

## Integrated Information Theorie IIT – auf der Ebene der Mechanismen

### Abstract:

**Über Bewusstsein wurden in den letzten 2,5 Jahrtausenden viele Theorien entwickelt. Einen wesentlichen Meilenstein setzte Rene Descartes mit „Ich denke also bin ich“. Aktuell sind es die von Neurowissenschaftlern vorgestellten Entwürfe. Davon sind drei im Fokus. Die „Global Workspace Theorie GWT“, die „Orchestrated objective reduction (Orch OR)“ und die „Integrated Informations Theorie IIT“. Bei IIT handelt es sich um eine Hypothese, die ausgehend von fünf Axiomen versucht Bewusstsein im generellen nachzuweisen. Ist also nicht beschränkt auf das menschliche Gehirn. Vielmehr postuliert es eine physikalische Struktur, die geeignet ist, um Bewusstsein hervorzubringen. Das könnten dann auch Maschinen, Gesellschaften oder etwa Galaxien sein. Als Ergebnis liefert IIT den Informationsgehalt  $\Phi$  der dem Grad des Bewusstseins entspricht und einen dazugehörigen Komplex, welcher die Qualia beschreibt. Die mathematische Beschreibung würde es ermöglichen solche Ergebnisse rechnerisch zu bestimmen. Aufgrund der dabei enorm exponentiellen Charakteristik ist das nur für sehr kleine Systeme und nicht für das menschliche Gehirn möglich. Meist sind es nur wenige Logik-Gatter anhand derer das Modell vorgestellt wird. Für ein allgemeineres Verständnis wird hier die Metapher eines Familiensystems zum Ansatz gebracht.**

Ein möglicher Weg, um Bewusstsein nachzuweisen wäre die Gehirne von verschiedenen Wesen zu untersuchen und den Ort dieser Erfahrung zu lokalisieren. Von Neurowissenschaftlern wird diese Methode zusehends häufiger angewendet. Dabei werden Probanden angehalten sich in einen bewussten Zustand zu versetzen, während gleichzeitig Aktivitäten in verschiedenen Hirnregionen gemessen werden. Das funktioniert bei Menschen aufgrund ihrer sprachlichen Kommunikation schon gut. Bei Koma Patienten nicht, bei anderen biologischen Wesen kaum und bei künstlichen Wesen schon gar nicht. Bei der integrierten Informationstheorie (IIT) wird nun der Versuchsaufbau umgedreht. Es wird eine Hypothese für Bewusstsein aufgestellt und diese anschließend durch Messungen am Gehirn validiert. IIT geht nun von fünf unverrückbaren Annahmen aus welche als Axiome aufzufassen sind. Das sind: "existence, structure, information, integration und exclusion". Sofern eine Materie, könnte ein Gehirn aber auch eine KI sein, diese fünf Eigenschaften aufweist hat es ein Bewusstsein mit einer von eben dieser bestimmten Quantität und Qualität. Bewusstsein ist ein System, welches sich mathematisch beschreiben lässt, in Lebewesen zu finden ist und wahrscheinlich auch konstruierbar ist.

Die Axiome von IIT für Bewusstsein sind bei weitem noch nicht so robust wie etwa die der Mathematik, der Lichtgeschwindigkeit oder der Gravitation. Sie sind aber gut nachvollziehbar und jeder Mensch kann diese durch Eigenversuche bestätigen aber auch widerlegen. Im Einzelnen haben die Axiome folgende Bedeutung:

Existence: "Bob" existiert dann, wenn er als Ursache auf "Alice" eine Wirkung ausüben kann oder umgekehrt, wenn "Bob" von "Alice" beeinflussbar ist. Ein Sonderfall von Existenz liegt dann vor, wenn sich diese Akteure selbst beeinflussen können. Die Ursache für die Wirkung liegt dann im Inneren. Obwohl es gegen den Hausverstand spricht, kann man nahezu jede Existenz in Zweifel ziehen. Ob der Mensch am Mond gelandet ist, es ein Corona Virus gibt oder Ameisen die größten Klima Schädlinge sind, kann in Frage gestellt werden. Illusionisten lehnen sogar die Existenz der physikalischen Welt ab. Was allerdings nicht in Zweifel gezogen werden kann ist das eigene Denken, auch wenn es falsch sein sollte - so Rene Descartes.

Structure: Wärme, Berührung, Nahrung, Stimme der Mutter, usw. sind Grunderfahrungen. Daraus und mit jeder neuen Erfahrung entwickelt sich im Laufe des Lebens ein riesiges Erfahrungsrepertoire. Es sind Strukturen und keine "EIN - Erfahrung" die alle anderen Erfahrungen überdeckt. Mein aktuelles Erleben von Skifahren beinhaltet die Teil-Erfahrungen von weiß, Schnee, Sonne, Schmerz, Geschwindigkeit, usw. Auch machen wir völlig neue Erfahrungen, die nur wenig in eine Struktur passen. Alle Erfahrungen haben eines Gemeinsam. Im kleinsten sind es Kompositionen von sehen, hören, fühlen, riechen und schmecken. Andere Erfahrungen kann der Mensch nicht machen.

Information: Es macht schon einen Unterschied, ob ich krank oder gesund bin. Wer noch nie Schmerzen hatte kennt diesen Unterschied nicht. Die Fähigkeit eines Wesens Unterschiede zu erkennen ist gleichbedeutend mit Informationen verarbeiten. Wenn ich in meinem Umfeld keine Unterschiede erkennen kann, reagiere ich auch nicht. Der Motor meines Autos macht immer das gleiche Geräusch. Selbst wenn es nur minimal anders ist und der Beifahrer das nicht bemerken bin ich alarmiert. Das ist eine von Äußere Unterscheidung, die dann noch auf eine innere Unterscheidung trifft: „Es ist nicht so wie immer“. Information kann auch intern entstehen. Im Traum ist meine Sensorik stark reduziert und trotzdem mache ich Erfahrungen. IIT beschäftigt sich hauptsächlich mit interner Informationsverarbeitung.

Integration: Die Erfahrung von einem Einsatzfahrzeug mit Blaulicht hat jeder schon einmal gemacht. Obwohl die Szene aus einem blauen Licht, einem fahrenden Auto, einer Straße usw. besteht erleben wir diese Komponenten nicht explizit. Verschieden Teile werden zu einer gesamten Erfahrung integriert. Das Gehirn könnte aus dieser Szene aber auch eine völlig andere Erfahrung generieren. Etwa ein Straßen Spektakel mit blauem Diskolicht. Die Fähigkeit aus eigener Kraft Information zu gewinnen die von außen nicht zugänglich ist wird bei IIT als Integration gesehen.

