Grundbegriffe der Entscheidungstheorie

§1. Die Entscheidungstheorie (ET) beruht auf der Grundfrage, was ein Akteur in einer Situation relativ zu seinem Wissen, seinem Wissen über seine Handlungen und ihre Konsequenzen in Anbetracht seiner Präferenzen rationalerweise tun soll. Betont man das "tun soll" spricht man von "normativer" ET. Demgegenüber ist die "empirische" ET mit der These verbunden, Akteure (etwa in der Mikroökonomie) verhielten sich *ceteris paribus* gemäß diesem Modell. Ein Baustein der ET ist die Wahrscheinlichkeitstheorie.[1] Diese dient dazu, die Wahrscheinlichkeit von *Ereignissen* zu modellieren. Dazu kann man (a) von einer nicht weiter definierten Menge von möglichen Welten ausgehen, oder (b) die möglichen Welten im Sinne Carnaps als Aussagenmengen aus einer gegebenen Sprache konstruieren. Im Falle (b) sind Aussagen schon vorgegeben. Im Falle (a) werden Propositionen als Mengen von möglichen Welten eingeführt (den Welten, in denen das, was die Proposition behauptet, der Fall ist). *Ereignisse* (etwas tritt ein, eine Veränderung findet statt, oft aber auch Sachverhalte einschließend verstanden) werden in beiden Ansätzen als Mengen von möglichen Welten definiert. Will man die Abfolge, die in einem Ereignis statt hat, ausdrücken, muss man Ereignisse als geordnete Tupel von möglichen Welten einführen (die erste Welt im Tupel ist der Ausgangszustand, das zweite Element des Tupels der unmittelbar nächste Gesamtzustand, mit dem sich nun evtl. [ist das zweite Element ungleich dem ersten] etwas geändert hat, usw.). Auf die Menge der Ereignisse wird nun eine Wahrscheinlichkeitsverteilung definiert. (Dies betrifft entweder die Potenzmenge der Menge der möglichen Welten, in der ja alle Teilmengen der Menge der möglichen Welten sind, oder eines der 2- bis n-fachen Cartesischen Produkte der Menge der möglichen Welten, in der sich alle 2- bis n-Tupel von möglichen Welten befinden. Welches Cartesische Produkt hierbei gewählt wird, richtet sich danach, wie tief man mit Ereignissen in die Vergangenheit oder Zukunft schauen will/kann.)
Jedem Ereignis wird so eine Wahrscheinlichkeit, dass es der Fall ist, zugewiesen. Werden die Ereignisse durch Aussagen repräsentiert (im Fall (a) entsprechen Mengen von möglichen Welten ja Propositionen), kann sich die Wahrscheinlichkeitsverteilung auf Aussagen beziehen.

---

1 Vgl. Goldberg, *Probability*; Skyrms, *Einführung in die induktive Logik*.

$\{p, q, r, ...\}$ sei eine Menge von Aussagen, wobei sich Aussagen wahrheitsfunktional auf die übliche Weise verbinden lassen. Die Wahrscheinlichkeitsverteilung P erfüllt die Kolmogorov-Axiome:

A1 $P(p) \geq 0$ (Es gibt keine negativen Wahrscheinlichkeiten.)

A2 $\vdash p$ dann $P(p) = 1$ (Tautologien sind maximal wahrscheinlich.)

A3 $\vdash (\neg(p \wedge q))$ dann $P(p \vee q) = P(p) + P(q)$

Axiom 3 besagt, dass im Falle, dass sich zwei Aussagen bzw. Ereignisse ausschließen, die Wahrscheinlichkeit, dass eines von beiden vorliegt, die Summe der Einzelwahrscheinlichkeiten ist. Dies leuchtet ein, da sich zwei sich ausschließende Ereignisse nicht überlappen, also bei der Betrachtung der disjunktiven Wahrscheinlichkeit zwei disjunkte Mengen von möglichen Welten zu vereinigen sind. Aus den Axiomen ergeben sind nun eine Reihe von Grundtheoremen bzw. -regeln:[2]

T1 $P(\neg p) = 1 - P(p)$

(Negationsregel)(aus A3 und A2 mittels "$p \vee \neg p$")

T2 $\vdash (\neg p)$ dann $P(p) = 0$

(Kontradiktionen als minimal wahrscheinlich.)(aus A2, T1)

T3 $\vdash (p \equiv q)$ dann $P(p) = P(q)$

(Äquivalente sind gleich wahrscheinlich.)

