markov-Kette erster Ordnung, da der aktuelle Zustand des Schülers (aufmerksam oder unaufmerksam) allein darüber entscheidet, wie wahrscheinlich es ist, welchen Zustand er im nächsten Moment haben wird. Die Übergangswahrscheinlichkeiten hängen nur vom direkten vorherigen Zustand ab, und es handelt sich um ein einfaches Beispiel für eine Markov-Kette erster Ordnung.

Frage: Wann ist eine Markov-Kette homogen?

Eine Markov-Kette ist homogen, wenn die Übergangswahrscheinlichkeiten zwischen den Zuständen im Laufe der Zeit konstant bleiben. Das bedeutet, dass die Wahrscheinlichkeit, von einem Zustand in einen anderen zu wechseln, u**nabhängig von der Zeit ist**.

Beispiel für eine nicht homogene Markov-Kette:

Nehmen wir an, wir modellieren das Wachstum einer Pflanze über die Zeit. Die Pflanze kann in verschiedenen Entwicklungsstadien sein, und wir betrachten, wie sie von einem Stadium zum nächsten übergeht. Die Übergangswahrscheinlichkeiten zwischen den Stadien hängen von der Jahreszeit ab. Im Frühling hat die Pflanze eine höhere Wahrscheinlichkeit, in das nächste Stadium überzugehen, während die Übergangswahrscheinlichkeiten im Sommer, Herbst und Winter geringer sind. In diesem Fall ist die Markov-Kette nicht homogen, da die Übergangswahrscheinlichkeiten von der Zeit, d.h., der Jahreszeit, abhängen.

Beispiel für eine homogene Markov-Kette:

Betrachten wir eine einfache Münzwurf-Markov-Kette. Sie befinden sich in einem Zustand, der entweder "Kopf" oder "Zahl" ist, abhängig von einer Münze, die Sie werfen. Wenn Sie in "Kopf" sind, werfen Sie die Münze, und abhängig von ihrem Ergebnis wechseln Sie entweder zu "Kopf" oder "Zahl". Die Übergangswahrscheinlichkeiten sind für beide Zustände gleich: Es gibt eine 50%ige Wahrscheinlichkeit, bei "Kopf" zu bleiben, und eine 50%ige Wahrscheinlichkeit, zu "Zahl" zu wechseln. Dies ist ein Beispiel für eine homogene Markov-Kette, da die Übergangswahrscheinlichkeiten konstant und unabhängig von der Zeit oder dem Zustand sind.

Frage: Was sind Markov-Ketten n-ter Ordnung?

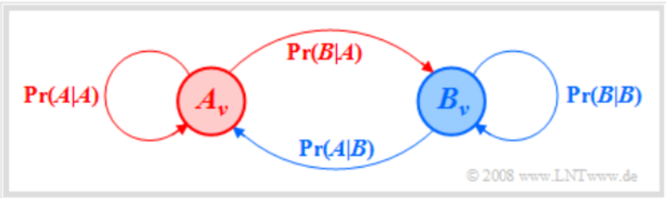
Eine Markov-Kette höherer Ordnung (n-te Ordnung) bezieht sich auf eine Markov-Kette, bei der die Übergangswahrscheinlichkeiten nicht nur von einem vorherigen Zustand abhängen, sondern von den letzten n Zuständen. In diesem Fall sind die zukünftigen Zustandsübergänge nicht nur von einem vorherigen Schritt, sondern von den letzten n Schritten abhängig. **Wir schauen uns also hier mal nicht nur den aktuellste Stand an, sondern auch n-Anzahl Zustände aus der Vergangenheit**.

Ein Beispiel für eine Markov-Kette zweiter Ordnung (eine Markov-Kette, bei der die Übergangswahrscheinlichkeiten von den letzten beiden Zuständen abhängen) könnte das Verhalten eines Roboters sein, der sich in einem Gitter bewegt. Der Roboter hat vier mögliche Bewegungsrichtungen: nach oben, unten, links und rechts. Die Wahrscheinlichkeit, in welche Richtung er sich als nächstes bewegt, hängt jedoch nicht nur von seinem aktuellen Standort ab, sondern auch von seinem vorherigen Standort.

Wenn sich der Roboter beispielsweise in der Nähe einer Wand befindet und sich zuvor von der Wand weg bewegt hat, ist es wahrscheinlicher, dass er sich weiterhin von der Wand weg bewegt. Das bedeutet, die Wahrscheinlichkeit für seine nächste Bewegung hängt von den beiden vorherigen Bewegungen ab. In diesem Fall handelt es sich um eine Markov-Kette zweiter Ordnung.

Wir beschäftigen uns in dieser Vorlesung nur mit erster Ordnung. Wir haben also nur einen Zustandsübergang jeweils.