Exclusion: Gerade jetzt wird mir bewusst, dass ich einen Freund anrufen sollte. Wenn das so ist, kann mir nicht auch gleichzeitig bewusst sein: "ein Ei kochen zu wollen". Zu einem Zeitpunkt, auch wenn er sehr kurz ist, kann ich mir nur einer Sache bewusst sein. Nach IIT wird jene Erfahrung bewusst die den größten Unterschied macht also die meiste Information erzeugt.

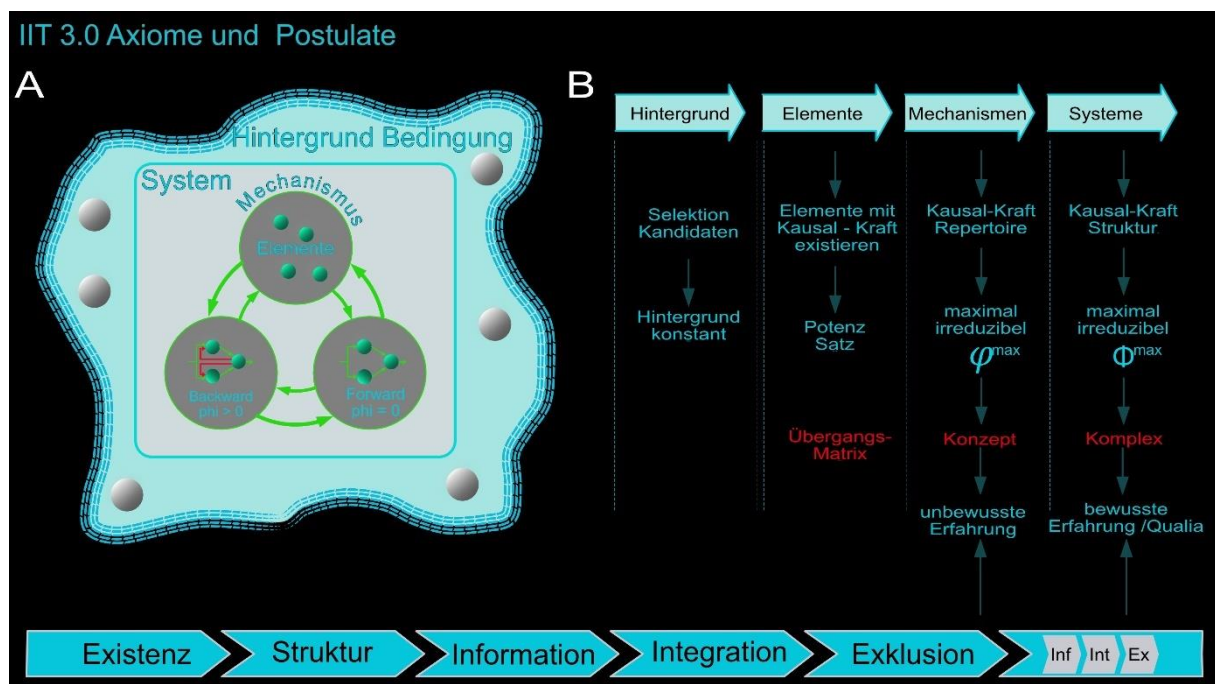


Abbildung 1 Systeme

IIT wurde 2004 von Giulio Toni (Tononi, 2014) vorgestellt und befindet sich aktuell in der Version 3.0. Aufbauend auf den Axiomen wird verlangt nach physikalischen Äquivalenten zu suchen (in allen Arten von Gehirnen) oder künstliche Strukturen zu designen die denen entsprechen. Ein Sandstrand, ein Fotosensor oder Fischeschwärme werden es nicht sein. Es braucht dazu ein komplexes System (Abbildung 20A) welches sich zu seinem Hintergrund differenziert und innerhalb aus wechselwirkenden Mechanismen besteht. Jeder Mechanismus setzt sich aus Elementen zusammen, die wiederum in Wechselwirkung stehen. Elemente können auf unterschiedlichste Weise aufeinander wirken. Bei IIT braucht es netzwerkartige Strukturen. Dabei unterscheidet man zwischen Feed - Forward Netzen wie bei KI oder im Kleinhirn und Feed-Backward Netzen wie im Neokortex. Lediglich Backward Typologie erzeugt interne Information und kann Bewusstsein generieren. Der dann vorliegende Informationsgehalt wird mit  $\Phi$  bezeichnet. Die Anzahl der Elemente, 100 Milliarden im Gehirn, zusammen mit dessen Vernetzung bestimmen den Grad des Bewusstseins. Neben dem Gehirn sollte man Bewusstsein dort suchen, wo noch Milliarden Elemente in Feedback Struktur angeordnet sind. Das könnte die Weltbevölkerung, das Internet oder eine Galaxie sein.

Aus einer Menge von Mechanismen (Abbildung 20B) wird für die Betrachtung ein Kandidatensatz selektiert. Die Mechanismen des Hintergrundes haben bei IIT vorerst keine Bedeutung. Für elementare Mechanismen ist gefordert, dass sie existieren. Das ist dann gegeben, wenn sie über eine kausale Kraft vermögen (**Potenzsatz**). Der aktuelle Zustand eines Mechanismus ist die Folge von vorhergegangenen Zuständen. Genauso determiniert der aktuelle Zustand den Zukünftigen (**Zustands-Übergangs-Matrix**). Innerhalb des Kandidaten-Satzes verfügen die Mechanismen über ein **Kausal-Kraft-Repertoire**. Die Wahrscheinlichkeitsverteilung der Repertoires und deren Abstand zu anderen ermöglicht die Berechnung der Informations Wahrscheinlichkeit von  $\varphi$ . Aus allen möglichen Wahrscheinlichkeiten wird die Höchste ausgewählt, mit  $\varphi^{\max}$  bezeichnet und mit einer Erfahrung gleichgesetzt. Die Struktur der Verteilung, hier **Konzept** genannt, ist der Quale einer Erfahrung zugeordnet. Auf Mechanismen Eben wahrscheinlich noch unbewusst.

Das menschliche Gehirn mit seinen Milliarden Neuronen (Elementen) ist zur Erklärung der Grundlagen von IIT zu komplex. Meistens sind es Logik-Gatter und hier auch nur die einfacheren (AND, OR, XOR) aus der Informatik. Diese bieten für Techniker einen guten Zugang. Für Nicht-Techniker wird begleitend, als Metapher ein Familiensystem geführt. Dieses Konstrukt ist aus der systemischen Therapie bekannt und vergleichbar mit dem Familienbrett. Dafür erhalten die Mitglieder der Familie "Simbolean" massiv vereinfachte mentale Grundausrichtungen.