T4 $P(p \vee q) = P(p) + P(q) - P(p \wedge q)$

(Generelle Disjunktionsregel)

Die bedingte Wahrscheinlichkeit $P(q/p)$ ist die Wahrscheinlichkeit, dass q der Fall ist, wenn p der Fall ist.

T5 Ist $P(p) > 0$, so $P(q/p) = P(p \wedge q)/P(p)$

Die bedingte Wahrscheinlichkeit von q gegeben p ist die Wahrscheinlichkeit, dass beide vorliegen relativ zur (d.h. geteilt durch die) Wahrscheinlichkeit von p. q ist von p probabilistisch unabhängig, wenn die Wahrscheinlichkeit von q nicht durch das Vorliegen von p steigt, d.h. wenn $P(q/p) = P(q)$.

T6 Ist q von p probabilistisch unabhängig, so $P(p \wedge q) = P(p) \bullet P(q)$

(Spezielle Konjunktionsregel)

T7 Ist $P(p) > 0$, so $P(p \wedge q) = P(p) \bullet P(q/p)$

(Generelle Konjunktionsregel)

---

2 Vgl. Skyrms, *Einführung in die induktive Logik*, Kap.5.

T8 $\vdash (p \supset q)$, so $P(p) \leq P(q)$

(Prinzip der logischen Folgerung)

T8 besagt, dass eine Konsequenz einer Aussage mindestens so wahrscheinlich ist, wie die Aussage selbst. Diese Regeln treten mit dem Anspruch auf, *unser intuitives Räsonieren mit Wahrscheinlichkeiten zu erfassen.* Das wird dadurch gezeigt, dass eine Person A, die gegen diese Regeln verstößt, das Opfer von Wetten werden kann (sogenannten "Dutch Books"), bei denen eine Person B, die A´s Zumessen von Wahrscheinlichkeiten kennt, A aufgrund von A´s devianten Wahrscheinlichkeitsregeln zum Abschluss von Wetten zwingen kann, bei denen A verlieren muss: Epistemische (d.h. unsere intuitiven) Wahrscheinlichkeiten *sind* mathematische (den Kolmogorov-Axiomen genügende) Wahrscheinlichkeiten.[3] Die Wahrscheinlichkeitstheorie, wie sie hier angewendet wird, betrifft den Meinungsaspekt der Entscheidungssituationen. Eine naheliegende Frage ist, woher wir um die genauen Wahrscheinlichkeitszuschreibungen eines Akteurs wissen. Das Grundmodell, das von idealisierten Akteuren ausgeht, die Geld wertschätzen und jede Mark genau so wertschätzen wie die andere, unabhängig davon, wieviel Geld sie besitzen, sieht dazu zwei Möglichkeiten vor:

Methode 1:

Man fragt oder interpretiert die Akteure bezüglich *komparativer* Urteile (d.h. ob sie p für höchstens so wahrscheinlich halten wie q [und der damit definierbaren anderen komparativen Begriffe; s.u. zu einem solchen Vorgehen]); daraus ergibt sich eine Rangordnung des Für-wahr-Haltens bezüglich der vorgelegten Aussagen und deren logischer Konsequenzen; man kann nun eine beliebige Metrisierung (d.h. Zuweisung quantitativer Wahrscheinlichkeitswerte) wählen, sofern diese die Rangordnung beibehält.[4]

Methode 2:

Man bietet dem Akteur (zumindest als Gedankenexperiment) einen Wettschein zum Kauf an; die Wette besagt: 1DM Gewinn, falls p wahr ist, 0DM sonst. Ist der Akteur bereit, 56 Pfennig für den Wettschein zu bezahlen, aber nicht 57 Pfennig, so ist die von ihm p zugeschriebene Wahrscheinlichkeit $P(p) \leq 0.57$. Zur Modellierung des Bewertungsaspektes der Entscheidungssituation bedarf es einer Präferenzordnung. Die

3 Vgl. Skyrms, *Einführung in die induktive Logik*, S.289-97.

4 Vgl. Jeffrey, *Logik der Entscheidungen*, S.25-41.

einfachste Form der Präferenzordnung wäre wieder eine *komparative* Ordnung: Man fragt oder interpretiert den Akteur, ob er den Sachverhalt p dem Sachverhalt q vorzieht (d.h. zunächst ob q "höchstens so wünschenswert ist wie" p, womit dann die Relation "gleich wünschenswert" definierbar ist [p ist gleichwünschenswert wie q genau dann, wenn p höchstens so wünschenswert ist wie q und q höchstens so wünschenswert ist wie p] sowie die Relation "weniger wünschenswert" [als "p höchstens so wünschenswert und nicht gleich wünschenswert wie q"]). Diese Ordnung muss man dann wieder metrisieren. Alternativ: Man bietet dem Akteur einen Wettschein zum Kauf an: die Wette besagt: p wird realisiert, falls eine Tautologie wahr ist, $\neg$p sonst. Das was der Akteur für diesen Wettschein, der sicher gewinnt, da eine Tautologie immer wahr ist, auszugeben bereit ist drückt seine Wertschätzung von p aus.

Im Prinzip – und das reicht für die Modellbildung der ET – lassen sich Akteuren also quantifizierte Wahrscheinlichkeitsverteilungen und Wünschbarkeitsgrade zuschreiben. Die Wünschbarkeitsfunktion (Bewertungsfunktion) bewertet *in den Grundmodellen* jeden möglichen Zustand mit einem monetären Bewertungsmaß (nicht zu verwechseln mit dem daraus zu definierenden *Nutzen*maß). Die Grundmodelle unterscheiden sich nun bezüglich des einem Akteur zugeschriebenen Weltwissens und der Frage, wie gewiss es ist, welche Konsequenzen eine Handlung hat. Die bekanntesten Grundmodelle, die auch den Ausgangspunkt für Nozicks Überlegungen bilden, sind die folgenden[5]:

(1) In einem ***Savage-Modell*** hat man

(i) eine endliche, nicht-leere Menge W möglicher Welten

(ii) eine endliche, nicht-leere Menge K von Konsequenzzuständen

(iii) eine nicht-leere Menge H von Funktionen von W in K (das sind die Handlungen)

(iv) ein Wahrscheinlichkeitsmaß P auf der Potenzmenge von W (d.h. bezüglich der Ereignisse)

(v) eine Bewertungsfunktion V von C in R (d.h. ein in reellen Zahlen ausgedrückter Situationswert)

5 Vgl. Resnik, *Choices*; Spohn, *Grundlagen der Entscheidungstheorie*.

- die Nutzenfunktion in R ist auf H definiert, sodass für alle $h \in H$,

$$U(h) = \sum_{w \in W} V(h(w)) \bullet P(\{w\})$$

Das heißt: der Nutzen einer Handlung ist die Summe der Werte, die sie relativ zu Ausgangssituationen realisieren kann (vermittels ihrer Konsequenz $k \in K$, wobei $k = h(w)$), *gewichtet* mit der Wahrscheinlichkeit der jeweiligen Ausgangssituation. Von einem *rationalen* Akteur wird nun gefordert, dass er den Nutzen maximiert, d.h. die Handlung wählt, für die U(h) maximal ist. (Oder es wird behauptet – in der empirischen Lesart der ET –, *dass* Akteure, die sich durch ein Savage-Modell beschreiben lassen, den Nutzen maximieren. Das Wahrscheinlichkeitsmaß in (iv) drückt die Wahrscheinlichkeit aus, dass wir uns in einem Zustand w befinden. In w hätte die Handlung die Konsequenzen h(w) mit der Bewertung V(h(w)). Dabei ist, gemäß (iii), *gewiss*, welche Konsequenzen eine Handlung hat. Auch beziehen sich auf Handlungen (bzw. Handlungsoptionen) keine Wahrscheinlichkeiten.

