

Petri-Netze
eine informelle Einführung in
die Grundideen

von Ulrich Grude, 1988

Inhaltsverzeichnis

0	Vorwort.....	3
1	Das Universum der Signale.....	4
2	Von den Signalen und Interaktionen zu den Bedingungen und Ereignissen der Netztheorie.....	7
3	Nebenläufigkeit ist nicht Gleichzeitigkeit, sondern Unabhängigkeit.....	9
4	Prozessnetze und Zeit: zwei unabhängige und frei kombinierbare Konzepte.....	12
5	Prozessnetze beschreiben Prozesse objektiv, d.h. unabhängig von speziellen Beobachtungen.....	14
6	Ereignisse sind atomar und sollten nicht in zwei Hälften gespalten werden.....	18
7	Die Schaltregel für einzelne Ereignisse sagt nichts über Nebenläufigkeit.....	22
8	Die Schritt-Semantik ist nur die zweitbeste Semantik für B/E-Netze.....	24
9	Kontakte stören die Prozessnetz-Semantik, aber nur ein bisschen.....	27
10	Kontakte lassen sich durch Komplemente beseitigen, aber ...	30
11	Konfusion: In Netzen klar erkennbar.....	35
12	S/T-Netze, zwei Ideen: Mehrere-Token und unterscheidbare Marken.....	38
13	Kann eine Marke einen Schaltvorgang überdauern?.....	40
14	Schlingen in Netzen, oder: Nebenbedingungen in der Wirklichkeit?.....	42
15	S/T-Netz-Problem 1: Prozesse unterscheiden ununterscheidbare Token.....	48
16	S/T-Netz-Problem 2: Wie nebenläufig sind S/T-Netze?.....	51
17	Netze in Matrix-Darstellung: Von den B/E-Netzen zu den farbigen Netzen.....	60
18	Farbige Netze und Pr/T-Netze: wie verhalten sich diese Netzarten zueinander?.....	69
19	Ist jedes Netz ein Graph oder jeder Graph ein Netz?.....	73
20	Theorie und Anwendungen von Petri-Netzen.....	76
21	Literatur-Angaben.....	77

0 Vorwort

Einige Grundideen der Petri-Netze wurden von Carl Adam Petri in seiner Dissertation ("Kommunikation mit Automaten", Technische Hochschule Darmstadt, 1961) veröffentlicht. Später leitete Dr. Petri ein Forschungsinstitut der Gesellschaft für Mathematik und Datenverarbeitung (GMD) in Bonn-Birlinghoven, in dem er und seine Mitarbeiter diese Ideen weiterentwickelten. Ab 1985 hatte ich fast drei Jahre lang Gelegenheit, mich an diesem Forschungsinstitut über Petri-Netze zu informieren. Dies geschah im Rahmen eines Technologie-Transfer-Projekts zwischen der Firma PSI (Gesellschaft für Prozess-Steuerungs- und Informationssysteme, Berlin) und der GMD in Bonn. Die PSI hatte ein Petri-Netz-Werkzeug entwickelt und war an "mehr Theorie" interessiert. Die GMD wollte wissen, was die Industrie mit Petri-Netzen macht und was konkrete Petri-Netz-Werkzeuge "in der Praxis" leisteten.

In Birlinghoven lernte ich vor allem "per Osmose": Beim Mittagessen oder Kaffeetrinken mit anderen Institutsmitarbeitern, bei Gesprächen mit Gästen etc. Prof. Petri war ein sehr freundlicher Institutsleiter und kümmerte sich um jeden seiner Mitarbeiter. Außerdem war ihm die Weitergabe und Verbreitung seiner Ideen sehr wichtig. Mehrmals nahm er sich selbst ein paar Stunden Zeit, um mir geduldig einige der schwierigeren Aspekte seiner Netze zu erläutern.

Wenn ich eine bestimmte Idee in informellen Gesprächen kennengelernt hatte, fand ich sie manchmal auch in Papieren und Büchern über Petri-Netze wieder. Vorher waren sie mir häufig (hinter viel Formalismus) verborgen geblieben. Gegen Ende meiner Zeit in Birlinghoven versuchte ich, wichtige Grundideen der Netztheorie informell und möglichst leicht verständlich darzustellen, sozusagen eine Zusammenfassung vieler Gespräche beim Kaffeetrinken. So entstand das folgende Papier.

Die elektronische Form des Papiers ging im Laufe der Jahre verloren. Anfang 2005 hat meine Frau Barbara eine der letzten Papier-Kopien eingescannt und überarbeitet. Vielen Dank dafür. Vielen Dank auch an meine damaligen Gesprächspartner in Birlinghoven von 1985 bis 1987, die mir so viel erklärt haben: Eike Best, Jörg Desel, Cesar Fernández, Hartmann Genrich, Ursula Goltz, Carlos Heuser, Kurt Lautenbach, Agathe Merceron, Wolfgang Reisig und Klaus Voss. Und besonderen Dank an Prof. Petri für seine Ideen und seine Erläuterungen.

Berlin, im März 2005

Ulrich Grude.

1 Das Universum der Signale

Das "Universum der Signale" ist eine Vorstellung, die auf den Physiker Wilhelm Reichenbach (und möglicherweise auf noch frühere Autoren) zurückgeht und im Anhang des Buches "Symbolic Logic" von Rudolf Carnap axiomatisch beschrieben wird. Dieses Universum ist ein gedanklicher Vorläufer der Petri-Netze und seine Kenntnis kann ein tieferes Verständnis der Netztheorie erleichtern.

Im Universum der Signale bestehen alle Dinge letztlich aus Signalen. Signale haben Ähnlichkeit mit physikalischen Elementarteilchen.

Ein Signal kann (nur) zweierlei tun:

1. es kann sich "allein für sich" ausbreiten
2. es kann mit anderen Signalen interagieren.

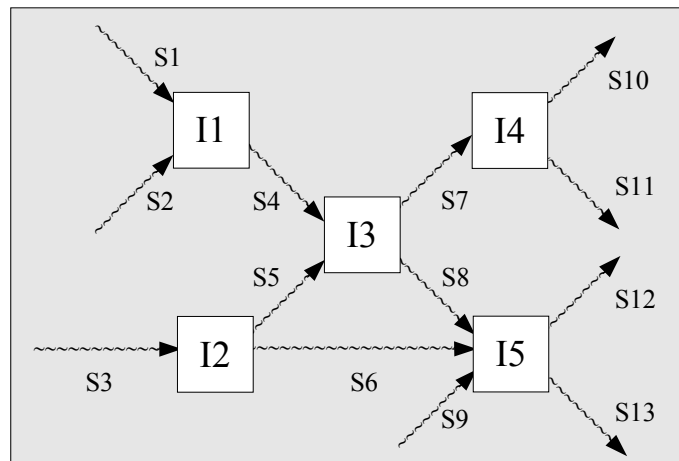


Abbildung 1.1: Signale und Interaktionen

In Abb. 1.1 sind die Signale, die sich allein für sich ausbreiten, durch gewellte Pfeile und Interaktionen von Signalen durch Rechtecke dargestellt.

Ein Signal entsteht durch eine Interaktion von Signalen und es vergeht, indem es seinerseits mit anderen Signalen interagiert.

In Abb. 1.1 entsteht z. B. das Signal S4 durch die Interaktion von S1 und S2 (d.h. durch die Interaktion I1). S4 erreicht sein Ende, indem es mit S5 interagiert (Interaktion I3).

Eine Interaktion ist durch ihre Eingangssignale und ihre Ausgangssignale bestimmt. In Abb. 1.1 ist z. B. I3 die Interaktion mit den Eingangssignalen S4 und S5 und den Ausgangssignalen S7 und S8.

Wären S4 und S5 Billiard-Kugeln und wäre I3 ein Zusammenstoß, dann könnte man z. B. fragen, ob die Kugel S7 (die aus dem Zusammenstoß herauskommt) nicht die selbe ist wie die Kugel S4 (die in den Zusammenstoß hineingeht). Um die Frage zu beantworten, könnte man die Kugeln mit individuellen Namen beschriften und den Zusammenstoß beobachten.

Da Signale im Universum der Signale die kleinsten Elementarbestandteile sind, kann man sie prinzipiell nicht individuell unterscheiden und nicht direkt beobachten (ganz ähnlich wie man etwa Photonen prinzipiell nicht sehen kann). Deshalb ist z. B. die Aussage: "Die Signale S4 und S7 sind identisch" weder wahr noch falsch, sondern nicht sinnvoll.

Allgemein gilt: ein Signal ist durch zwei Interaktionen "begrenzt" oder bestimmt. Durch die erste dieser Interaktionen entsteht das Signal, durch die zweite hört es auf zu existieren.