**EXISTENCE:** Der Kandidatensatz der Familie Simbolean (Abbildung 21B) rekrutiert sich aus PAPA, MAMA und KIND. Der Hintergrund mit OMA, CHEF und FREUND bleibt unberücksichtigt. Diese sind da, haben aber keinen direkten Einfluss und werden als ein Hintergrundrauschen gesehen. PAPA ist sehr flexibel für Alternativen, vergleichbar mit einer ODER Funktion. Jede Alternative ist ihm recht sofern zumindest eine vorhanden ist. Im Gegensatz dazu ist MAMA sehr anspruchsvoll. Für eine Entscheidung müssen alle Bedingungen erfüllt sein, insbesondere was geordnete Abläufe angeht und ist vergleichbar mit einer UND Funktion. Es müssen alle Bedingungen erfüllt sein, um einen nächsten Schritt zu machen. Das KIND ist sehr wählerisch. Wenn ihm etwas angeboten wird, ist davon auszugehen, dass es genau das Gegenteil davon haben möchte. Ähnlich wie eine XOR Funktion. Es ist nicht flexibel und akzeptiert jeweils nur eine Möglichkeit, alle anderen bleiben ausgeschlossen. Ganz klar, reale Menschen lassen sich nicht so vereinfachen. Obwohl erkennen kann man das schon hin und wieder.

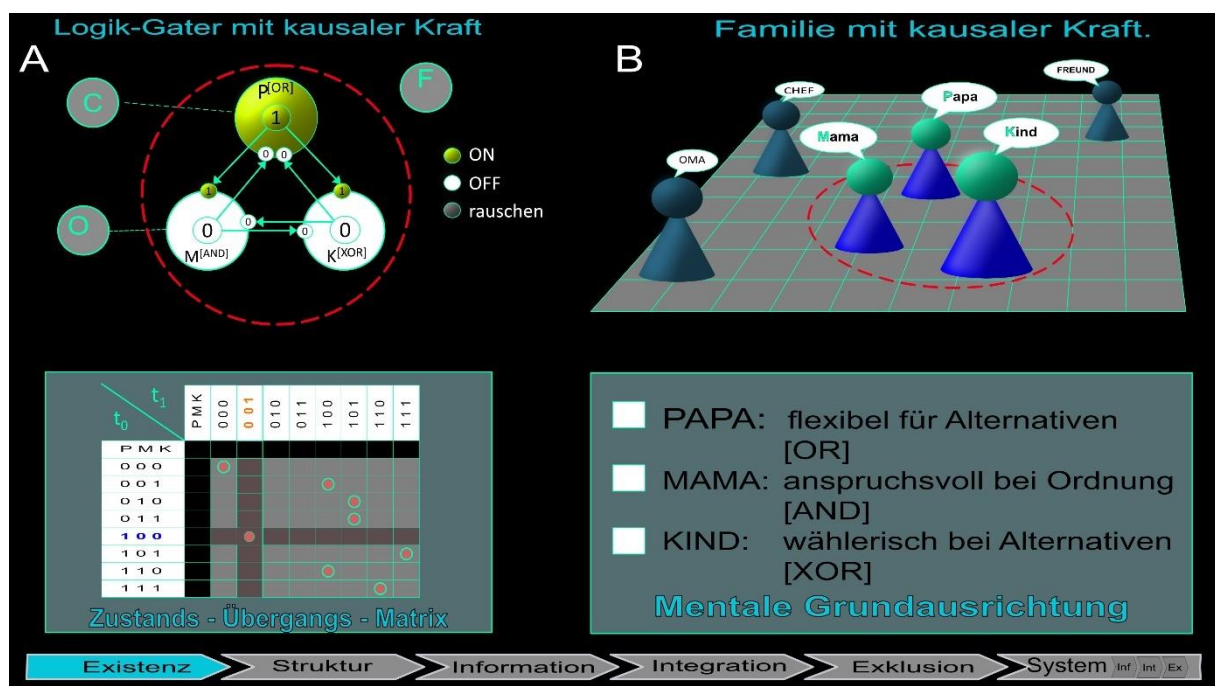


Abbildung 2 EXISTENCE

In Abbildung 21B nun die Simbolean's transferiert auf die Logik Gattern  $P^{[OR]}$ ,  $M^{[AND]}$  und  $K^{[XOR]}$ . Um dem Existence-Postulat einer kausalen Kraft gerecht zu werden sind alle Ausgänge mit den Eingängen der anderen Gatter verbunden. Lediglich der eigene Ausgang ist mit dem eigenen Eingang nicht verbunden. Ein Gatter kann nach binärer Logik die Zustände 0 oder 1 (1 bit) annehmen. Daraus ergibt sich für den Mechanismus PMK mit drei bits ein Repertoire von 000 bis 111. Für die weitere Analyse wird angenommen, dass sich der Mechanismus gerade im Zustand 100 befindet. Dieser Zustand bestimmt gleichzeitig den nächst Folgenden. Mit  $M=0$  und  $K=0$  an den Eingängen von  $P$  kann logisch nur  $PMK=001$  kommen. Dazu sucht man auf y-Achse (Zeitpunkt  $t_0$ ) der Zustands-Übergangs-Matrix den Wert 100 und erhält dann auf der x-Achse ( $t_1$ ) den Folgezustand von 001. Das kann nun für alle künftigen Zustände gemacht werden. Genau so kann der vergangene Zustand aus der Matrix abgelesen werden. Hätte man im Zeitpunkt  $t_1=100$  so wäre dieser aus  $t_0$  mit 001 oder 110 hervorgegangen.

**STRUCTURE:** Wenn man an die Familie Simbolean denkt, ist man geneigt alle Mitglieder gleichzeitig zu betrachten. Sofern die Beziehungen aufrecht sind, können Familienmitglieder auch einzeln auftreten und sie bleiben trotzdem eine Familie (Abbildung 22 B). Als Personen haben sie neben den Rollen in der Familie auch noch andere. PAPA ist in seinem Beruf Designer, die MAMA betreibt einen Online Shop, und das KIND geht zur Schule. Fallweise treten Familienmitglieder auch paarweise auf. PAPA und MAMA gehen gerne tanzen. Die MAMA lernt dem Kind das Schminken und weil der PAPA ein passionierter Fischer ist, nimmt er das KIND dazu mit. Manche Erfahrungen kann die Familie nur gemeinsam erleben. Etwa Urlaub, Geburtstagsfeiern, Möblieren, usw.