Homogene Markov-Kette 1. Ordnung (Beispiel mit 2 Zuständen):



ν ist der betrachtete Zeitpunkt (Zukunft)
Pr == P

Die absoluten Zustandswahrscheinlichkeiten lauten:

$\boldsymbol{P(A_\nu)} = P(A|A) * P(A_{\nu-1}) + P(A|B) * P(B_{\nu-1})$ — $\boldsymbol{A_\nu}$ *kann aus* $A_{\nu-1}$ *oder* $B_{\nu-1}$ *folgen*

$\boldsymbol{P(B_\nu)} = P(B|B) * P(B_{\nu-1}) + P(B|A) * P(A_{\nu-1})$ — $\boldsymbol{B_\nu}$ *kann aus* $B_{\nu-1}$ *oder* $A_{\nu-1}$ *folgen*

Die bedingten Übergangswahrscheinlichkeiten lauten:

$\boldsymbol{P(A|A)} = P(A_\nu|A_{\nu-1})$ $\quad$ $\boldsymbol{P(A|B)} = P(A_\nu|B_{\nu-1})$

$\boldsymbol{P(B|B)} = P(B_\nu|B_{\nu-1})$ $\quad$ $\boldsymbol{P(B|A)} = P(B_\nu|A_{\nu-1})$

Die Summe der abgehenden Pfeile eines Ereignisses ist 1!
Dies gilt nicht für die Summe der ankommenden Pfeile.

Frage: Wie berechne ich die zukünftige Wahrscheinlichkeit für Zustand A? Was muss ich dafür in die Formel einbeziehen?

Das Ganze schauen wir uns in einem Beispiel mit zwei Zuständen an. Wir haben den Zustand A und den Zustand B. Um die Wahrscheinlichkeit zu berechnen, dass ein Zustand A in der Zukunft eintreten wird, benötigen Sie zwei Schlüsselelemente:

Sie müssen die Wahrscheinlichkeit kennen, dass sich das System in der Gegenwart im Zustand A befindet.

Sie müssen die Wahrscheinlichkeit kennen, dass, wenn das System sich in der Gegenwart im Zustand A befindet, es auch im zukünftigen Zustand A verbleibt (Wahrscheinlichkeit, dass aus dem gegenwärtigen Zustand A auch der zukünftige Zustand A wird) .

Diese beiden Wahrscheinlichkeiten multiplizieren Sie miteinander, um die gesamte Wahrscheinlichkeit zu erhalten, dass der Zustand A in der Zukunft eintreten wird. Dies ergibt sich aus der Bedingten Wahrscheinlichkeit, bei der Sie die Wahrscheinlichkeit von A in der Gegenwart mit der bedingten Wahrscheinlichkeit von A in der Zukunft gegeben A in der Gegenwart p(A|A) multiplizieren.

Dann gibt es aber noch die zweite Möglichkeit, die Sie dazu addieren müssen. Die gegenwärtige Wahrscheinlichkeit für B multipliziert mit der Wahrscheinlichkeit, dass aus dem gegenwärtigen Zustand B der Zustand A wird.

Wenn ich beide Möglichkeiten addiere, kriege ich die Wahrscheinlichkeit für A in der Zukunft.

Um die Wahrscheinlichkeit für das Eintreten von B in der Zukunft in einem Markov-Ketten-Kontext zu berechnen, sind die folgenden Schritte erforderlich:

Sie benötigen die bedingte Wahrscheinlichkeit, dass B aus A resultiert, symbolisiert als P(B|A).
Sie müssen die Wahrscheinlichkeit kennen, dass sich das System in der gegenwärtigen Situation A befindet, P(A).
Sie benötigen auch die Wahrscheinlichkeit, dass sich das System in der gegenwärtigen Situation B befindet, P(B).
Schließlich benötigen Sie die bedingte Wahrscheinlichkeit, dass B aus B resultiert, symbolisiert als P(B|B).
Die Gesamtwahrscheinlichkeit für das Eintreten von B in der Zukunft wird dann berechnet, indem Sie die Wahrscheinlichkeit in der Gegenwart, P(A) und P(B), mit den bedingten Wahrscheinlichkeiten P(B|A) und P(B|B) multiplizieren, um die entsprechenden Beiträge zum zukünftigen Eintreten von B zu berücksichtigen.

Ich habe also jedes mal den aktuellen Zustand multipliziert mit der Übergangswahrscheinlichkeit.

Frage: Was sind die bedingten Wahrscheinlichkeiten bei Markov? Was zeigen Sie?