Eine Interaktion ist durch n Eingangssignale und m Ausgangssignale begrenzt oder bestimmt. Die Eingangssignale verschwinden und die Ausgangssignale entstehen durch die Interaktion. Entsprechend den Werten von n und m kann man eine Reihe von interessanten Spezialfällen von Interaktionen unterscheiden:

Interaktionen mit $n = 0$ und $m > 0$ entsprechen einem spontanen *Erscheinen* von Signalen, z. B.:

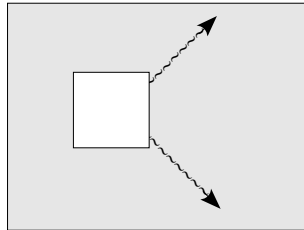


Abbildung 1.2: Spontanes Erscheinen zweier Signale

Interaktionen mit $n > 0$ und $m = 0$ entsprechen einem spontanen *Verschwinden*, z. B.:

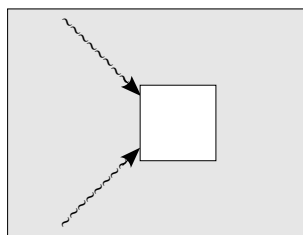


Abbildung 1.3: Spontanes Verschwinden zweier Signale

Ist $n = 1$ und $m = 2$, so entspricht das einem spontanen *Zerfall* eines Signals in zwei Signale:

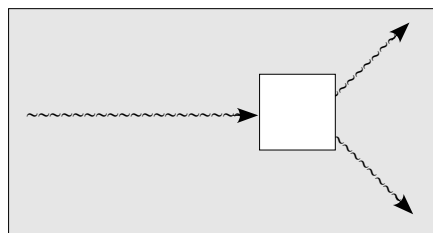


Abbildung 1.4: Ein spontaner Zerfall

Umgekehrt entspricht eine Interaktion mit $n = 2$ und $m = 1$ einer *Verschmelzung* von zwei Signalen zu einem Signal:

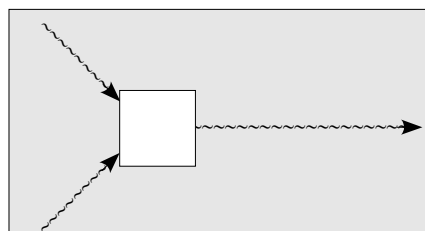


Abbildung 1.5: Eine Verschmelzung

Ein weiterer Spezialfall ist die Interaktion eines Signals "nur mit sich selbst", d.h. eine Interaktion mit $n = 1$ und $m = 1$:

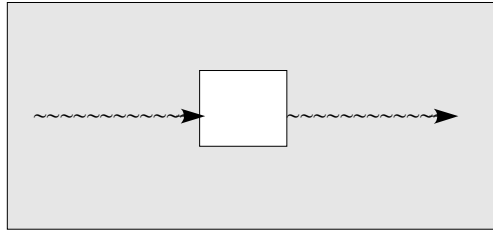


Abbildung 1.6: Eine Selbst-Interaktion

Man kann nun fragen, ob es spontanes Erscheinen und Verschwinden, spontane Zerfälle, Verschmelzungen und Selbstinteraktionen "in Wirklichkeit" gibt oder ob das Universum der Signale der physikalischen Realität besser entspricht, wenn man diese Spezialfälle ausschließt.

In einem Gas stoßen die Moleküle zufällig zusammen. Zusammenstöße zwischen zwei Molekülen sind dabei viel häufiger als Zusammenstöße, an denen drei oder noch mehr Moleküle beteiligt sind. Entsprechend kann man auch für Interaktionen zwischen Signalen annehmen, dass solche mit $n = 2$ viel häufiger auftreten als solche mit $n > 2$. Vielleicht genügt es sogar, wenn man nur Interaktionen mit $n = 2$ und $m = 2$ annimmt:

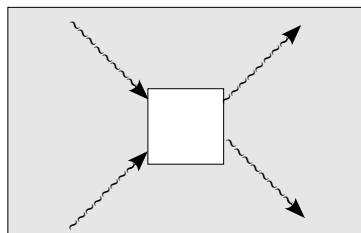


Abbildung 1.7: Eine "Normal-Interaktion"

Solche und ähnliche Fragen und Probleme sollen hier aber nicht weiter verfolgt werden. Sie wurden nur erwähnt, um die mit dem *Universum der Signale* verbundenen Vorstellungen etwas deutlicher zu machen.

2 Von den Signalen und Interaktionen zu den Bedingungen und Ereignissen der Netztheorie

In der Netztheorie unterscheidet man *Bedingungen* und *Ereignisse* von ihren *Vorkommnissen*. Z. B. ist B1: "Es ist Frühling" eine Bedingung und E1: "Es wird Frühling" ein Ereignis. Dagegen sind z. B. B1.1987: "Es ist Frühling 1987", B1.1988: "Es ist Frühling 1988" etc. Vorkommnisse der Bedingung B1. Entsprechend sind E1.1987: "Es wird Frühling 1987" und E1.1988: "Es wird Frühling 1988" etc. Vorkommnisse des Ereignisses E1.

Anstelle von "Vorkommnis einer Bedingung" bzw. "Vorkommnis eines Ereignisses" werden im folgenden die Bezeichnungen "unwiederholbare Bedingung" bzw. "unwiederholbares Ereignis" verwendet. Im Kontrast dazu werden Bedingungen und Ereignisse auch als "wiederholbare Bedingungen" bzw. als "wiederholbare Ereignisse" bezeichnet.

Somit ist B1 eine wiederholbare Bedingung (sie kann wiederholt erfüllt werden) und E1 ist ein wiederholbares Ereignis (es kann wiederholt eintreten). Dagegen sind B1.1987 und B1.1988 etc. unwiederholbare Bedingungen und E1.1987 und E1.1988 etc. unwiederholbare Ereignisse.

Die Bezeichnungen "Bedingung" und "Ereignis" ohne qualifizierendes Adjektiv werden (wie üblich) mehrdeutig verwendet. Aus dem jeweiligen Kontext folgt hoffentlich stets, ob wiederholbare oder unwiederholbare Bedingungen und Ereignissen oder beide Arten gemeint sind.

Netze aus wiederholbaren Bedingungen und Ereignissen werden auch als Bedingungs-Ereignis-Netze, kurz: B/E-Netze (englisch: condition event nets, short: C/E-nets) bezeichnet. Netze aus unwiederholbaren Bedingungen und Ereignissen werden auch als Prozessnetze bezeichnet (englisch: occurrence nets, zu deutsch etwa: Vorkommnis-Netze).

Als Beispiel für ein (markiertes) B/E-Netz und das zugehörige Prozessnetz dient häufig das folgende Netz "Die vier Jahreszeiten":

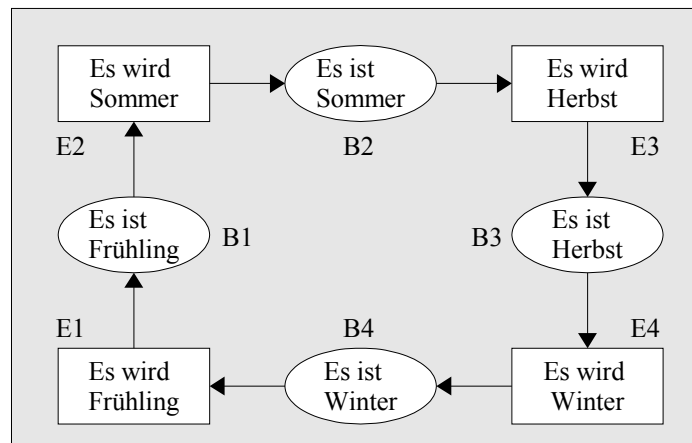


Abbildung 2.1: Das B/E-Netz "Die vier Jahreszeiten"

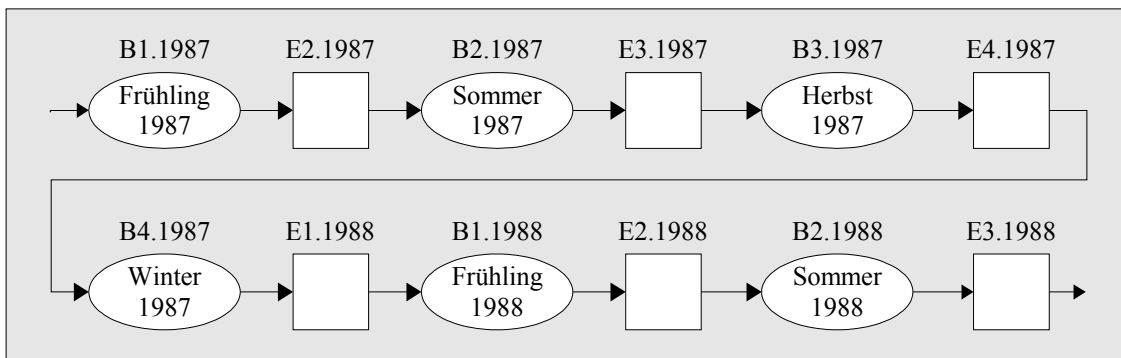


Abbildung 2.2: Das Prozessnetz "Die vier Jahreszeiten"

Unwiederholbare Bedingungen und Ereignisse haben sehr viel Ähnlichkeit mit Signalen und Interaktionen:

Eine unwiederholbare Bedingung ist durch zwei unwiederholbare Ereignisse *begrenzt* oder bestimmt. Durch das erste dieser Ereignisse wird die Bedingung *gültig* und durch das zweite wird sie wieder *ungültig*. Ein unwiederholbares Ereignis ist durch n unwiederholbare Vorbedingungen und m unwiederholbare Nachbedingungen begrenzt oder bestimmt. Die Vorbedingungen werden durch das Ereignis ungültig und die Nachbedingungen werden durch das Ereignis gültig.