(2) In einem ***Fishburn-Modell*** hat man

(i) eine nicht-leere, endliche Menge K von Konsequenzzuständen

(ii) eine nicht-leere, endliche Menge H von Handlungen

(iii) eine Funktion P auf H, die für jedes $h \in H$ ein Wahrscheinlichkeitsmaß P_h auf der Potenzmenge von K bildet.

(iv) eine Bewertungsfunktion V von K in R

- die Nutzenfunktion von H in R ist definiert, sodass für alle $h \in H$

$$U(h) = \sum_{k \in K} V(k) \bullet P_h(\{k\})$$

Das heißt: Ereignisse werden eingeführt als Mengen von möglichen Welten aus K. Es gibt keine Wahrscheinlichkeitsfunktion bezüglich der Ausgangsweltzustände (wie in einem Savage-Modell). Wir wissen also, in welchem Zustand wir uns befinden. Die Ergebnisse von Handlungen sind *nicht* gewiss (auch abweichend von Savage-Modellen), sondern nur wahrscheinliche Konsequenzen im Maß $P_h(\{k\})$. Die Definition des rationalen Akteurs bzw. die empirische Behauptung entspricht der bei Savage-Modellen.

(3) In einem ***Jeffrey-Modell*** hat man

(i) eine nicht-leere, endliche Menge W von möglichen Welten

(ii) eine Partition H von W

(iii) ein Wahrscheinlichkeitsmaß P auf der Potenzmenge von W

- die Nutzenfunktion ist eine Funktion von Potenzmenge(W)\ { } in R, sodass für $A,B \subset W$ mit $P(A \cap B) = 0$ und nicht $P(A \cup B) = 0$ gilt:

$$U(A \cup B) = U(A) \bullet P(A/A \cup B) + U(B) \bullet P(B/A \cup B)$$

Extrahiert man eine Bewertungsfunktion aus der Nutzenfunktion durch $V(w) = U(\{w\})$, so gilt für die Nutzenfunktion:

$$U(A) = \sum_{w \in A} V(w) \bullet P(w/A)$$

Handlungen werden in Jeffrey-Modellen repräsentiert als Ereignisse, die ebenso wie alle anderen Ereignisse, gemäß (iii), eine Wahrscheinlichkeit ihres Vorliegens oder Eintretens haben. Die Konsequenzen einer Handlung A (als Ereignis eine Menge von möglichen Welten, d.h. eine Teilmenge von W und Element der Potenzmenge von W) sind relativ zu A auch nur wahrscheinlich, und diese Beziehungen werden als bedingte Wahrscheinlichkeiten P(w/A) modelliert. $w \in A$ gilt, da der Konsequenzzustand ein Zustand ist, in dem Handlung A vollzogen wurde (d.h. vorliegt). Dass, gemäß (ii), die Menge der Handlungen H eine Partition von W ist, besagt, dass die Handlungen untereinander disjunkt sind und insgesamt die Möglichkeiten erschöpfen. Dass Handlungen selbst bloß eine Wahrscheinlichkeit haben, kann man so deuten, dass der Akteur nicht *weiß*, in welchem Weltzustand, mit welchen offenstehenden Handlungsoptionen, er sich befindet, oder so dass der Akteur zwar weiß, in welchem Zustand er sich befindet, aber nicht sicher sein kann, ob er die Handlung überhaupt *vollziehen* kann - unabhängig davon, dass, selbst wenn er die Handlung vollziehen kann, die Konsequenzen bloß wahrscheinlich im Maße x eintreten. Die Definition rationaler Akteure entspricht wieder (1) und (2).