Wir haben hier also auch bedingte Wahrscheinlichkeiten und zwar die bedingten Wahrscheinlichkeiten, dass au einem Zustand in der Gegenwart derselbe oder ein anderer Zustand in der Zukunft eintritt.

Die Bedingten Wahrscheinlichkeiten:

P (***A*** |***A***) = *P* ($A_v|A_{v-1}$): Das bedeutet, dass dies die Wahrscheinlichkeit ist, dass aus dem gegenwärtigen Zustand A auch in der Zukunft der Zustand A wird. Dies impliziert, dass der Zustand A im Zeitpunkt v von seinem vorherigen Zustand A im Zeitpunkt v-1 abhängt.

P (***A***|***B***) = *P*($A_v|B_{v-1}$): Das bedeutet, dass dies die Wahrscheinlichkeit ist, dass aus dem gegenwärtigen Zustand B der Zustand A wird. Dies impliziert, dass der Übergang von B zu A im Zeitpunkt v stattfindet, wobei der vorherige Zustand B im Zeitpunkt v-1 berücksichtigt wird.

P (***B*** |***B***) = *P*($B_v|B_{v-1}$): Dies ist die Wahrscheinlichkeit, dass aus dem gegenwärtigen Zustand B auch in der Zukunft der Zustand B wird. Dies bedeutet, dass der Zustand B im Zeitpunkt v nur von seinem vorherigen Zustand B im Zeitpunkt v-1 abhängt.

P (***B***|***A***) = *P* ($B_v|A_{v-1}$): Das bedeutet, dass dies die Wahrscheinlichkeit ist, dass aus dem gegenwärtigen Zustand A der Zustand B wird. Dies impliziert, dass der Übergang von A zu B im Zeitpunkt v stattfindet, wobei der vorherige Zustand A im Zeitpunkt v-1 berücksichtigt wird.

Frage: Warum beträgt die Gesamtwahrscheinlichkeit aller abgehenden Ereignisse eines Zustands in einem Markov-Prozess immer 1 und wie hängt dies mit den möglichen zukünftigen Zustandsübergängen zusammen?

Die Gesamtwahrscheinlichkeit aller möglichen Ereignisse, die von einem Zustand ausgehen, beträgt immer 1. Dies ergibt sich aus der Logik. Wenn ein Zustand vorliegt und es zwei mögliche Wege gibt, wie sich dieser Zustand in der Zukunft entwickeln kann, dann ergibt die Summe der Wahrscheinlichkeiten beider Wege natürlich 100 %.

In einem Markov-Modell repräsentieren die "abgehenden Pfeile" die Wahrscheinlichkeiten für Zustandsübergänge von einem gegebenen Zustand zu anderen möglichen Zuständen. Diese Pfeile zeigen an, wohin der Prozess von einem bestimmten Zustand aus gehen kann.

Die "ankommenden Pfeile" oder "einkommenden Pfeile" hingegen sind die Wahrscheinlichkeiten, wie ein bestimmter Zustand erreicht wurde. Diese Pfeile zeigen an, von welchen vorherigen Zuständen der aktuelle Zustand erreicht werden kann.

F**rage: Sind die ankommenden Pfeile in der Summe auch 1? Begründe**.

Die Summe der ankommenden Pfeile zu einem bestimmten Zustand ist normalerweise nicht auf 1 festgelegt, weil ein Zustand von verschiedenen vorherigen Zuständen erreicht werden kann, und die Wahrscheinlichkeiten für diese Übergänge können unterschiedlich sein.

Die Wahrscheinlichkeit, einen bestimmten Zustand zu erreichen, hängt von den Wahrscheinlichkeiten der vorherigen Zustandsübergänge ab. Wenn es mehrere vorherige Zustände gibt, die zu einem gegebenen Zustand führen können, dann können die Wahrscheinlichkeiten dieser vorherigen Zustandsübergänge unterschiedlich sein. Daher addieren sich die Wahrscheinlichkeiten der ankommenden Pfeile nicht notwendigerweise zu 1, da sie die Wahrscheinlichkeiten aus verschiedenen vorherigen Zuständen repräsentieren.

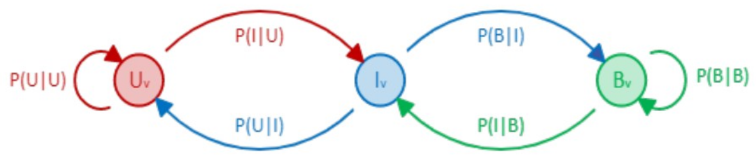
In einem Markov-Modell liegt die Bedingung, dass die Summe der abgehenden Pfeile eines Zustands immer 1 ergibt, weil ein Zustand immer in einen anderen Zustand übergeht, und die Wahrscheinlichkeiten der Übergänge alle möglichen zukünftigen Zustände abdecken müssen. Aber die ankommenden Pfeile repräsentieren die Wahrscheinlichkeiten, wie ein Zustand aus verschiedenen vorherigen Zuständen erreicht wird, und diese können variieren.