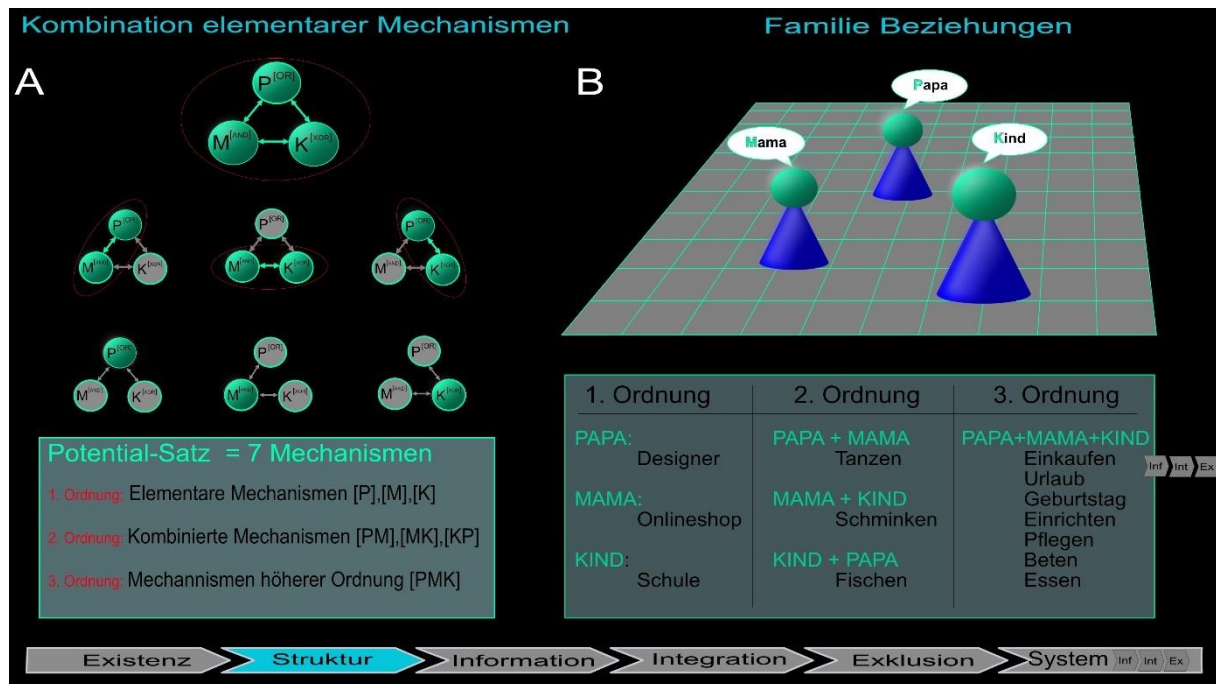


Abbildung 3 Struktur

So ähnlich ist es bei den Mechanismen der Logik-Gatter. Jedes Element wie  $P^{[OR]}$ ,  $M^{[AND]}$  und  $K^{[XOR]}$  kann als eigener elementarer Mechanismus betrachtet werden.  $P^{[OR]}$  wirkt als Mechanismus erster Ordnung, erfüllt aber gleichzeitig die Existence-Anforderung. Bei  $P^{[OR]} = 1$  in  $t_0$  kann die Frage, wie die Zustände von  $M^{[AND]}$  und  $K^{[XOR]}$  in  $t_1$  gewesen sein mussten, mit  $MK = 01, 10, 11$  beantwortet werden. Das gleiche Verfahren kann ebenfalls für alle Paare von Kombinationen der Mechanismen 2. Ordnung, angewendet werden. Letztendlich auch noch für den gesamten Mechanismus oder dem Mechanismus höherer Ordnung. Die Menge der Kombinationen  $M$  aus drei elementaren Mechanismen  $M\{P, M, K, PK, PM, MK, PMK\}$  ergibt sieben Möglichkeiten.  $M$  steigt exponentiell mit der Anzahl von Elementen in der 1. Ordnung. Mechanismen bis zu zwanzig Elementen lassen sich auf heutigen Computern noch berechnen. Für die Berechnung der 302 Neuronen des *C.elegans* (Wurm) reicht die Zeitspanne unseres Universums nicht aus. Daher ist die Berechnung unseres Gehirns in der höheren Ordnung vorerst ausgeschlossen. Sehr wohl aber für eine Handvoll Logik Gatter.



**INFORMATION:** Der Online Shop von MAMA ist schon ganz gut angelaufen. Für die weitere Entwicklung braucht sie ein Web-Portal. Dazu wiederum ist es erforderlich, dass sie einen Web Designer Kurs macht. Sie ist gezwungen dieses Vorhaben innerhalb der Familie abzustimmen. PAPA Simbolean hat schon immer im Haushalt mitgearbeitet. Jetzt muss er das allein bewerkstelligen **und** das KIND wird auf die Hilfe von MAMA bei den Hausübungen verzichten müssen **und** die OMA wird für den Shuttle-Dienst zur Musikschule benötigt. MAMA ist in der Zwangslage von allen das Einverständnis zu bekommen. Wie wahrscheinlich ist das?