(4) In einem ***Luce-Krantz-Modell*** hat man

(i) eine nicht-leere, endliche Menge W von möglichen Welten

(ii) eine nicht-leere, endliche Menge K von Konsequenzzuständen

(iii) eine nicht-leere Menge H von Funktionen in C deren Definitionsbereiche $A \subset W$ sind, wobei für jedes $h_A \in H$, nicht $P(A) = 0$

(iv) für jedes $h_A \in H$ ein $h_B \in H$ mit $h_A \subset h_B$, und mit $h_B \in H$ ist für jedes $A \subset W$, sodass nicht $P(A) = 0$, die Einschränkung von h_B auf A in H

(v) ein Wahrscheinlichkeitsmaß P auf der Potenzmenge von W

- die Nutzenfunktion ist eine Funktion von H in R, so dass: wenn $(h_A \cup g_B) \in H$, wobei $A \cap B = \{\ \}$ und nicht $P(A) = 0$, nicht $P(B) = 0$, dann ist:

$U(h_A \cup g_B) = U(h_A) \bullet P(A/A \cup B) + U(g_B) \bullet P(B/A \cup B)$.

Extrahiert man wieder eine auf dem Cartesischen Produkt $W \times K$ definierte Bewertungsfunktion V, für die gilt:

Wenn es für $w \in W$, wobei nicht $P(w) = 0$, und $k \in K$ ein $h_{\{w\}} \in H$ gibt mit $h_{\{w\}}(w)=k$, dann $V(w,k)=U(h_{\{w\}})$, so gilt für die Nutzenfunktion:

$$U(h_A) = \sum_{w \in A} V(w,h_A(w)) \bullet P(w/A)$$

Offensichtlich sind Luce-Krantz-Modelle schon etwas komplexer. "h_A" steht dafür, dass die Handlung h erfolgt relativ dazu, dass der Akteur erfährt, dass A der Fall ist. Luce-Krantz-Modelle beziehen so die Erweiterung des Situationswissens ein. Handlungen erfolgen, gemäß (iii), unter diesem Zuwachs an Wissen um wahrscheinliche Ereignisse. Die in (iv) erwähnte Einschränkung "auf A" besagt, dass A, das vorher nur eine Option innerhalb der Option B war, nun der Fall ist. Die Konsequenzen von Handlungen sind, gemäß (iii), wieder (anders als in (2) und (3)) gewiss. Auch die Handlungen selbst haben keine Wahrscheinlichkeiten (anders als in (3)). P(w/A) sagt, wie wahrscheinlich *relativ zum gewußten Sachverhalt A* es ist, dass man sich in Welt w befindet. Es handelt sich also nicht (wie in (3)) um die Wahrscheinlichkeit, dass A als Handlung einen anderen Zustand nach sich zieht. Der Nutzen einer Handlung h unter dem Wissen, dass A vorliegt, ist die Summe der Werte aller Konsequenzen dieser Handlung gewichtet damit, welche Wahrscheinlichkeit die Ausgangssituationen für die Handlung gemäß dem Umstand, dass A vorliegt, haben. Rationale Akteure, die durch Luce-Krantz-Modelle beschreibbar sind, maximieren den Nutzen (wie in (1) – (3)).

Was haben wir hier gesehen?

Die verschiedenen Grundmodelle machen recht verschiedene Annahmen über die Entscheidungssituation. Man mag aus theoretischen Gründen eines von ihnen vorziehen, man kann indessen auch *je nach Vorwissen über die anstehenden Entscheidungssituation* eines als das in einer Situation *dieses Typs* angemessenste auswählen. Insofern wiedersprechen sich die Grundmodelle nicht. Der ET stehen alle Modellierungen offen. Sie stimmen alle in der Definition des rationalen Akteurs als eines Nutzenmaximierers überein. Komplexere Modelle entwicklen die Grundmodelle in die naheliegenden Richtungen:

(a) Betrachtung nicht einzelner Entscheidungen, sondern von Abfolgen von Handlungen (d.h. von *Strategien*), wobei zu den jetzt bedachten zukünftigen Situationen, die mit temporallogischen Strukturen zu modellieren sind, auch gemäß der ET strukturierte Entscheidungssituationen gehören;
(b) Betrachtung nicht ganzer weiterer Weltverläufe, sondern Isolation der *kausalen* Konsequenzen einer Handlung durch "small worlds";
(c) Betrachtung nicht eines einzelnen Akteurs, sondern der Abfolge von Entscheidungen von *verschiedenen* Akteuren (in einem Baum aller verschiedenen Handlungsabläufe);
(d) Betrachtung nicht nur einer Abfolge von Handlungen (wie in (a)), sondern Berücksichtigung möglicher zukünftiger *Erfahrungen* (d.h. neuer Wahrscheinlichkeitsverteilungen, z.B. durch Konditionalisierung).[6]

§2. Nozick hat keine ausgearbeitete Theorie zur Integration symbolischen Wertes bzw. symbolischen Nutzens, da er Schwierigkeiten mit dem Nutzenmaximierungsprinzip sieht.[7] Nozick will auch allein symbolischen Nutzen einführen, nicht den symbolischen Wert eines Zustandes.[8] Als erste Näherung zur Einbeziehung des symbolischen Nutzens und eines (analog einführbaren) symbolischen Wertes lassen sich jedoch folgende Überlegungen anstellen:

Zustände werden bewertet. Bei einigen läßt sich direkt ein monetärer Wert zuordnen. Andere Zustände werden zunächst wertgeschätzt in Hinblick auf außermonetäre Werte (wie Identität des Selbst, Zugehörigkeit zu einer Gemeinschaft etc.). Solche Zustände seien Zustände mit "symbolischen Wert" genannt. Handlungen, die *kausal* auf solche

---

6 Vgl. zu (a), (d) und (b): Spohn, *Grundlagen der Entscheidungstheorie*; zu (c) Sen, *Collective Choice and Social Welfare*, s.u.
7 Vgl. Nozick, *The Nature of Rationality*, S.34f.
8 Vgl. ebd. S.48

Zustände bezogen sind, erhalten ihre Nutzenbestimmung auf die übliche Weise. Andere Handlungen, die *symbolisch* auf solche Zustände bezogen sind, die monetären oder symbolischen Wert besitzen, haben "symbolischen Nutzen". Nutzenmaximierung muss nun symbolischen Nutzen einschließen. Entlang der Symbolisierungsrelation muss Nutzenzumessung an die symbolischen Handlungen zurückfließen. Dies kann nur so geschehen, dass symbolischer Nutzen schließlich mit "herkömmlichen" Nutzen verrechnet werden kann. Denn nur so kann es zu einer Handlung (insbesondere in Konfliktfällen) kommen. Zwei Wege lassen sich hier beschreiten:

(a)

Zur Erstellung der Korrelation kann man wieder komparative Werturteile verwenden. Ausgangspunkt der Korrelierung können zwei Situationen s_1 und s_2 sein, wobei s_1 hauptsächlich monetären und s_2 hauptsächlich symbolischen Wert hat. Ist der Akteur zwischen diesen beiden Situationen *indifferent*, müssen sie in einer *kombinierten* Präferenzordnung, die monetäre und symbolische Werte umfaßt, auf derselben Stufe stehen. Auf diese Weise läßt sich insgesamt eine komparative Präferenzordnung, die beide Werttypen umfaßt, aufstellen. Diese läßt sich dann wieder metrisieren, ohne dass die quantifizierten Werte jetzt als monetäre Werte zu verstehen sind. Der rationale Akteur ist wieder ein Nutzenmaximierer. Die Beziehungen, die Handlungen zu Zuständen haben, müssen nun konventionelle (d.h. symbolische) Beziehungen oder *gewisse* Kausalbeziehungen (etwa eines Savage-Modells) sein, um diese Beziehungen einheitlich zu modellieren. Oder die Menge der Handlungen muss entsprechend geteilt und unterschiedlich behandelt werden, so dass auch probabilistische Kausalbeziehungen möglich sind. Die Nutzenfunktionsdefinition wäre entsprechend komplizierter.