Ein Beispiel, um den Unterschied zu verdeutlichen: Angenommen, wir haben einen Zustand C und zwei mögliche vorherige Zustände A und B. Die ankommenden Pfeile für C sind:

Von Zustand A mit einer Wahrscheinlichkeit von 0,3.
Von Zustand B mit einer Wahrscheinlichkeit von 0,5.
Die Summe der ankommenden Pfeile für C beträgt 0,3 + 0,5 = 0,8. Dies zeigt, dass die Summe der ankommenden Pfeile nicht 1 beträgt, da Zustand C aus verschiedenen vorherigen Zuständen mit unterschiedlichen Wahrscheinlichkeiten erreicht werden kann.

Beispiel (mit 3 Zuständen):

Eine Person ist entweder: ungeimpft [hat keinen Impfschutz (mehr)], geimpft [hat Impfschutz] oder geboostert [Impfschutz aufgefrischt].
Der zukünftige Impfstatus ist bedingt vom aktuellen Impfstatus.



Die absoluten Zustandswahrscheinlichkeiten lauten:

$\boldsymbol{P(U_v)} = P(U|U) * P(U_{v-1}) + P(U|I) * P(I_{v-1})$ — $\boldsymbol{U_v}$ *kann aus* U_{v-1} *oder* I_{v-1} *folgen*

$\boldsymbol{P(I_v)} = P(I|U) * P(U_{v-1}) + P(I|B) * P(B_{v-1})$ — $\boldsymbol{I_v}$ *kann aus* U_{v-1} *oder* B_{v-1} *folgen*

$\boldsymbol{P(B_v)} = P(B|B) * P(B_{v-1}) + P(B|I) * P(I_{v-1})$ — $\boldsymbol{B_v}$ *kann aus* B_{v-1} *oder* I_{v-1} *folgen*

Die bedingten Übergangswahrscheinlichkeiten lauten:

$\boldsymbol{P(U|U)} = P(U_v|U_{v-1})$ $\quad \boldsymbol{P(U|I)} = P(U_v|I_{v-1})$ $\quad \boldsymbol{P(I|U)} = P(I_v|U_{v-1})$

$\boldsymbol{P(B|B)} = P(B_v|B_{v-1})$ $\quad \boldsymbol{P(B|I)} = P(B_v|I_{v-1})$ $\quad \boldsymbol{P(I|B)} = P(I_v|B_{v-1})$

Betrachten wir das anhand eines Beispiels: Angenommen, wir haben eine Gruppe von Personen, die sich in drei Zuständen befinden können: ungeimpft, geimpft oder geboostert. Diese Zustände werden durch eine Markov-Kette modelliert. In diesem Modell sind die Zustandsübergänge wie folgt festgelegt:

1. Aus einem ungeimpften Zustand kann entweder der ungeimpfte Zustand beibehalten werden oder in einen geimpften Zustand übergegangen werden.

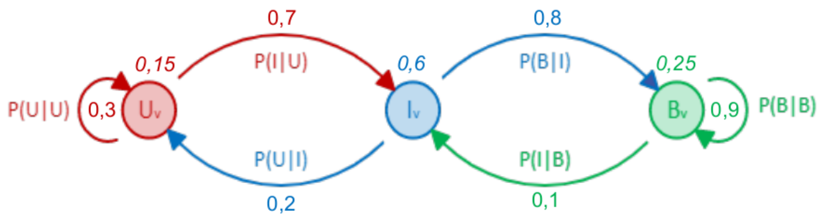
2. Aus einem geimpften Zustand kann entweder in einen geboosterten Zustand übergegangen werden oder in einen ungeimpften Zustand (was den Impfschutz verlieren bedeutet).

3. Aus einem geboosterten Zustand kann entweder der geboosterte Zustand beibehalten werden oder es kann in den normalen Impfschutz gewechselt werden. Dies geschieht, wenn zu viel Zeit seit der Booster-Impfung vergangen ist.

Es ist wichtig zu beachten, dass in diesem Modell direkte Übergänge vom ungeimpften Zustand in den geboosterten Zustand und vom geboosterten Zustand in den ungeimpften Zustand nicht vorgesehen sind. Dies bedeutet, dass es keine direkten Wechsel zwischen diesen beiden Zuständen gibt.