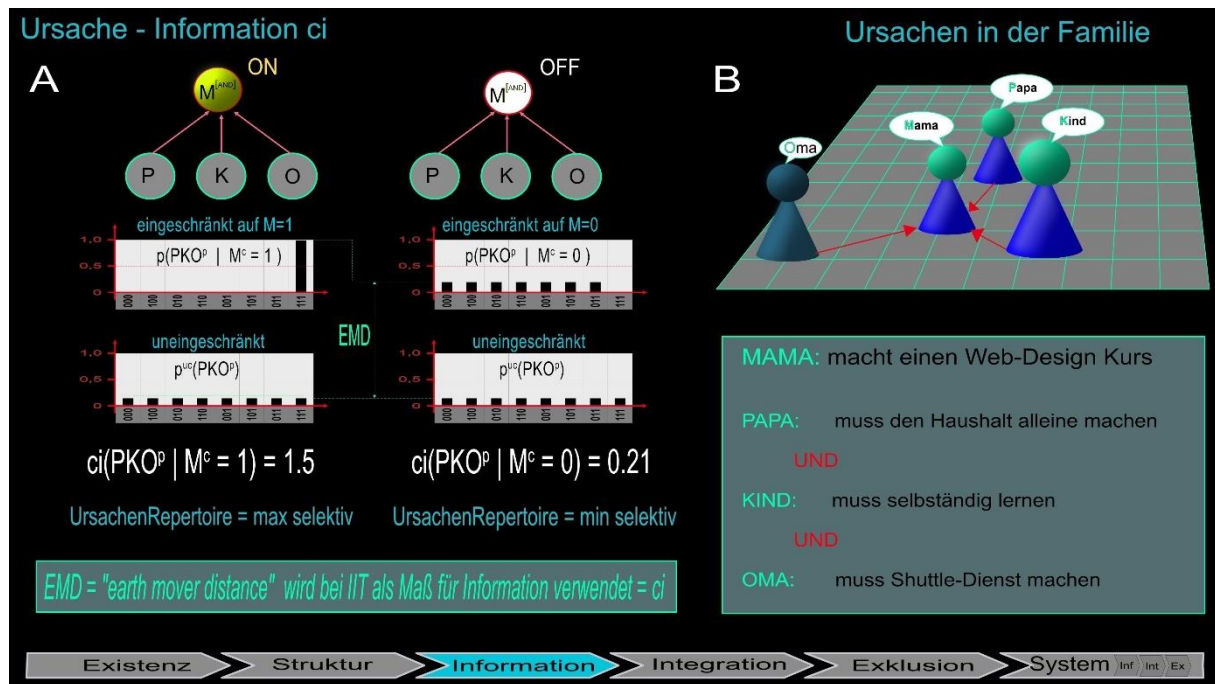


Abbildung 4 INFORMATION ci

Wir stellen diese Situation mit den Logik Gattern nach.  $M^{[AND]}$  hat dann die drei Eingänge PKO. Könnte nun  $M^{[AND]}$  alle seine möglichen Zustände (0,1) ungezwungen (unconstrained cause = uc) annehmen dann sind vorher ( $p$  = past) die Eingangszustände gleich verteilt  $p^{uc}(PKO^p | M = \text{zufällig})$ . Daraus ergibt sich eine Informations Wahrscheinlichkeit  $H$  (eta) nach Shannon mit 3 bit. Ist jedoch  $M^{[AND]}$  jetzt ( $c$  = current) im erzwungenen Zustand von ON (=1) dann mussten alle Eingänge in der Vergangenheit auf 111 gewesen sein. Mit der Wahrscheinlichkeitsverteilung von  $p(PKO^p | M^c = 1)$  ergibt sich ein  $H$  (eta) von 0bit. Aus intrinsischer Sicht generiert der Unterschied von 3bit zu 0bit wiederum eine Information. Seit der Version IIT 3.0 wird dieser aber nicht mehr nach der Shannon'schen Informations-Entropie berechnet, sondern nach der statistischen Methode des EMD (earth mover distance). Mittels EMD wird die Distanz zwischen der ungezwungenen Verteilung  $p^{uc}(PKO^p | M = \text{zufällig})$  und der erzwungenen Verteilung  $p(PKO^p | M^c = 1)$  errechnet. Der Wert von 1,5 ist dann die Metrik für die Ursachen-Information und wird in IIT als ci (cause information) bezeichnet.

#### Die Berechnung der Information mittels EMD zieht sich durch die ganze IIT.

EMD berechnet den Aufwand, den man investieren müsste, um eine Wahrscheinlichkeitsverteilung in eine andere zu transformieren. Im Falle von  $M^{[AND]} = 0$  ergibt sich ein D von 0,21. Alleine die visuelle Begutachtung zwischen den Verteilungen von  $p(PKO^p | M^c = 1)$  und  $p(PKO^p | M^c = 0)$  zeigt dass für  $M=0$  weniger Verschiebe-Aufwand erforderlich ist. Die Tatsächliche Berechnung von EMD ist aufwändig und wird meist programmiert (Python phi) oder in Mathematica designed.

Kürzlich hat das KIND der Familie Simbolean ein Nicht-Genügend auf die Mathematikarbeit heimgebracht (Abbildung 24B). Auf Grund des Web-Portals hat MAMA in den vergangenen Monaten gut verdient. PAPA bucht daraufhin den Urlaub. Der Grund für seine Entscheidung liegt in der Vergangenheit, wobei er dazu noch sehr flexibel ist und ihm schon eine positive Nachricht genügt. Nun ist der Urlaub gebucht, was auf das künftige Agieren von MAMA und KIND einen zwingenden Einfluss hat. Die MAMA verlangt vom KIND ab sofort bessere Noten und das KIND kann nun nicht mehr mit der Freundin in den Urlaub fahren. Es hat nur die entweder - oder Möglichkeit.

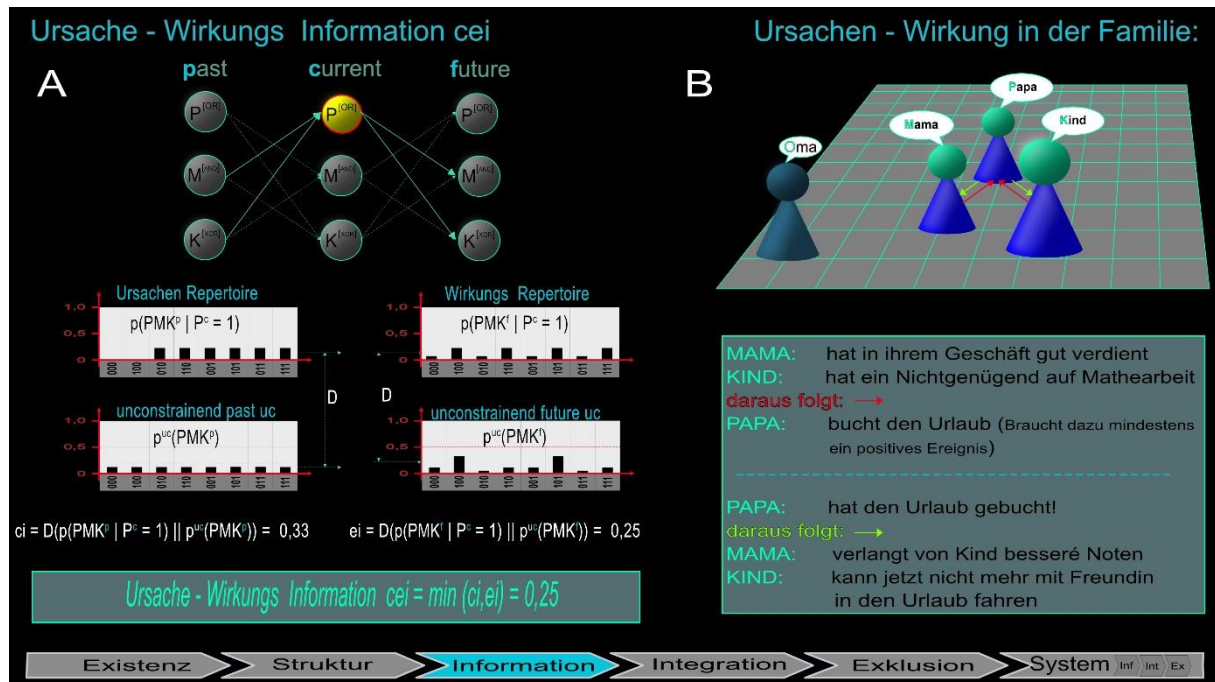


Abbildung 5 INFORMATION cei

Analog zu PAPA der Familie Simbolean wird nun bei den Logik-Gattern der elementare Mechanismus  $P^{(OR)}$  angeschaut und fix auf ON (=1) gesetzt (Abbildung 24A). Die dabei generierte Information ist zweigeteilt. Einerseits die Ursachen-Information ci (past -> current) und andererseits die Effekt-Information ei (current -> future). In beiden Fällen erfolgt die Berechnung über die "earth mover distance" D. Im Falle ci wird die Distanz (D) zwischen der ungezwungenen Wahrscheinlichkeits-Verteilung  $p^{uc}(PMK^p)$  und der erzwungen Wahrscheinlichkeitsverteilung  $p(PMK^p | P^c = 1)$  mit ( $ci = 0,33$ ) errechnet. Für ei wird das gleiche Verfahren angewendet, allerdings zwischen der aktuellen (c) und der künftigen Verteilung (f). Der mathematische Ausdruck dafür ist  $ei = D(p(PMK^f | P^c = 1) || (p^{uc}(PMK^f))) = 0,25$ . Lediglich der minimale dieser beiden Werte gilt als Ursache-Wirkungs-Information  $cei = 0,25$ . Dazu stellt man sich den Mechanismus  $P^{(OR)}=1$  als Trichter vor, der jeweils nur das Minimum durchlässt.