Gemeinsam werden (unwiederholbare bzw. wiederholbare) Bedingungen und Ereignisse auch als (unwiederholbare bzw. wiederholbare) *Netz-Elemente* oder *Netzknoten* bezeichnet.

Unwiederholbare Netz-Elemente (oder: Vorkommnisse von Bedingungen und Ereignissen) sind unter anderem deshalb so wichtig, weil der Begriff der Nebenläufigkeit (englisch: concurrency) als Relation zwischen ihnen definiert ist: in einem Prozessnetz sind zwei Netz-Element X1 und X2 *nebenläufig* zueinander, wenn kein Pfeilweg von X1 nach X2 oder von X2 nach X1 führt.

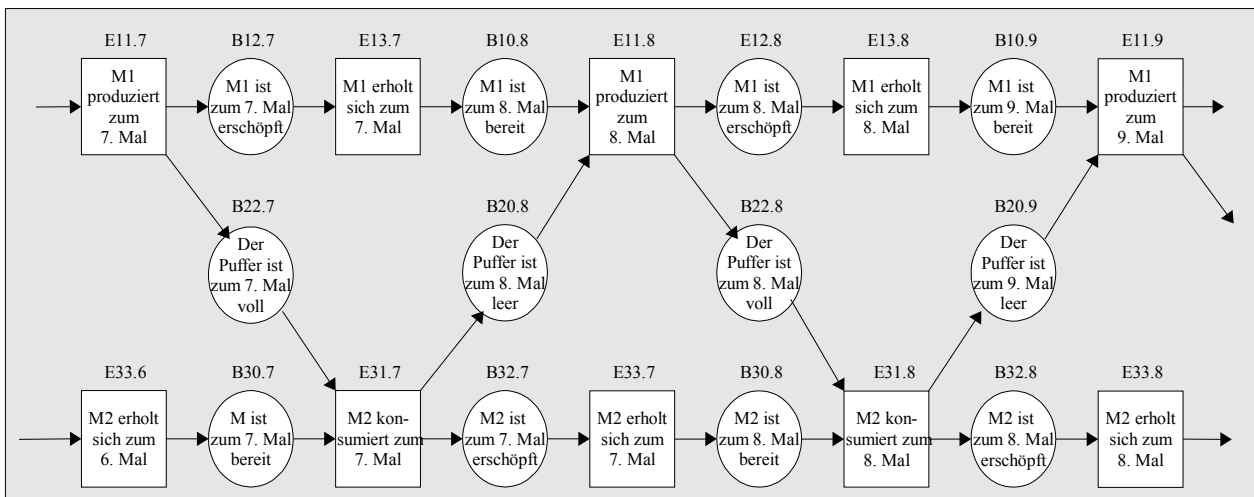


Abbildung 2.3: Ein Prozessnetz, welches die Zusammenarbeit zweier Maschinen beschreibt

In diesem Prozessnetz sind u.a. folgende Netz-Elemente nebenläufig zueinander: B12.7 und B22.7, B12.7 und E31.7, E13.7 und B30.7, E13.7 und E31.7 etc. *Nicht-nebenläufig* zueinander sind u.a. folgende Netz-Elemente: E11.7 und B22.7, E11.7 und E31.7, E11.7 und E33.8, B12.7 und B10.9 etc.

3 Nebenläufigkeit ist nicht Gleichzeitigkeit, sondern Unabhängigkeit

Fragt man Computerfachleute nach einer informellen und möglichst simplen Definition des Begriffs "Nebenläufigkeit", dann erhält man häufig eine Antwort, in der der Begriff "gleichzeitig" eine wesentliche Rolle spielt, etwa: "Ereignisse sind nebenläufig, wenn sie nicht nacheinander, sondern gleichzeitig passieren".

Eine solche Definition basiert auf der plausiblen Annahme, dass schon bekannt ist, was Zeit ist bzw. was "gleichzeitig" bedeutet.

In der Netztheorie wird diese Annahme ausdrücklich vermieden. Statt den Begriff Nebenläufigkeit auf Gleichzeitigkeit zurückzuführen geht man dort davon aus, dass Nebenläufigkeit ein fundamentaleres Phänomen ist als Zeit oder Gleichzeitigkeit. Ein Netz-Theoretiker würde deshalb eher versuchen, "Gleichzeitigkeit" mithilfe von "Nebenläufigkeit" zu definieren als umgekehrt.

Formal geht man in der Netztheorie von einer *Abhängigkeitsrelation* zwischen (unwiederholbaren) Netz-Elementen aus und definiert Nebenläufigkeit als *Unabhängigkeit*.

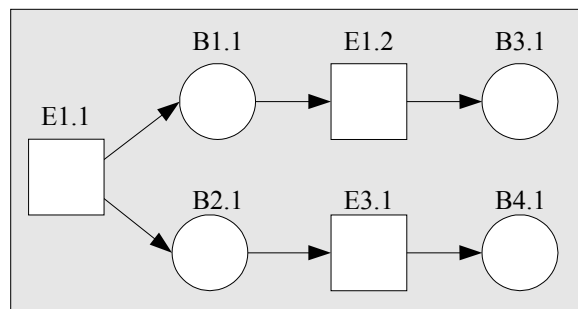


Abbildung 3.1: Ein Prozessnetz

In Abb. 3.1 ist ein (Ausschnitt aus einem) Prozessnetz dargestellt. Die Abhängigkeiten zwischen den Netz-Elementen (d.h. zwischen den unwiederholbaren Bedingungen und Ereignissen) sind durch Pfeile dargestellt. Die beiden Ereignisse E2.1 und E3.1 sind *unabhängig voneinander* (d.h. nicht durch einen Weg entlang den Pfeilen verbunden) und somit *nebenläufig*. Ebenso ist das Ereignis E2.1 nebenläufig zur Bedingung B4.1 und die beiden Bedingungen B3.1 und B4.1 sind nebenläufig zueinander. Das Ereignis E2.1 ist *abhängig* von B1.1 und von E1.1. Also sind E2.1 und B1.1 bzw. E2.1 und E1.1 *nicht nebenläufig* zueinander, etc.

Was "Abhängigkeit" inhaltlich bedeutet, muß von Anwendungsfall zu Anwendungsfall geklärt werden. Besonders wichtig sind physikalisch-kausale Abhängigkeiten. Aber auch andere Abhängigkeiten, von denen keineswegs offensichtlich ist, ob und wie sie auf physikalisch-kausale Abhängigkeiten zurückgeführt werden können, lassen sich durch Netze darstellen, z. B. die Abhängigkeit zwischen einer *Bitte* und ihrer *Erfüllung* oder einem *Befehl* und seiner *Ausführung*, die Abhängigkeit der vier Jahreszeiten voneinander u.s.w.

Nebenläufigkeit bedeutet in einem bestimmten Anwendungsfall die Abwesenheit der für diesen Anwendungsfall relevanten Abhängigkeiten.

Im folgenden soll N=U (bzw. N=G) die Auffassung bezeichnen, dass "Nebenläufigkeit" soviel wie "Unabhängigkeit" (bzw. soviel wie "Gleichzeitigkeit") bedeutet. Die beiden Auffassungen N=U und N=G sollen etwas genauer miteinander verglichen werden.

Die Netztheorie geht davon aus, dass zwischen nebenläufigen und nicht-nebenläufigen Netz-Elementen (d.h. unwiederholbaren Bedingungen und Ereignissen) ein Unterschied besteht, der 1. praktisch relevant, 2. objektiv und 3. qualitativ (oder: fundamental) ist. Von daher hat die Netztheorie drei Einwände gegen die Auffassung $N=G$ bzw. Argumente für die Auffassung $N=U$.