Wir setzen nun die Zustände in die zuvor besprochene Formel ein. Die Wahrscheinlichkeit, dass in der Zukunft ein ungeimpfter Zustand vorliegt (P(Uv)), ergibt sich aus der Wahrscheinlichkeit, dass in der aktuellen Gegenwart ein ungeimpfter Zustand vorliegt (P(Uv-1)), multipliziert mit der Wahrscheinlichkeit, dass der gegenwärtige ungeimpfte Zustand in einen ungeimpften Zustand in der Zukunft übergeht (P(U|U)). Dies wird anschließend zur Wahrscheinlichkeit addiert, dass in der Gegenwart ein geimpfter Zustand vorliegt (P(Iv-1)), multipliziert mit der Wahrscheinlichkeit, dass der gegenwärtige geimpfte Zustand in einen ungeimpften Zustand in der Zukunft übergeht (P(U|I)).
Das macht man auch für die anderen Zustände.

Beispiel (mit 3 Zuständen):



Der aktuelle Zustand:
- *15%* ungeimpft
- *60%* geimpft
- *25%* geboostert

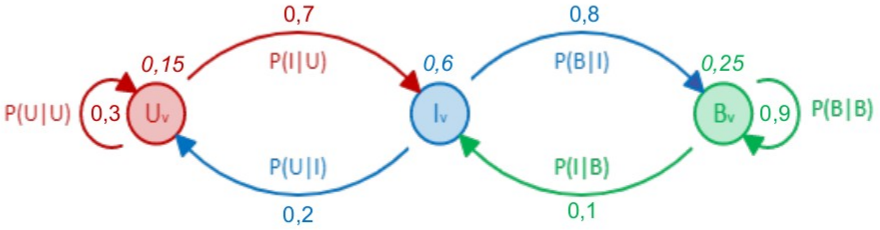
Die bedingten Übergangswahrscheinlichkeiten (dieses Beispiel):

- eine ungeimpfte Person bleibt mit 30%er Wahrscheinlichkeit ungeimpft
- eine ungeimpfte Person lässt sich mit 70%er Wahrscheinlichkeit impfen
- eine geimpfte Person lässt sich mit 80%er Wahrscheinlichkeit boostern
- eine geimpfte Person verliert mit 20%er Wahrscheinlichkeit ihren Impfschutz
- eine Person mit Booster frischt mit 90%er Wahrscheinlichkeit den Impfschutz auf
- eine Person mit Booster frischt mit 10%er Wahrscheinlichkeit den Impfschutz nicht auf

Zur Berechnung der Wahrscheinlichkeiten des zukünftigen Zustandes wird nur der aktuelle Zustand herangezogen. Die Zustände davor interessieren nicht.

Die folgende Beispiel verdeutlicht das Beispiel mit der Impfung mit Zahlen.

Beispiel (mit 3 Zuständen):



Der aktuelle Zustand:
- *15%* ungeimpft
- *60%* geimpft
- *25%* geboostert

Berechnung der nächsten Zustandswahrscheinlichkeiten:

$\boldsymbol{P(U_v)} = 0{,}3 * 0{,}15 + 0{,}2 * 0{,}6 = 0{,}045 + 0{,}12 = \mathbf{0{,}165}$ 16,5% ungeimpft

$\boldsymbol{P(I_v)} = 0{,}7 * 0{,}15 + 0{,}1 * 0{,}25 = 0{,}105 + 0{,}025 = \mathbf{0{,}130}$ 13,0% geimpft

$\boldsymbol{P(B_v)} = 0{,}9 * 0{,}25 + 0{,}8 * 0{,}6 = 0{,}225 + 0{,}48 = \mathbf{0{,}705}$ 70,5% geboostert

Die bedingten Übergangswahrscheinlichkeiten lauten:

$\boldsymbol{P(U|U)} = 0{,}3$ $\boldsymbol{P(U|I)} = 0{,}2$ $\boldsymbol{P(I|U)} = 0{,}7$

$\boldsymbol{P(B|B)} = 0{,}9$ $\boldsymbol{P(B|I)} = 0{,}8$ $\boldsymbol{P(I|B)} = 0{,}1$

Sollen die darauffolgenden Zustandswahrscheinlichkeiten berechnet werden, wird $P(U_v)$ als $P(U_{v-1})$ gesetzt, $P(I_v)$ als $P(I_{v-1})$ und $P(B_v)$ als $P(B_{v-1})$.

Wenn wir die Zahlen in die Formeln einsetzen, erhalten wir folgende Werte. Man sieht, dass die Summe der ausgehenden Pfeile 1 ergibt: z.B. ergibt ungeimpft mit 30 % und 70 % 100 % als Summe. Die ankommenden Pfeile sind aber nicht in der Summe 100 %. Das sieht man, wenn man sich die ankommenden Pfeile von den Impfungen Iv anschaut. Man hat zwei Pfeile – einer ergibt 0,1 und der andere 0,7 = 0.8 insgesamt.