**Integration:** Das KIND der Familie Simbolean ist sehr an Musik interessiert und hat dafür auch eine Begabung (Abbildung 25C). Gleichzeitig sind MAMA und PAPA bestrebt dem KIND die bestmögliche Ausbildung zu garantieren. Sie haben in der Vergangenheit auch schon Geld für eine private Musikschule gespart. Andererseits denkt das KIND auch über seine musikalische Fortbildung nach. Nun stehen zwei Alternativen an. Das KIND, weil es selbständig sein möchte, beschließt ein Musikinstrument im Selbststudium zu lernen - es handelt **autonom**. Andererseits könnte es das Geld der Eltern nehmen - es handelt **integriert**. Unter professionelle Begleitung wird die Entwicklung wahrscheinlich besser. Die Simbolean' s erreichen mehr, wenn sie integriert handeln.

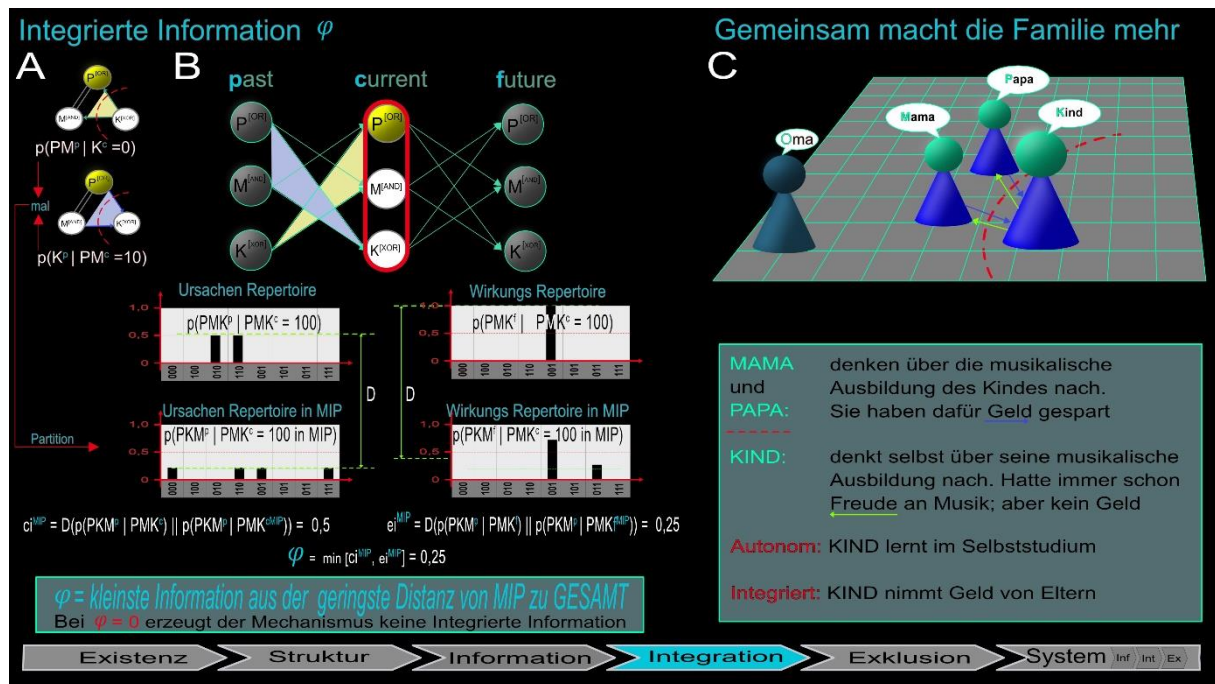


Abbildung 6 INTEGRIRT

Integrierte Information liegt nur dann vor, wenn keiner seiner elementaren Mechanismen das gleiche Ergebnis erzielt wie der gesamte Mechanismus. Dazu muss jede mögliche Kombination (Partition) daraufhin überprüft werden. Eine Informations-Partition entsteht durch Aufschneiden. In Abbildung 25A erfolgt der Schnitt zwischen K und PM wobei die wechselseitigen Ursachen bei  $p \rightarrow c$  ermittelt werden. Die Wahrscheinlichkeits-Verteilung der Partition ergibt sich aus  $p(PM^p | K^c = 0) \times p(K^c | PM^c = 10)$  also durch Multiplikation. Aus den vielen möglichen Partitionen ist jene auszuwählen die den geringsten Unterschied zum gesamten Mechanismus macht. Anschließend (Abbildung 25B) kann die Distanz zwischen MIP (Minimaler Informations Partition) und dem Gesamt-Mechanismus berechnet werden. Für die Ursachen-Information  $ci^{MIP} = D(p(PMK^p | PMK^c = 1) || p(PMK^p | PMK^{cMIP}))$  ergibt das 0,5. Mit gleicher Methode erhält man für die minimale Effekt-Information  $ei^{MIP}$  einen Wert von 0,25. Die geringste Distanz im Mechanismus ist wieder das Minimum zwischen  $ci^{MIP}$  und  $ei^{MIP}$  also 0,25 was bei IIT mit  $\phi$  bezeichnet wird.  $\phi$  ist die kleinste Information aus der geringsten Distanz von MIP zu GESAMT. Sofern  $\phi > 0$  ist, ist der Mechanismus nicht auf einen seiner elementaren Mechanismen reduzierbar also irreduzibel. Bei  $\phi = 0$  liegt keine integrierte Information vor, ist reduzierbar.



**Exclusion:** Die Familie Simbolean (Abbildung 26C) überlegt die Anschaffung eines neuen Wohnzimmers. Weil PAPA Designer ist, steht ihm die größte Kompetenz bei der Planung zu. Bevor er allerdings einen Auftrag an einen Tischler gibt, möchte er das Einverständnis sowohl der MAMA als auch des KINDES. Ursache für seine Entscheidung sind beide. Jetzt muss das Ganze noch finanziert werden. Die MAMA hat die Finanzen der Familie über. Damit hat PAPA eine sehr starke Wirkung auf sie. Design und Finanzen sind der Kern für die Entscheidung. Alles andere hat kaum einen Einfluss. PAPA ist von Ursachen der Vergangenheit beeinflusst und erzeugt gleichzeitig Wirkung in die Zukunft.