Der Einfachheit halber beschränken wir uns im folgenden auf die Diskussion von (unwiederholbaren) Ereignissen. Für Bedingungen gilt jeweils entsprechendes.

Einwand 1: Wenn zwischen zwei unwiederholbaren Ereignissen $E1.n$ und $E2.m$ keine (für den betreffenden Anwendungsfall wesentliche) Abhängigkeit besteht, dann ist es häufig auch unwesentlich, ob $E1.n$ kurz vor $E2.m$, gleichzeitig mit $E2.m$ oder kurz nach $E2.m$ eintritt.

Beispiel:

$E1.27$: ich schreibe zum 27. Mal einen Brief an meinen Geschäftspartner José in Südamerika.

$E2.25$: José schreibt zum 25. Mal an mich.

Da wir weit voneinander entfernt schreiben und nur Briefe austauschen, ist es praktisch irrelevant, ob José ein paar Minuten vor mir zu schreiben beginnt oder ein paar Minuten nach mir.

Die Auffassung $N=G$ führt dazu, dass die Unterscheidung zwischen nebenläufigen und nicht-nebenläufigen Ereignissen zumindest in manchen Fällen nicht praktisch relevant ist. Dagegen unterscheidet die Auffassung $N=U$ immer nur zwischen der Anwesenheit und der Abwesenheit von praktisch relevanten Abhängigkeiten.

Einwand 2: Es ist eine Erkenntnis der modernen Physik, dass "Gleichzeitigkeit" keine objektive Eigenschaft von Ereignissen ist, sondern auch vom Beobachter abhängt (von seinem physikalischen Zustand, nicht von seiner Willkür). D.h. zwei Ereignisse $E1.n$ und $E2.m$ können von *einem* Beobachter in der Reihenfolge "erst $E1.n$, dann $E2.m$ ", von einem *zweiten* Beobachter in der umgekehrten Reihenfolge "erst $E2.m$, dann $E1.n$ " und von einem *dritten* Beobachter gleichzeitig beobachtet werden, wobei alle drei Beobachter korrekt beobachten.

In vielen Fällen ist die Abhängigkeit der "Gleichzeitigkeit" von den Beobachtern so gering, dass man sie praktisch vernachlässigen kann. Andererseits nehmen die Anwendungsfälle zu, in denen dies nicht ohne weiteres möglich ist (z. B. wenn Signale über große Entfernungen via Satellit übertragen werden, oder wenn in neuartigen Schaltelementen mit extrem kurzen Signalen gearbeitet wird). Die Netztheorie hat den Anspruch und das Ziel, auch in diesen Fällen gültig und nützlich zu sein.

Einwand 3: Der zeitliche Abstand zwischen zwei Ereignissen kann beliebig klein sein. Durch Messungen kann man höchstens feststellen, dass er über bzw. nicht über einer gewissen Schranke liegt. Die Schranke ist durch das Messverfahren bestimmt. Praktisch heißt das: durch Messungen kann man höchstens zwischen "sicherlich nicht gleichzeitigen" und "möglicherweise gleichzeitigen" Ereignissen unterscheiden, aber nicht zwischen "gleichzeitigen" und "nicht-gleichzeitigen".

Dagegen kann man einwenden: Ein (durch die jeweilige Anwendung bestimmter) kleiner Zeitunterschied mache doch nicht so viel aus und kann in der Praxis ignoriert werden. Das ist in vielen Fällen sicher richtig, bedeutet aber, dass für diese Praxis kein fundamentaler Unterschied zwischen gleichzeitigen und nicht-gleichzeitigen Ereignissen (d.h. zwischen nebenläufigen und nicht-nebenläufigen Ereignissen entsprechend der Auffassung $N=G$) besteht bzw. bestehen sollte.

Zum Abschluss dieses Abschnitts sei der Unterschied zwischen der Auffassung $N=U$ und $N=G$ noch anhand eines (Gedanken-) Experiments erläutert.

Gegeben sei ein schwarzer Kasten mit 2 Lämpchen an der Oberseite. Zwei Ereignisse werden

angekündigt und treten dann ein:

E1.1: Lämpchen 1 leuchtet auf.

E2.1: Lämpchen 2 leuchtet auf.

Zu entscheiden ist die Frage: Traten die Ereignisse *nebenläufig* ein oder nicht?

Ein Vertreter der Auffassung $N=G$ muss mit einer Stoppuhr oder einer ähnlichen Messapparatur den zeitlichen Abstand zwischen E1.1 und E2.1 messen. Wenn dieser Abstand groß ist, dann kann er mit Sicherheit "nicht-nebenläufig" entscheiden. Ist der Abstand dagegen klein, dann muss er mit einer Fehlentscheidung aufgrund eines Messfehlers rechnen.

Ein Vertreter der Auffassung $N=U$ braucht zwar keine Messungen (des zeitlichen Abstands zwischen E1.1 und E2.1) durchzuführen, er muss aber den schwarzen Kasten öffnen und untersuchen. Findet er darin zwei elektrische Schaltungen, die unabhängig voneinander die beiden Lämpchen zum Aufleuchten gebracht haben, dann wird er mit Sicherheit "nebenläufig" entscheiden, auch dann, wenn zwischen den Ereignissen ein deutlicher Zeitunterschied zu beobachten war. Findet er dagegen, dass das eine Lämpchen eine Photozelle beleuchtet hat, durch die das andere Lämpchen angeschaltet wurde, dann wird er mit Sicherheit "nicht-nebenläufig" entscheiden, auch dann, wenn keine Messapparatur zur Verfügung steht, mit der man einen zeitlichen Abstand zwischen den beiden Aufleucht-Ereignissen feststellen kann.

Auch ein Vertreter der Auffassung $N=U$ wird möglicherweise eine falsche Entscheidung bezüglich der Nebenläufigkeit der beiden Ereignisse treffen. Es könnte z. B. sein, dass seine elektrotechnischen Kenntnisse mangelhaft sind und er deshalb die Schaltung(en) in dem schwarzen Kasten missversteht.

Aber sein Fehler wäre kein Messfehler, sondern ein Fehler anderer Art. Entsprechend werden sich Maßnahmen zur Vermeidung solcher Fehler von Maßnahmen zur Minimierung von Messfehlern (bzw. zur Vermeidung von Fehlentscheidungen aufgrund von Messergebnissen) grundsätzlich unterscheiden.

4 Prozessnetze und Zeit: zwei unabhängige und frei kombinierbare Konzepte

Prozessnetze beschreiben Abhängigkeiten zwischen Netz-Elementen. Direkt sagt ein Prozessnetz nichts über die *zeitlichen* Aspekte der dargestellten Bedingungen und Ereignisse aus. Nur indirekt kann man gewisse Schlüsse ziehen: wenn ein Element X2 abhängig ist von einem Element X1, dann kann X2 nicht vor X1 passieren (d.h. eintreten bzw. gültig werden, je nachdem, ob es sich um eine Bedingung oder ein Ereignis handelt). Sind X1 und X2 dagegen nebenläufig (d.h. unabhängig voneinander), dann sagt das Prozessnetz allein nichts darüber aus, ob X1 vor X2, gleichzeitig mit X2 oder nach X2 passiert. In diesem Sinn abstrahiert ein Prozessnetz "so weit wie möglich" von allen zeitlichen Eigenschaften und Messgrößen.

Gerade deshalb ist es besonders einfach, Netze mit verschiedenen "zeitlichen Vorstellungen und Größen" zu kombinieren. Z. B. kann man frei wählen, ob man nur zeitlich ausgedehnte oder nur zeitlich "punktförmige" Ereignisse oder beide Arten zugrunde legen will. E1: "Ich schreibe einen Brief" ist ein Ereignis der ersten Art, es dauert eine gewisse Zeit an. E2: "Ich beginne einen Brief zu schreiben" kann dagegen als zeitlich punktförmig aufgefasst werden. Offensichtlich kann man jedes zeitlich ausgedehnte Ereignis durch zwei zeitlich punktförmige Ereignisse und eine Bedingung dazwischen darstellen.

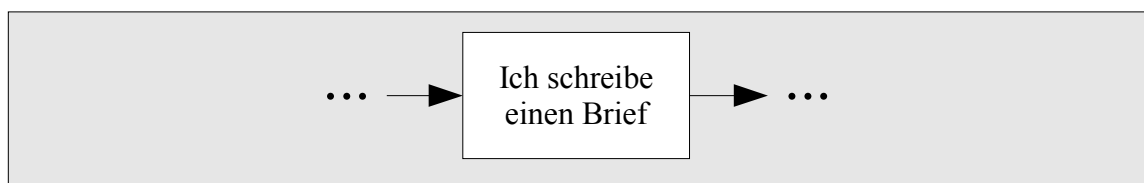


Abbildung 4.1: Ein zeitlich ausgedehntes Ereignis

Ganz entsprechend kann man festlegen, ob man zeitlich ausgedehnte Bedingungen oder zeitlich punktförmige Bedingungen beschreiben will.