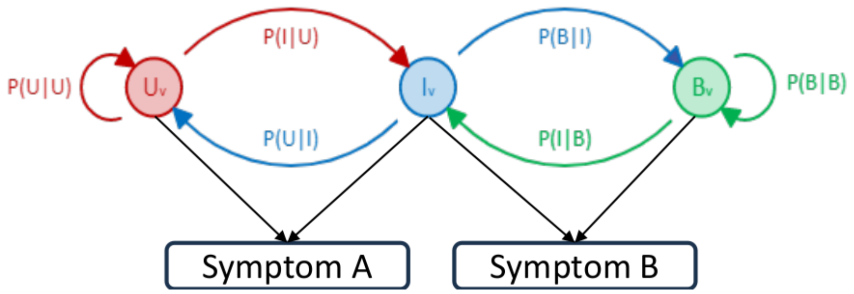
Wenn wir die Wahrscheinlichkeit für zukünftige Zustände berechnen möchten, müssen wir die gegenwärtigen Zustände verwenden, um die zukünftigen Zustände abzuleiten. Dies bedeutet, dass die Wahrscheinlichkeit, einen bestimmten Zustand in der Zukunft zu haben, von den Wahrscheinlichkeiten der gegenwärtigen Zustände und deren Übergänge in diesen Zustand abhängt. Die gegenwärtigen Zustände dienen als Ausgangspunkt für die Vorhersage, wie sich das System im Laufe der Zeit entwickeln wird.
Wenn ich also noch weiter in die Zukunft rechnen möchte, muss ich aus den „Zukünftigen Zuständen“ die gegenwärtige Zustände machen, um noch weiter in die Zukunft zu rechnen.
> So kann man Schrittweise in die Zukunft rechnen, wobei man nur die Zustände aus der Gegenwart betrachtet. Zukünftige Ereignisse werden dabei bei weiteren Schritten zu gegenwärtigen Zuständen.

Die Vergangenheit ist zwar implizit mit dabei, weil die gegenwärtigen Wahrscheinlichkeiten auch aus etwas resultiert sind, aber explizit (das heißt, in der Formel) tauchen nur die Wahrscheinlichkeit für die Gegenwart auf und ihre Übergangswahrscheinlichkeiten.

Teil 3: Markov Modell: Verstecktes Markov-Modell

- „Hidden Markov Model“
- Markov-Kette mit nur teilweise beobachtbaren Zuständen
- es gibt also bekannte (beobachtete) Zustände sowie versteckte
- die Übergangswahrscheinlichkeiten sind bekannt

<u>Beispiel:</u>
Bei Probanden werden bei Krankheit 2 Symptome beobachtet (Symptom A und B).
Es ist bekannt, welche Symptome je nach Impfstatus bei Erkrankung auftreten können.
Es ist jedoch nicht bekannt, welchen Impfstatus die Probanden haben.
Dies kann aus dem vorhandenen Wissen geschlossen werden.



Frage: Wann nutzt man Markov-Ketten und wann Hidden Markov?

Antwort: Markov-Ketten nutzt man, wenn alle Zustände bekannt sind. Hidden Markov nutzt man, wenn nicht alle Zustände bekannt sind.

Lange Antwort:

Markov-Ketten:

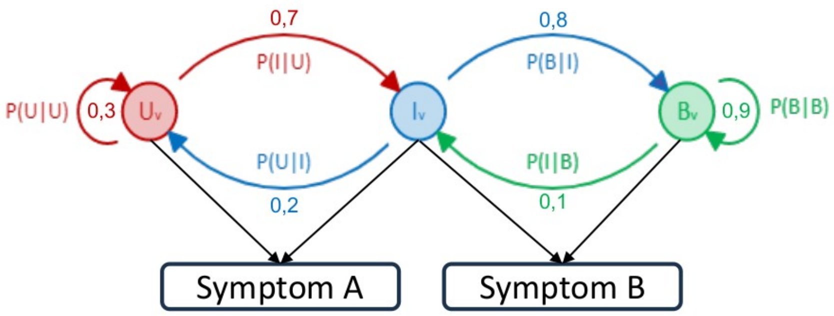
Sichtbare Zustände: In Markov-Ketten sind alle Zustände sichtbar und beobachtbar. Das bedeutet, dass Sie den aktuellen Zustand des Systems direkt beobachten können.
Anwendungsbereiche: Markov-Ketten werden häufig in Problemen verwendet, bei denen die Zustände offen sichtbar und die Übergangswahrscheinlichkeiten zwischen den Zuständen bekannt sind. Beispiele sind Textverarbeitung (z. B. Vorhersage des nächsten Wortes in einem Text), Sprachverarbeitung, Finanzmodelle und Simulationen.