(b)

Man könnte auch *getrennte* Entscheidungssituationen für kausale Handlungen relativ zu (monetär) wertvollen Situationen und für symbolisch nutzvolle Handlungen betrachten. Der Aufbau einer symbolischen Präferenzordnung und der entsprechenden Entscheidungssituation könnte einem Savage-Modell folgen und eine entsprechende Handlung als optimal ausgeben. Daneben könnte z.B. ein Jeffrey-Modell eine bezüglich monetären Wertes optimale Handlung ergeben. Verglichen wird dann der monetäre Nutzen von h_2 mit dem symbolischen Nutzen von h_1 (analog den Überlegungen in (1) – irgendeine Weise der Verrechnung muss es geben!). Die so getrennt berechneten Nutzenwerte wären dann in einer allgemeineren Formel der Nutzenmaximierung zu verrechnen. Es bedarf in Nozicks Variante, den symbolischen Nutzen in die Entscheidungstheorie einzubeziehen, einer allgemeineren Formel der Nutzenkalkulation, die symbolischen Nutzen mit Nutzen wie er von der kausalen oder epistemischen Wahrscheinlichkeitstheorie berechnet wird, verrechnet. In diese Formel gehen die Gewichtungen ein, die man der entsprechenden Methode der Nutzenberechnung zumisst. Nozicks Formel[9] lautet:

$$(DV)\ U(h) = P(KET) \bullet U_{KET}(h) + P(EET) \bullet U_{EET}(h) + P(SET) \bullet U_{SET}(h)$$

Der Gesamtnutzen einer Handlung h ergibt sich als Summe des Nutzens wie er gemäß (a) der kausalen Entscheidungstheorie (KET), die auch das *Dominanzprinzip* (die Handlung zu wählen, die in allen alternativen Handlungsverläufen bessere Resultate liefert als die Handlungsalternativen) einschließt, (b) der epistemologischen (EET), die Umstände *wechselseitigen Wissens* einbezieht, und (c) der symbolische Verknüpfung betreffenden Entscheidungstheorie (SET) berechnet wird, gewichtet mit der Wahrscheinlichkeit (dem Grad des Wahrseins), den man der entsprechenden Entscheidungstheorie – eventuell relativ zum betrachteten Fall – zumisst.

Insofern symbolischer Nutzen in die Kalkulation des Nutzens eingeht, können Abweichungen von der herkömmlichen Logik der Nutzenmaximierung auftreten. Dafür wird Handeln nach ethischen Prinzipien oder symbolischen Wertzumessungen innerhalb der Entscheidungstheorie modellierbar. Repräsentiert man epistemologische Maximen mit einem Wert, den sie für das Verfolgen der Wahrheit oder das Handeln besitzen, so können auch Methoden des rationalen Für-wahr-Haltens mit (DV) bewertet werden. Eingeführt werden muss *für diese Weise* der Vermittlung von kausalen und symbolischen Konsequenzen einer Handlung die Gewichtung der verschiedenen entscheidungstheoretischen Methoden. Nozick führt diese Gewichtung im zweiten Kapitel von *The Nature of Rationality* vor anhand unseres – vermeintlich – schwankenden Vertrauens in eine entsprechende entscheidungstheoretische Methodik in verschiedenen Szenarien von *Newcombs Problem* oder des *Gefangenen-Dilemmas.*[10] Wären die entscheidungstheoretischen Methoden selbst ungewichtet (also die [begründete] Wahl einer Methode eine „alles oder nichts"-Entscheidung), so dürften Variationen in Entscheidungsszenarien (insbesondere bezüglich der zu erreichenden Werte und des damit einhergehenden Risikos) nicht zu einer Unsicherheit oder Unzufriedenheit mit dem entsprechenden Räsonement führen, solange dieses immer noch zutrifft. Solche Unsicher-

9 Vgl. ebd.

10 Vgl. zu beiden auch Sainsbury, *Paradoxien*, S.73-102.

heiten treten aber, *obwohl* dieselbe Argumentation – etwa gemäß dem Dominanzprinzip – noch greift, anscheinend auf; im Fall des *Gefangenen-Dilemmas*:

> In general, when the cooperative solution payoffs are very much higher than the dominance ones, and when payoffs for the nonmatching actions offer only slight gains or losses over these two, then we strongly will think that cooperation is rational and will find that the dominance argument has little force. [If] [a]lternatively [...] the cooperation solution is only slightly better than the dominant one, and the extreme values in the payoffs for the nonmatching actions diverge greatly[, and] [w]hen we have no special ties to the other party or particular knowledge of the other party´s probabilities of action, then we will think it is rational to perform the dominant action in [such] a situation, not running any risk that the other party will perform his dominant action, which he has a large incentive to do.[11]

Nozick deutet diese Unsicherheit und das Wechseln der Methode als bedingt durch eine Zuweisung von Glaubwürdigkeit oder Wahrscheinlichkeit an die jeweilige Methode. Andererseits könnte sich hier auch ausdrücken, dass relativ zu Entscheidungssituationen *gemäß der Merkmale dieser Situationen* eine Auswahl einer entscheidungstheoretischen Methode erfolgt, die dann strikt befolgt wird. Je nach Situation würde dann eine Methode verwendet, nicht in jeder Situation jede Methode gewichtet mit einem Grad ihrer Wahrscheinlichkeit. Der von Nozick beschriebene Umgang mit Varianten von Entscheidungssituationen entspricht eher einem solchen Vorgehen, insofern bei einer geringen Ausgangswahrscheinlichkeit der epistemologischen Entscheidungstheorie die Werte im *Gefangenen-Dilemma* extrem umgewichtet werden müssen, um eine kooperative Handlung rational zu machen, während die Schwelle zum Erfüllen der (qualitativen) Kriterien, nicht mehr dem Dominanzprinzip zu folgen, geringer sein könnte.

§3 Soziale Entscheidungstheorien müssen insbesondere soziale (kollektive) Wohlfahrtsfunktionen voraussetzen. Dies ist bei Akteuren mit unterschiedlichen Präferenzordnungen nicht trivial. Zur Vereinfachung des Ausgangspunktes kann man ein Savage-Modell wählen, bei dem bekannt ist, in welcher Situation man sich befindet (d.h. es gibt eine Welt $w \in W$ mit $P(w)=1$) oder man wählt ein Fishburn-Modell, in dem die Konsequenzen der Handlungen gewiss sind (d.h. $P(k/h)=1$ oder $P(k/h)=0$), so dass zunächst die Wahrscheinlichkeiten aus den Überlegungen herausfallen. Soziale Wohlfahrtsfunktionen oder Soziale Entscheidungsfunktionen müssen aus einer Ansammlung individueller Präferenzordnungen eine soziale Präferenzordnung generieren. Die soziale/kollektive Entscheidungstheorie untersucht, welche Bedingungen sich konsistent an eine solche Soziale Entscheidungsfunktion stellen lassen bzw. welche Kombination von

11 Nozick, *The Nature of Rationality*, S.53.

Bedingungen. Zu den Bedingungen, die untersucht werden, gehören z.B.: Vermeidung von Diktatoren (das hieße die Präferenzordnung eines Individuums würde übernommen), Pareto-Bedingungen (wie Realisierung eines Präferenzkonsenses und/oder individuelle Veto-Rechte), Beliebigkeit der Ausgangspräferenzen der Gruppenmitglieder, Stabilität der Sozialen Entscheidungsfunktion unter Hinzufügung irrelevanter Zusatzalternativen, Reflexivität und Konnexität der entstehenden Anordnung, verschiedene Formen der Transitivität, wechselseitige Übertragung von Ausgewähltsein – usw. Die sogenannten "Impossibility" und "Possibility" Theoreme, etwa von Arrow, Harsanyi und Sen, betreffen die Kombinierbarkeit solcher Bedingungen.[12]

12 Vgl. allgemein: Sen, *Collective Choice and Social Welfare*; Resnik, *Choices*.