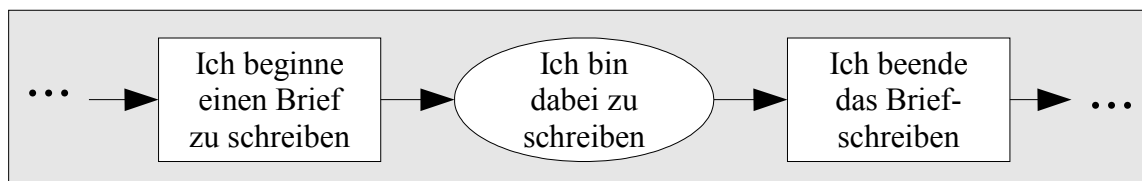


Abbildung 4.2: Eine zeitlich punktförmige Bedingung

Entsprechend diesen Festlegungen kann man dann Prozessnetze mit zeitlichen Angaben anreichern. Z. B. kann man zeitlich ausgedehnte Netz-Elemente mit ihrem Anfangs- und ihrem End-Zeitpunkt, oder mit ihrem Anfangszeitpunkt und ihrer Dauer, oder nur mit ihrer Dauer etc. beschriften. Zeitlich punktförmige Netz-Elemente kann man z. B. mit dem *Zeitpunkt*, an dem sie passieren, oder mit einem *Zeitintervall*, innerhalb dessen ihr Zeitpunkt liegt (gelegen hat bzw. liegen soll) beschriften.

Natürlichen sollten die Zeitbeschriftungen eines Netzes in sich widerspruchsfrei sein. Z. B. sollte der

Endzeitpunkt eines Elements nicht vor seinem Anfangszeitpunkt liegen. Außerdem sollten die Zeitangaben mit der Netz-Struktur, d.h. mit den Abhängigkeiten zwischen den Elementen, verträg-

lich sein. Ist z. B. X2 (direkt oder indirekt) abhängig von X1. so sollte aus den Zeitangaben nicht hervorgehen, dass X2 zeitlich vor X1 passiert ist bzw. passieren wird oder passieren soll. Unabhängig von den Einzelheiten der Zeitangaben gilt darüber hinaus (oder: sollte darüber hinaus gelten):

1. Die Netz-Struktur ist primär, die Zeitangaben sind sekundär.
2. Die Netz-Struktur kann unabhängig von den Zeitangaben betrachtet und untersucht werden.
3. Folgt aus einem Netz, dass ein Element X2 zeitlich *nach* einem Element X1 passiert, so kann man feststellen, ob dies "nur aus zeitlichen Gründen" oder aus Gründen einer "tieferen Abhängigkeit" so ist.

5 Prozessnetze beschreiben Prozesse objektiv, d.h. unabhängig von speziellen Beobachtungen

Als Ergänzung zum Begriff eines Prozessnetzes gibt es einen sehr anschaulichen und intuitiven Begriff einer Beobachtung (eines Prozesses). Eine Beobachtung ist eine Folge von "Momentaufnahmen" des beobachteten Prozesses. Anstelle der Bezeichnung "Momentaufnahme" wird in der Netztheorie der technische Begriff "Schnitt" (englisch: cut) verwendet.

Ein Schnitt durch ein Prozessnetz PN ist eine maximale Menge von paarweise unabhängigen (d.h. nebenläufigen) Netz-Elementen. Ein Schnitt kann Bedingungen und Ereignisse enthalten. Ein B-Schnitt ist ein Schnitt, der nur Bedingungen enthält. Ein E-Schnitt enthält entsprechend nur Ereignisse.

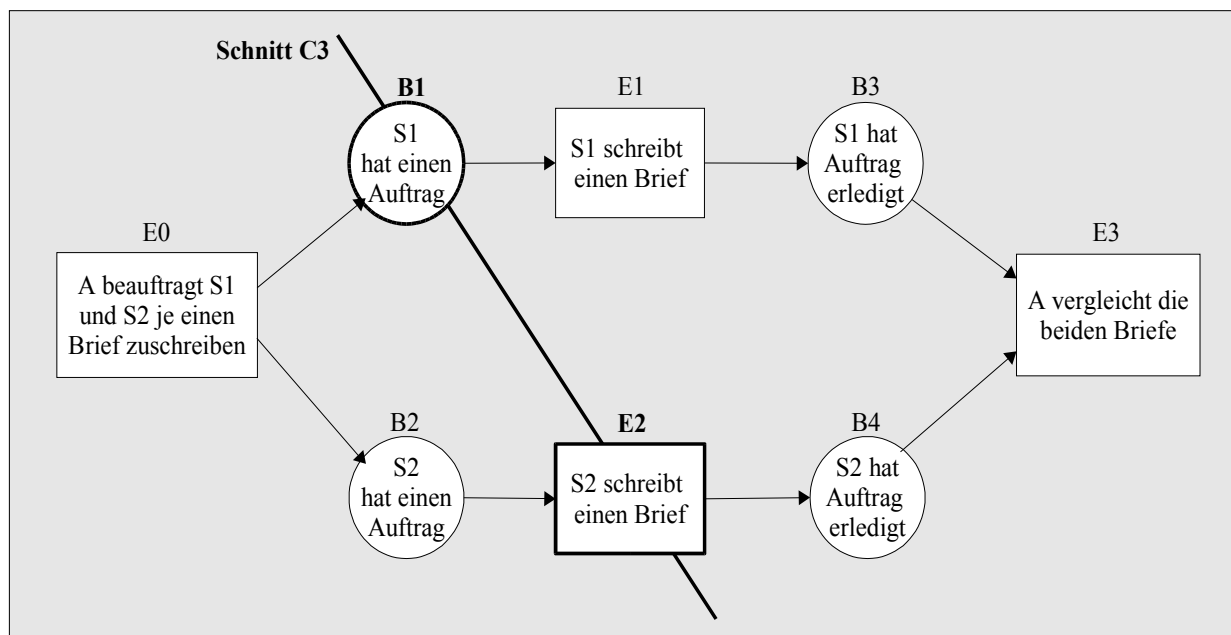


Abbildung 5.1: Ein Schnitt durch ein Prozessnetz

Der in Abb. 5.1 dargestellte Prozess läuft in einem Büro ab. Der Schnitt $C3 = \{B1, E2\}$ zeigt den Prozess in dem Moment, in dem S1 von A schon beauftragt wurde, aber noch nicht zu schreiben begonnen hat, während S2 schon schreibt. C3 ist ein Schnitt (d.h. maximal), da es kein weiteres Netz-Element gibt, welches sowohl zu B1 als auch zu E2 nebenläufig ist.

Die folgende Abbildung zeigt das Netz aus Abb. 5.1 mit all seinen Schnitten:

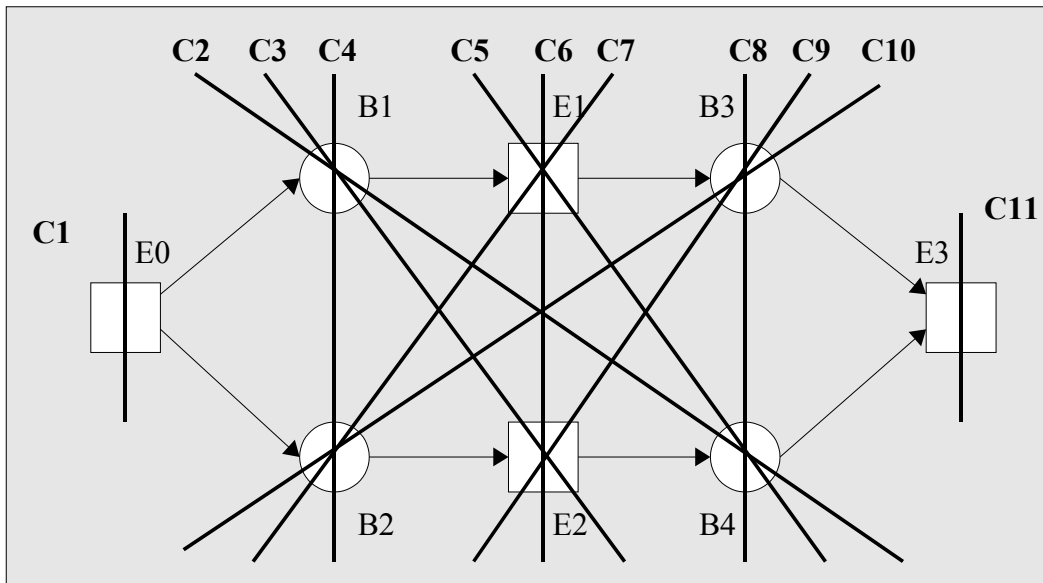


Abbildung 5.2: Das Prozessnetz und all seine Schnitte

Bestimmte Momentaufnahmen (Schnitte) kann man unmittelbar nacheinander machen. Z. B. sieht man unmittelbar nach dem Schnitt C3 entweder den Schnitt C2 oder den Schnitt C6 oder den Schnitt C5. C5 liegt unmittelbar nach C3, weil S1 in dem Moment zu schreiben beginnen kann, in dem S2 mit dem Schreiben aufhört.