Hidden Markov-Modelle (HMMs):

Verborgene Zustände: Im Gegensatz dazu sind bei HMMs die Zustände nicht direkt beobachtbar. Stattdessen werden Beobachtungen generiert, die von den versteckten Zuständen abhängen, und Sie müssen die versteckten Zustände aufgrund der Beobachtungen schätzen oder vorhersagen.
Anwendungsbereiche: HMMs werden in Situationen eingesetzt, in denen Sie den wahren Zustand des Systems nicht direkt beobachten können, sondern nur Beobachtungen erhalten. Typische Anwendungsfälle sind Spracherkennung (bei der Sie den gesprochenen Text aus den akustischen Signalen erkennen), maschinelles Lernen (z. B. in Natural Language Processing und Bioinformatik), sowie in der Kryptografie und in der Zeitreihenanalyse (zur Vorhersage von Trends und Muster in Daten).

Wenn wir also nur teilweise beobachtbare Zustände haben, dann verwenden wir Hidden Markov. Schauen wir uns dazu ein Beispiel an. **Angenommen wir wissen gar nicht, ob eine Person geimpft, ungeimpft oder geboostert ist**. Wir können nur die Symptome der Person beobachten. (Voraussetzung:)Das heißt, es ist **bekannt, wie die beobachtbaren Symptome (beobachtbaren Zustände) mit unseren Hidden Zuständen zusammenhängen**. Wir wissen die Übergangswahrscheinlichkeit zwischen allen Zuständen.

Ziel: **Anhand dieses Wissen können wir über die beobachtbaren Zustände auf die versteckten Zustände schließen**.



Die Summe der bedingten Übergangswahrscheinlichkeiten ist:
0,3+0,7+0,2+0,8+0,1+0,9 = 3.

(hier auch logisch, da 3x1=3)

Berechnung der Zustandswahrscheinlichkeiten (nur durch Übergänge):

$P(U_v)$ = 0,3 + 0,2 / 3 = $0{,}1\overline{6}$ ≈ 16% entweder durch $P(U|U)$ oder $P(U|I)$

$P(I_v)$ = 0,7 + 0,1 / 3 = $0{,}2\overline{6}$ ≈ 26% entweder durch $P(I|U)$ oder $P(I|B)$

$P(B_v)$ = 0,8 + 0,9 / 3 = $0{,}5\overline{6}$ ≈ 56% entweder durch $P(B|I)$ oder $P(B|B)$

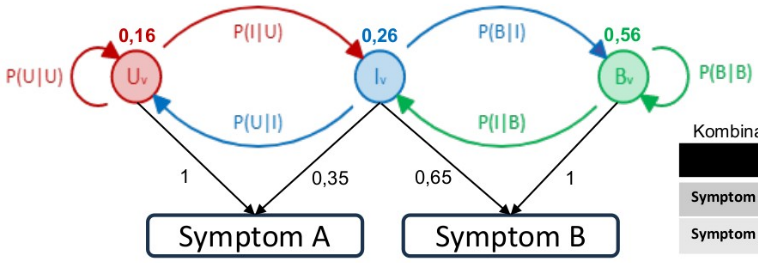
Die allgemeine Wahrscheinlichkeit $P(U_v)$, dass keine Impfung vorliegt ist **16%**.
Die allgemeine Wahrscheinlichkeit $P(I_v)$, dass eine Impfung vorliegt ist **26%**.
Die allgemeine Wahrscheinlichkeit $P(B_v)$, dass ein Booster vorliegt ist **56%**.

Lassen Sie uns dies näher untersuchen. Wir haben drei Knoten, von denen wir die Übergangswahrscheinlichkeiten kennen. Wie bereits erwähnt, sollte die Summe der ausgehenden Pfeile für jeden Knoten 1 ergeben. Insgesamt, wenn wir die Übergangswahrscheinlichkeiten aller Knoten zusammenzählen, ergibt dies 3, da wir drei Knoten haben.

Nun möchte ich die Zustandswahrscheinlichkeiten der versteckten Zustände berechnen, deren genaue Werte mir unbekannt sind. Das **mache ich jetzt nur über die bekannten Übergangswahrscheinlichkeiten**.ich weiß nicht die Wahrscheinlichkeit, mit der ein aktueller Zustand vorliegt.
> **Voraussetzung: Muss die Übergangswahrscheinlichkeiten kennen**.