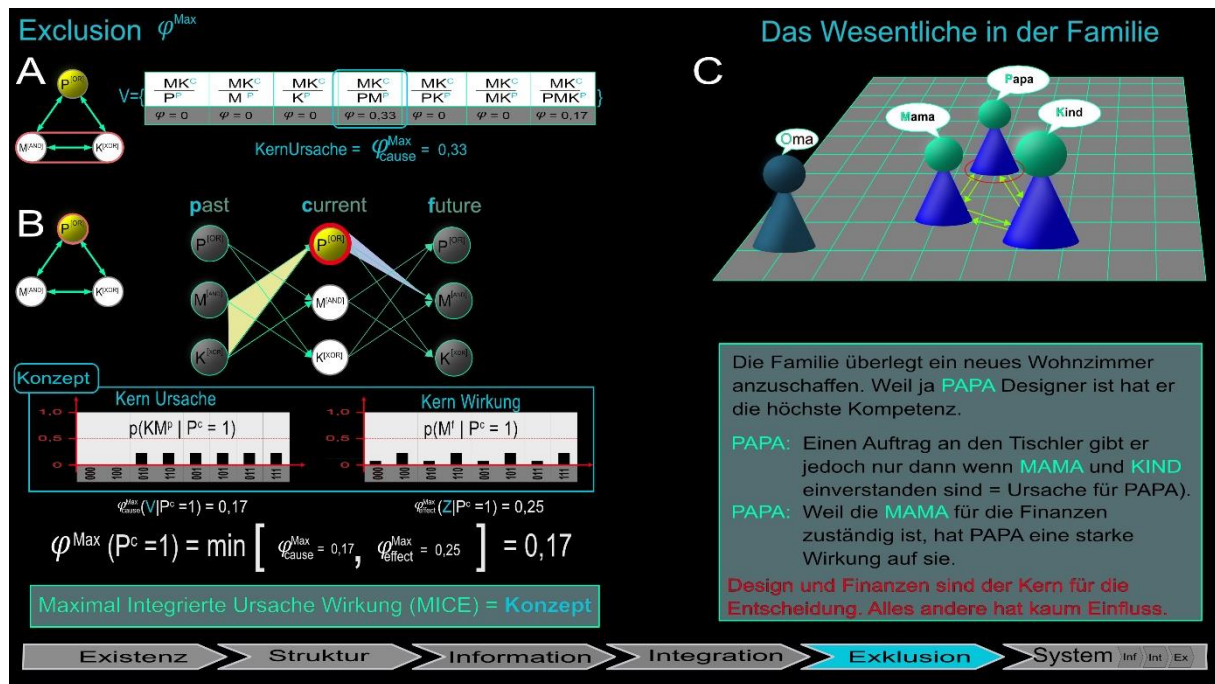


Abbildung 7 EXCLUSION

In Abbildung 26A wird vorerst der Mechanismus MK in jeder möglichen Wahrscheinlichkeitsverteilung angeschaut. Was heißt die Distanz der jeweiligen Wahrscheinlichkeits-Verteilungen MK zu V wobei  $V = \{P, M, K, PM, PK, MK, PMK\}$  zu berechnen. Der Fall MK zu PM erzeugt die größte Information mit  $\varphi = 0,33$ . Alle anderen sind kleiner oder Null. Die Kern-Ursache ist das  $\varphi_c^{Max}$ . Auch hier muss sowohl die Kern-Ursache als auch die Kern-Wirkung ermittelt werden (Abbildung 26B). Für den Mechanismus P bei 100 ergibt sich eine maximale Kernursache  $\varphi_c^{Max}(V | P^c = 1) = 0,17$  und ein Kern-Effekt von  $\varphi_e^{Max} = 0,25$ . Wiederum kommt nur der geringste Wert zum Ansatz was dann dem  $\varphi^{Max}$  dieses Mechanismus bestimmt. Im nächsten Kapitel, bei der systemischen Bestimmung muss für jeden möglichen Zustand (000 bis 100) sowohl ein  $\varphi_c^{Max}$  als auch ein dazugehöriges Konzept vorliegen. Ein Konzept eines Mechanismus sind die Ausprägungen der Wahrscheinlichkeits-Verteilungen von  $\varphi_c^{Max}$  und  $\varphi_e^{Max}$ . Transformiert auf Bewusstsein bedeutet  $\varphi^{Max}$  den Grad und Konzept die Qualia desselben. Auf Mechanismus-Ebene entspricht dies unbewussten Erfahrungen. Bewusstsein entsteht erst im Rahmen der systemischen Betrachtung.

Erkenntnisse:

- ☐ Ein System von Mechanismen sind über Feed-Back
- ☐ Mechanismen verfügen über intrinsische Kausal-
- ☐ Information als Distanz von Wahrscheinlichkeit-Verteilung
- ☐ Quantität =  $\varphi^{\text{Max}}$  und Qualität = Konzept von Erfahrung
- ☐ Berechenbar nur für kleine deterministische Systeme
- ☐ Ist eine Hypothese, die schon teilweise bestätigt

Literatur:

- |  |      |  |
|--|------|--|
| Masafumi Oizumi,<br>Larissa Albantakis,<br>Giulio Tononi               | 2014 | Von der Phänomenologie zu den Mechanismen des Bewusstseins:<br>Integrierte Informationstheorie 3.0 |
| Giulio Tononi,<br>Melanie Boly,<br>Marcello Massimini<br>Christof Koch | 2016 | Integrated information theory: from consciousness to its physical<br>substrate                     |
| Giulio Tononi  | 2012 | Phi: A Voyage from the Brain to the Sou  |
| Christof Koch  | 2019 | The Feeling of Life Itself: Why Consciousness Is Widespread but Can't Be<br>Computed               |
| Susan Blackmore  | 2018 | Consciousness: An Introduction   |

Glossar:

**Axiom:** Selbstverständliche Wahrheit über das Bewusstsein (Erfahrung existiert, sie ist nicht reduzierbar usw.). Die einzigen Wahrheiten, die bei Descartes nicht angezweifelt werden können und keinen Beweis benötigen. Sie sind Existenz, Zusammensetzung, Information, Integration und Ausschluss (siehe Text).

**Hintergrundbedingungen:** Externe Einschränkungen für einen Kandidatensatz von Elementen wurden behoben. Der vergangene und aktuelle Zustand der Elemente außerhalb des Kandidatensatzes ist auf ihre tatsächlichen Werte festgelegt.

**Kandidatenmenge:** Die Menge der betrachteten Elemente. Elemente innerhalb des Kandidatensatzes werden in alle möglichen Zustände versetzt, um das TPM des Kandidatensatzes zu erhalten.

**Ursache-Wirkungs-Repertoire:** Die Wahrscheinlichkeitsverteilung potenzieller vergangener und zukünftiger Zustände eines Systems, die durch einen Mechanismus in seinem aktuellen Zustand eingeschränkt wird.