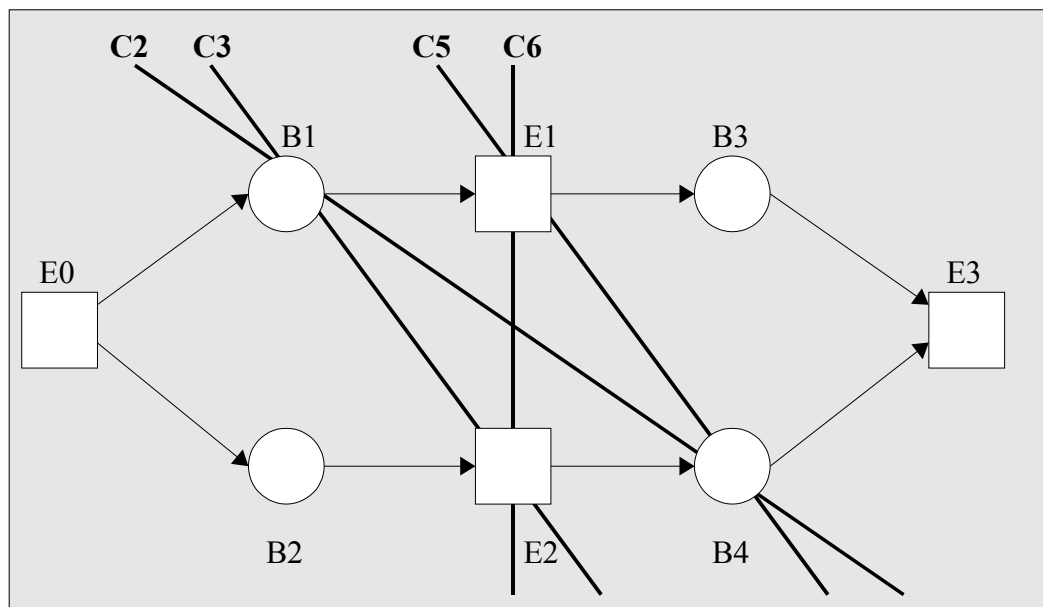


Abbildung 5.3: Nach C3 sieht man C2, C6 oder C5

Andere Momentaufnahmen (Schnitte) kann man zwar *nacheinander*, aber nicht *unmittelbar nacheinander* machen. Z. B. liegt zwischen dem Schnitt C3 und dem Schnitt C8 noch mindestens ein Schnitt zu dem das Ereignis E1 gehört:

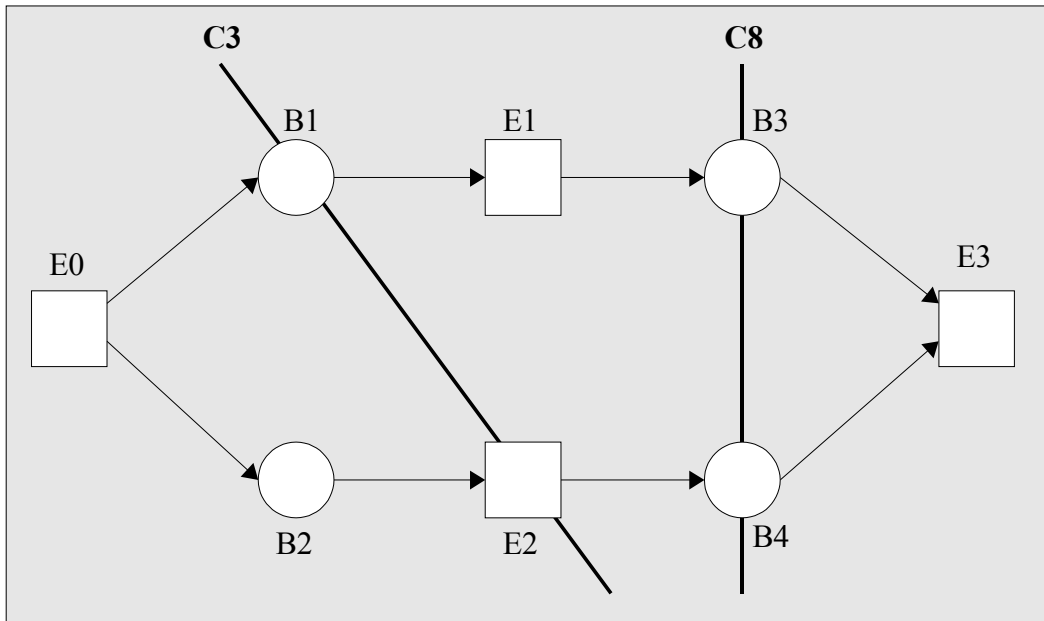


Abbildung 5.4: C3 und C8 liegen hintereinander, aber nicht unmittelbar

Noch andere Momentaufnahmen (Schnitte) kann man auf keinen Fall nacheinander machen. Z. B. liegt der Schnitt C3 nicht nach dem Schnitt C7, aber C7 liegt auch nicht nach C3:

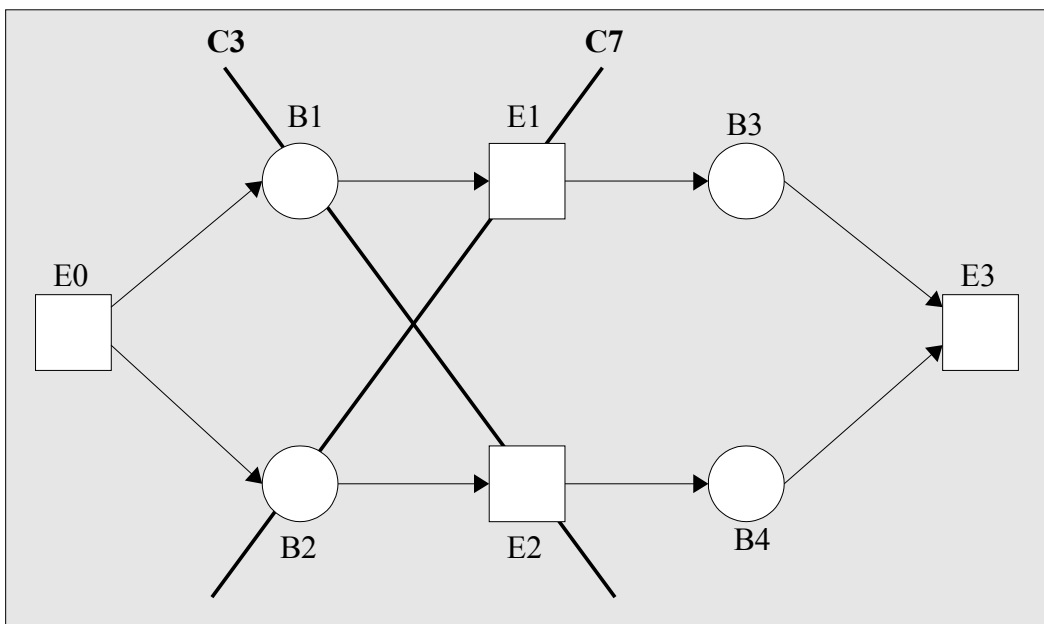


Abbildung 5.5: C3 liegt nicht hinter C7 und C7 liegt nicht hinter C3

In der folgenden Abbildung stellt jeder Knoten einen Schnitt durch das oben angegebene Prozessnetz dar. Schnitte, die man unmittelbar nacheinander "sehen kann" sind durch einen Pfeil verbunden:

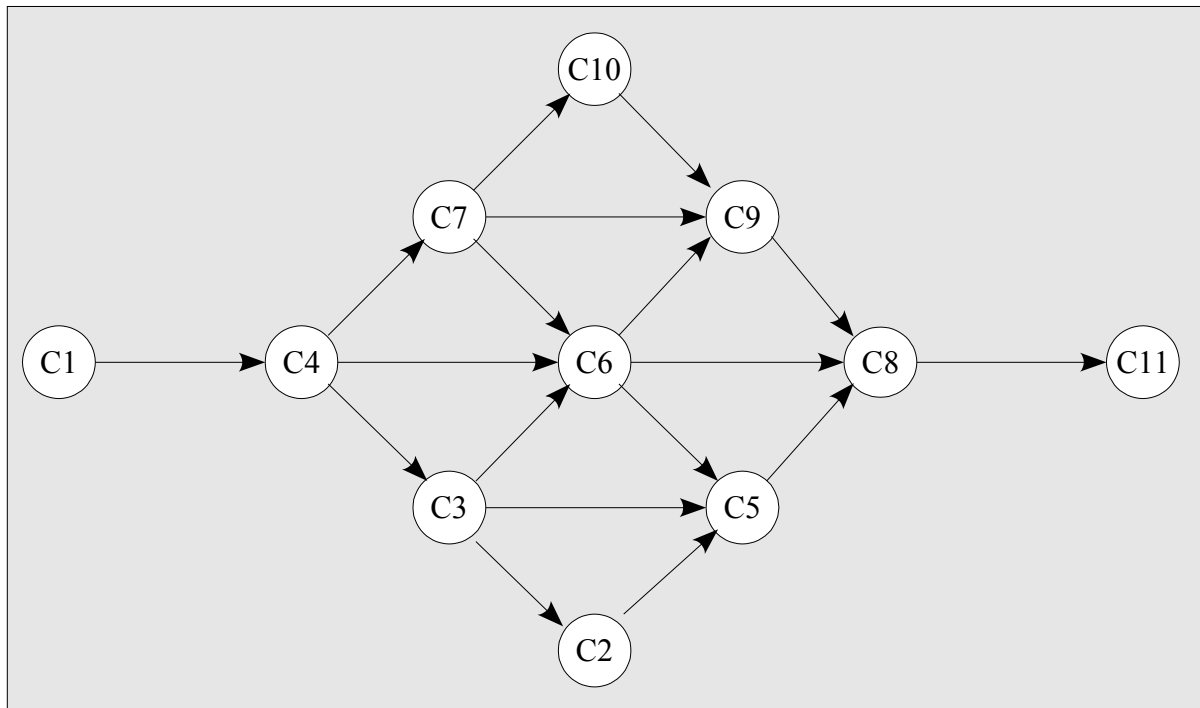


Abbildung 5.6: Ein Graph von Schnitten

Jeder Weg (entlang den Pfeilen) in einem Graph von Schnitten entspricht einer **Beobachtung** des betreffenden Prozesses. Jeder maximale Weg (den man weder vorn noch hinten verlängern kann) entspricht einer **vollständigen Beobachtung** des betreffenden Prozesses.

In Abb. 5.6 ist $O_1 = (C_4, C_7, C_9)$ eine Beobachtung, aber nicht vollständig. $O_2 = (C_1, C_4, C_7, C_9, C_8, C_{11})$ ist eine vollständige Beobachtung. Die kürzeste vollständige Beobachtung ist $O_3 = (C_1, C_4, C_6, C_8, C_{11})$.

Ein Prozessnetz sagt nichts aus über die Reihenfolge bzw. Gleichzeitigkeit seiner nebenläufigen Elemente. Eine (vollständige) Beobachtung zu einem Prozessnetz besagt, welche nebenläufigen Elemente des Netzes gleichzeitig und welche in einer bestimmten Reihenfolge beobachtet wurden. Eine Beobachtung enthält im Allgemeinen also auch Information, die nicht aus dem Prozessnetz stammt und die man als "subjektive Information des Beobachters" bezeichnen kann. Im Allgemeinen gibt es zu einem Prozessnetz viele verschiedene (vollständige) Beobachtungen, die sich alle durch ihre subjektive Information voneinander unterscheiden.

In diesem Sinn ist ein Prozessnetz abstrakter als eine (vollständige) Beobachtung und beschreibt einen Prozess objektiv, d.h. unabhängig von den verschiedenen möglichen Beobachtungen. Prozessnetze sind insbesondere mit der physikalischen Erkenntnis verträglich, dass die selben unwiederholbaren Ereignisse von verschiedenen Beobachtern in verschiedenen Reihenfolgen beobachtet werden können (siehe auch Einwand 2 im Abschnitt 3.).

6 Ereignisse sind atomar und sollten nicht in zwei Hälften gespalten werden

Sei E ein Ereignis in einem Bedingungs-Ereignis-Netz, und seien V_1, V_2, \dots, V_n die Vorbedingungen und N_1, N_2, \dots, N_m die Nachbedingungen von E .

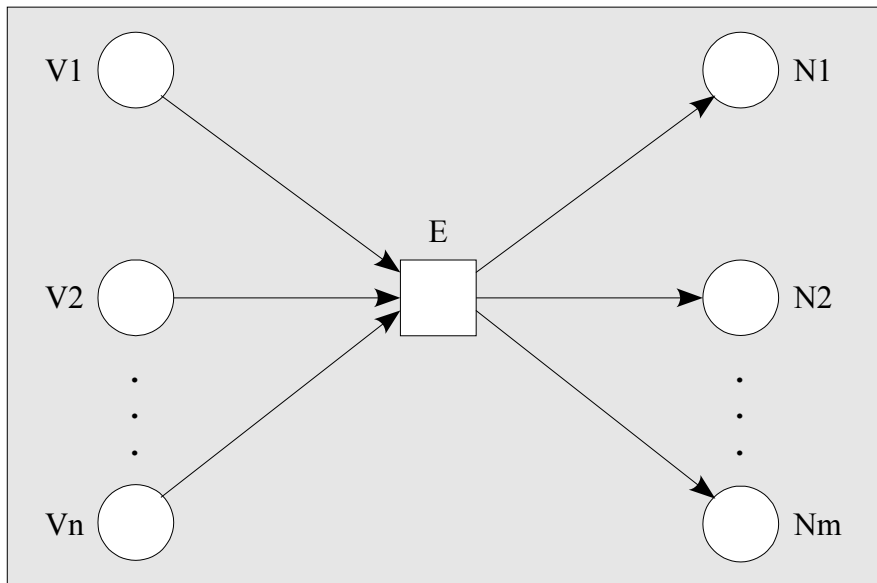


Abbildung 6.1: Ein Ereignis E

"Das Ereignis E tritt ein" bedeutet soviel wie:

- die Vorbedingungen V_i gehen von "gültig" zu "ungültig" über und
- die Nachbedingungen N_j gehen von "ungültig" zu "gültig" über.

Für das Durchspielen von Netzen haben sich sogenannte Token (z. B. kleine Münzen, Damesteine, Reißzwecken etc.) als sehr praktisch erwiesen. Ein Token auf einer Bedingung bedeutet: die Bedingung ist momentan gültig. Bedingungen ohne Token darauf sind momentan ungültig. Das Eintreten eines Ereignisses besteht dann darin, dass

- von jeder Vorbedingung V_i das dort liegende Token entfernt und
- auf jede Nachbedingung N_j ein Token gelegt wird.

Wenn man ein markiertes B/E-Netz von Hand ausführt, dann ist es ganz natürlich, daß man beim Eintreten eines Ereignisses E zuerst von jeder Vorbedingung V_i ein Token entfernt und dann auf jede Nachbedingung ein Token legt. Dagegen ist nichts einzuwenden.

Allerdings entwickelt sich aus dieser praktischen Vorgehensweise leicht die Vorstellung, dass ein Ereignis auch in einem tieferen Sinn aus zwei Teilereignissen besteht, die nacheinander eintreten: a) nimm-Token und b) gib-Token.

Diese Vorstellung widerspricht einer wichtigen Intuition der Netztheorie: dass ein Ereignis etwas atomares, unzerlegbares ist und somit das Ungültigwerden der Vorbedingungen V_i und das Gültigwerden der Nachbedingung N_j eine unauflösbare Einheit bilden. Diese Intuition ist schon in dem einfachen Beispiel der vier Jahreszeiten erkennbar (siehe Abschnitt 2, Abb. 2.1 und 2.2). Es ist wenig sinnvoll anzunehmen, dass z. B. zuerst der Winter aufhört und dann der Frühling anfängt. Vielmehr bilden das Aufhören des Winters und das Anfangen des Frühlings eine Einheit.

Das folgende Beispiel aus der Chemie soll ebenfalls diese Intuition illustrieren.