Ich möchte wissen, mit welcher Wahrscheinlichkeit ein Zustand vorliegt. Dazu schaue ich mir an, wie hoch die Wahrscheinlichkeit ist, dass der Zustand aus sich selbst resultiert. Dann schaue ich mir die Übergangswahrscheinlichkeit an, dass ein anderer Zustand vorher vorlag und das der Zustand aus dem vorherigen Zustand hervorgehen kann. **Diese beiden Wahrscheinlichkeiten addiere ich und dividiere durch die Summe der Übergangswahrscheinlichkeiten**.

In der Abbildung habe ich für ungeimpft 16 %. Das heißt, allgemein ist die Wahrscheinlichkeit, dass ungeimpft vorliegen kann, 16 %. Das mache ich auch für andere Zustände genauso.



Kombination aus den Zuständen und den Wahrscheinlichkeiten:

	ungeimpft	geimpft	geboostert
Symptom A	1*$P(U_v)$ ≈ **16%**	0,35*$P(I_v)$ ≈ **9,1%**	
Symptom B		0,65*$P(I_v)$ ≈ **16,9%**	1*$P(B_v)$ ≈ **56 %**

Wenn ein Proband **Symptom A** aufweist,
liegt die Wahrscheinlichkeit keiner Impfung damit bei $\frac{16\%}{(16\%+9,1\%)}$ ≈ **64%**
und die Wahrscheinlichkeit einer Impfung bei $\frac{9,1\%}{(9,1\%+16\%)}$ ≈ **36%**.

Wenn ein Proband **Symptom B** aufweist,
liegt die Wahrscheinlichkeit einer Impfung damit bei $\frac{16,9\%}{(16,9\%+56\%)}$ ≈ **23%**
und die Wahrscheinlichkeit eines Boosters bei $\frac{56\%}{(56\%+16,9\%)}$ ≈ **77%**.

Es konnte anhand der beobachteten Zustände Symptom A und Symptom B
auf die versteckten Zustände ungeimpft, geimpft und geboostert geschlossen werden.

Dann kenne ich die Wahrscheinlichkeit, mit der aus ungeimpft Symptom A resultiert, wenn Patient ungeimpft ist. Ich kenne auch die Wahrscheinlichkeit für geimpfte Patienten. Weiterhin habe ich dieselben Informationen für Symptom B vorliegen.Wenn ich weiß, dass Symptom A aus dem ungeimpften Status resultiert, kann ich mir den jetzigen Zustand von Ungeimpft anschauen. Der liegt bei 16 %. Da es sich um eine direkte Beziehung handelt, ist die Wahrscheinlichkeit für Symptom A beim ungeimpft-Status auch bei 16 %. Die Wahrscheinlichkeit für geboosterte kann man auch sehr leicht berechnen. Die Wahrscheinlichkeit, dass der geboosterte Zustand auftritt, ist 56 %. Dann ist die Wahrscheinlichkeit für Symptom B beim geboosterten Zustand auch 56 %. Differenzieren muss ich das Ganze bei geimpft, weil geimpft zu 35 % zu Symptom A und zu 65 % zu Symptom B führen kann. Für A muss ich bei geimpft 0,35 multiplizieren mit der Wahrscheinlichkeiten für den geimpften Zustand (26 %). bei Symptom B mache ich das Ganze mit 65 %.
Jetzt möchte ich aber wissen, wenn ein Proband Symptom A aufweist, wie hoch ist die Wahrscheinlichkeit, dass er geimpft oder ungeimpft ist?
Für keine Impfung ergibt sich die Wahrscheinlichkeit für Symptom A aus der Wahrscheinlichkeit, dass keine Impfung bei A vorliegt, geteilt durch die Wahrscheinlichkeit, dass keine Impfung bei Symptom A vorliegt, plus die Wahrscheinlichkeit, dass eine Impfung bei Symptom A vorliegt. Das heißt, ich nutze die Wahrscheinlichkeiten von geimpft und ungeimpft – das sind die zwei Zustände, die zu Symptom A führen können. So kriege ich raus, dass die Wahrscheinlichkeit für keine Impfung bei 64 % liegt bei Symptom A.
Das mache ich analog für keine Impfung und Symptom B. So konnte ich auf die versteckten Zustände schließen.

- Hidden Markov Models — scikit-learn 0.16.1 documentation (sourceforge.net)
- Das Modul Hidden Markov Models in scikit-learn ist deprecated. Das bedeutet, es wird nicht mehr weiterentwickelt und wird in Zukunft aus scikit-learn entfernt werden.

- hmms · PyPI
- GitHub - hmmlearn/hmmlearn: Hidden Markov Models in Python, with scikit-learn like API
- Weitere Bibliotheken für Hidden Markov Models

- oder ohne Einsatz einer speziellen Bibliothek programmieren
- hierzu lassen sich online diverse Lösungen finden…