**Ursache-Wirkungs-Informationen (  $cei$  ):** Die Informationsmenge, die von einem Mechanismus in einem Zustand angegeben wird, gemessen als Minimum an Ursacheninformationen (  $ci$  ) und Wirkungsinformationen (  $ei$  ).

**Ursacheninformationen (  $ci$  ) und Wirkungsinformationen (  $ei$  ):** Informationen über Vergangenheit und Zukunft, die als Abstand zwischen dem Ursachenrepertoire und dem unbeschränkten Ursachenrepertoire gemessen werden (auf der Effektseite gleich).

**Anlage:** Ein Satz von Elementen in einem System, das ein lokales Maximum des integrierten konzeptionelle Informationen  $\Phi$  erzeugt  $^{Max}$ . Nur ein Komplex existiert als Einheit aus seiner eigenen intrinsischen Perspektive.

**Konzept:** Eine Reihe von Elementen innerhalb eines Systems und das maximal nicht reduzierbare Ursache-Wirkungs-Repertoire, das es spezifiziert, mit dem zugehörigen Wert der integrierten Information  $\varphi^{max}$ . Das Konzept drückt die kausale Rolle eines Mechanismus innerhalb eines Komplexes aus.

**Konzeptionelle Struktur, Konstellation von Konzepten (  $C$  ):** Eine konzeptionelle Struktur ist die Menge aller Konzepte, die von einer Kandidatenmenge mit ihren jeweiligen  $\varphi^{Max}$ - Werten spezifiziert werden und als Konstellation im Konzeptraum dargestellt werden können.

**Conceptual Information (  $CI$  ):** Ein Maß dafür, wie viele verschiedene Konzepte von einem Elementsystem generiert werden.  $CI$  wird durch den Abstand  $D$  zwischen der Konstellation von Konzepten und dem „Null“-Konzept, dem uneingeschränkten Ursache-Wirkungs-Repertoire  $p^{uc}$ , quantifiziert.

**Konzeptraum:** Der Konzeptraum ist ein hochdimensionaler Raum mit einer Achse für jeden möglichen vergangenen und zukünftigen Zustand des Systems, in dem eine konzeptionelle Struktur dargestellt werden kann.

**Entfernung (  $D$  ):** In IIT 3.0 die Wasserstein-Entfernung, auch als Earth Mover-Entfernung (EMD) bezeichnet. Es gibt die Metrik des Konzeptraums und damit den Abstand zwischen Wahrscheinlichkeitsverteilungen (  $\varphi$  ) und zwischen Konstellationen von Konzepten (  $\Phi$  ) an.

**Integrierte konzeptionelle Informationen (  $\Phi$  ):** konzeptionelle Informationen, die von einem System über die konzeptionellen Informationen hinaus generiert werden, die durch seine (minimalen) Teile generiert werden.  $\Phi$  misst die Integration oder Irreduzibilität einer Konstellation von Konzepten (Integration auf Systemebene).

**Integrierte Information (  $\varphi$  ):** Information, die durch einen Mechanismus erzeugt wird, der über die Information hinausgeht, die durch seine (minimalen) Teile erzeugt wird.  $\varphi$  misst die Integration oder Irreduzibilität von Mechanismen (Integration auf Mechanismenebene).

**Eigeninformationen:** Unterschiede, die innerhalb eines Systems einen Unterschied machen.

**Mechanismus:** Jedes Subsystem eines Systems, einschließlich des Systems selbst, das eine kausale Rolle innerhalb des Systems spielt, z. B. ein Neuron im Gehirn oder ein Logikgatter in einem Computer.

**MICE (maximal irreduzibles Ursache-Wirkungs-Repertoire):** Das Ursache-Wirkungs-Repertoire eines Konzepts, dh das Ursache-Wirkungs-Repertoire, das unter allen möglichen Zuständen ein Maximum an integrierter Information  $\phi$  erzeugt.

**MICS (maximal irreduziblen konzeptuelle Struktur):** Die begriffliche durch einen Komplex in einem Zustand erzeugt Struktur, die entspricht ein lokales Maximum des integrierten konzeptionelle Informationen  $\Phi^{\text{Max}}$  (gleichbedeutend mit „quale“ oder „Konstellation“ in „qualia space“).

**MIP (Minimum Information Partition):** Die Partition, die den geringsten Unterschied macht (mit anderen Worten, die minimale "Differenz" -Partition).

**Null-Konzept:** Das uneingeschränkte Ursache-Wirkungs-Repertoire  $p^{uc}$  der Kandidatenmenge mit  $\phi = 0$ .

**Partition:** Unterteilung einer Reihe von Elementen in kausal / informativ unabhängige Teile, durchgeführt durch Rauschen der Verbindungen zwischen den Teilen.

**Potenzsatz:** Der Satz aller Teilmengen eines Kandidatensatzes von Elementen.

**Postulate:** Aus Axiomen abgeleitete Annahmen über die physischen Substrate des Bewusstseins (Mechanismen müssen kausale Kraft haben, nicht reduzierbar sein usw.), die formalisiert werden können und die Grundlage des mathematischen Rahmens der IIT bilden. Sie sind Existenz, Zusammensetzung, Information, Integration und Ausschluss (siehe Text).

**Aufgabenbereich:** Jeder Satz von Elementen eines Kandidatensatzes, über den das Ursache-Wirkungs-Repertoire eines Mechanismus in einem Zustand berechnet wird.

**Quale:** Die konzeptionelle Struktur, die von einem Komplex in einem Zustand erzeugt wird, der einem lokalen Maximum integrierter konzeptioneller Informationen entspricht.  $\Phi^{\text{Max}}$  (synonym mit „MICS“ oder „Konstellation“ im „Qualia-Raum“).

**Qualia-Raum:** Wenn eine Reihe von Elementen einen Komplex bildet, wird ihr Konzeptraum als Qualia-Raum bezeichnet.

**System:** Eine Reihe von Elementen / Mechanismen.

**TPM (Übergangswahrscheinlichkeitsmatrix):** Eine Matrix, die die Wahrscheinlichkeit angibt, mit der ein Zustand eines Systems in einen anderen Systemzustand übergeht. Das TPM wird durch die Mechanismen eines Systems bestimmt und durch Störung des Systems in alle möglichen Zustände erhalten.

**Unbeschränktes Repertoire (  $p^{uc}$  ) :** Die Wahrscheinlichkeitsverteilung potenzieller vergangener und zukünftiger Systemzustände ohne Einschränkungen aufgrund eines Mechanismus in einem Zustand. Das uneingeschränkte Ursachenrepertoire ist die gleichmäßige Verteilung der Systemzustände. Das uneingeschränkte Effektrepertoire wird erhalten, indem uneingeschränkte Eingaben für alle Systemelemente angenommen werden.