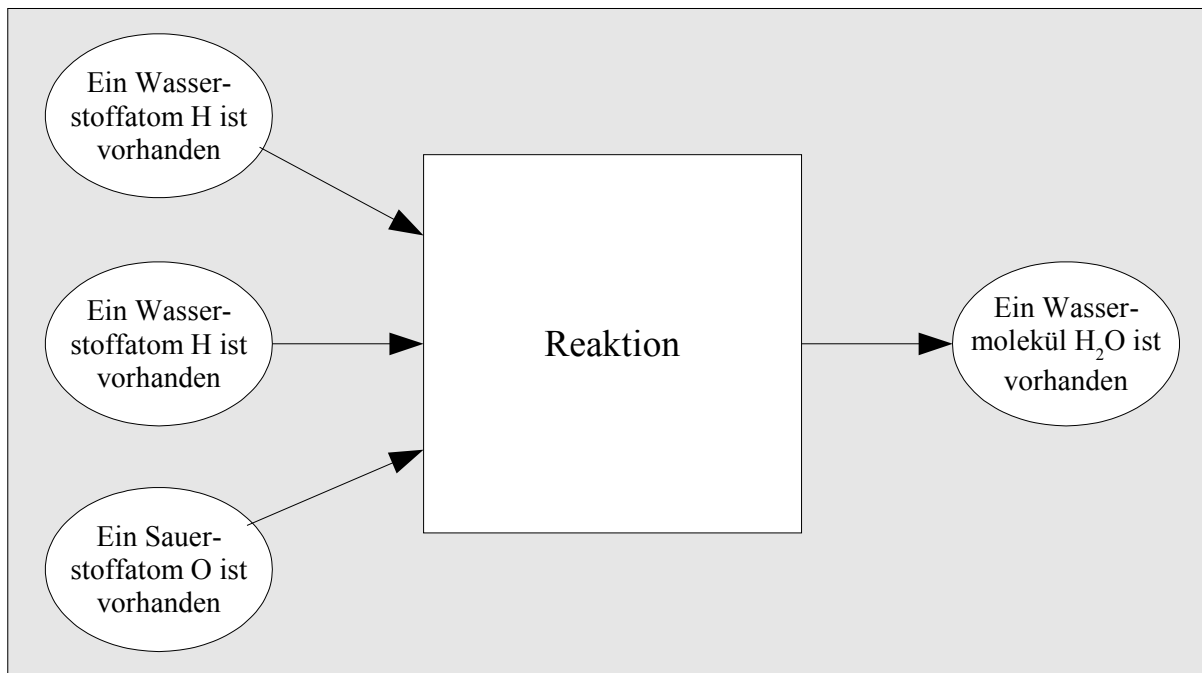


Abbildung 6.2: Ein paradigmatisches Ereignis

Damit ein Wassermolekül H_2O entstehen kann müssen zwei Wasserstoff-Atome H und ein Sauerstoff-Atom O vorhanden sein. Wenn diese drei Bedingungen (und einige weitere, hier nicht weiter wichtige Bedingungen) erfüllt sind, dann kann eine Reaktion eintreten. Das Reaktions-Ereignis besteht darin, dass drei Atome "verschwinden" und ein Molekül "erscheint".

Das Verschwinden der Atome und das Erscheinen des Moleküls sind nicht zwei Teilereignisse, die nacheinander ablaufen. Vielmehr erscheint das Molekül genau in dem Maße, in dem die Atome verschwinden. Oder: das Wassermolekül erscheint genau in dem Moment, in dem die Atome verschwinden, und nicht erst "im nächsten Moment". Diese Vorstellung ist unabhängig davon, ob man sich einen "Moment" als einen Zeitpunkt (ohne Dauer) oder als ein Zeitintervall vorstellt.

Wenn in einem bestimmten Anwendungsfall ein Ablauf dargestellt werden soll, bei dem zuerst gewisse (Vor-) Bedingungen ungültig und dann gewisse (Nach-) Bedingungen gültig werden, dann sollte man diesen Ablauf nicht durch **ein** Ereignis E darstellen (wie im linken Teil der folgenden Abb. 6.3), sondern durch **zwei** Ereignisse E1 und E2 und eine Bedingung B (wie im rechten Teil von Abb. 6.3):

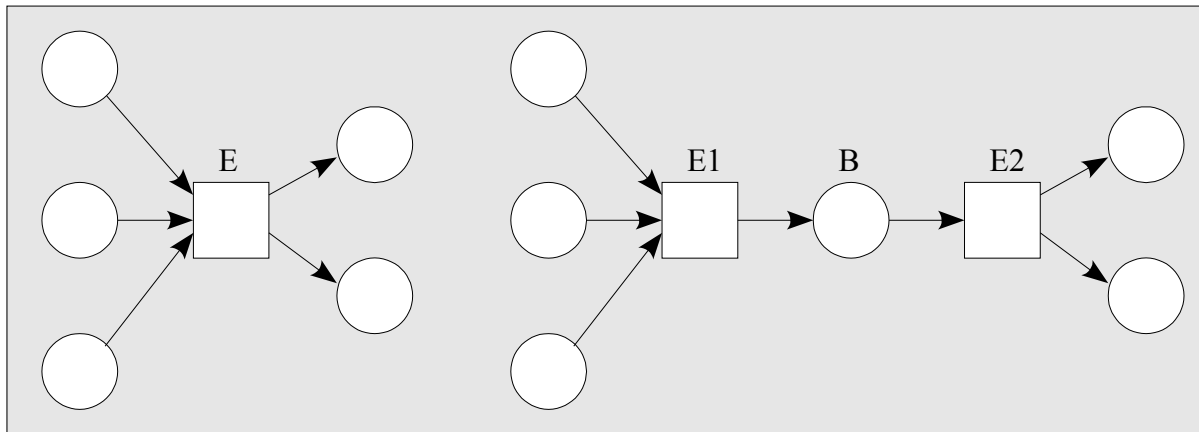


Abbildung 6.3: Zerlegung eines Ereignisses E in zwei Ereignisse E1 und E2

D. h. wenn man atomare Ereignisse als Grundlage nimmt, dann kann man daraus leicht "molekulare" (zusammengesetzte) Abläufe bilden.

Ob man sich ein Ereignis als eine atomare Einheit vorstellt oder als eine Folge von zwei Teilereignissen (nimm-Token, gib-Token) macht einen wesentlichen Unterschied, wenn man sogenannte **Schlingen** betrachtet.

Schlingen können nicht nur in B/E-Netzen, sondern auch in Stellen-Transitions-Netzen, in farbigen Netzen, Prädikat-Transitions-Netzen etc. vorkommen bzw. explizit verboten werden. Eine Schlinge besteht immer aus einem runden Knoten, einem eckigen Knoten und zwei Pfeilen: einem vom runden zum eckigen und einem vom eckigen zum runden Knoten.



Abbildung 6.4: Eine Schlinge

Das Eintreten von E besteht darin, dass

- a) ein Token von B genommen und
- b) ein Token auf B gelegt wird.

Geht man davon aus, dass zuerst a) und dann b) geschieht, dann kann E genau dann eintreten, wenn ein Token auf B liegt.

Die Reihenfolge "zuerst a) und dann b)" ist aber nicht natürlicher als die umgekehrte Reihenfolge. Geht man davon aus, daß zuerst b) und dann a) geschieht, dann kann E genau dann eintreten, wenn kein Token auf B liegt.

Also gilt: wenn man davon ausgeht, dass a) und b) nacheinander geschehen, dann kann E eintreten. Je nachdem, welche Reihenfolge von a) und b) man annimmt, kann E eintreten, wenn auf B ein Token liegt bzw. wenn auf B kein Token liegt.

Wenn man aber von der grundlegenden netztheoretischen Vorstellung ausgeht, dass das Nehmen und Geben der Token beim Eintreten eines Ereignisses eine Einheit bilden, dann kann das Ereignis E in Abb. 6.4 in keinem der beiden Fälle eintreten. Denn wenn auf B ein Token läge, dann müsste es zuerst entfernt werden, damit Platz vorhanden ist, um auf B ein Token zu legen. Läge dagegen kein Token auf B, dann müsste zuerst eines hingelegt werden, damit es danach weggenommen werden kann. In beiden Fällen könnte das Nehmen und Geben nur nacheinander und nicht als unzerstrennliche Einheit geschehen.

7 Die Schaltregel für einzelne Ereignisse sagt nichts über Nebenläufigkeit

Bedingungs-Ereignis-Netze, Stellen-Transitions-Netze, farbige Netze etc. sind verschiedene Klassen (oder: Arten) von Petri-Netzen. Für jede Klasse von Netzen sind u.a. zwei "Regeln" definiert:

1. Die **Aktivierungs-Regel**. Sie gibt an, unter welchen Umständen ein eckiger Netzknoden (d.h. ein Ereignis, eine Transition etc.) aktiviert ist, d.h. schalten kann.
2. Die **Schalt-Regel**. Sie gibt an, was im Einzelnen passiert, wenn ein eckiger Knoten schaltet. Es ist eine einfache, aber leicht zu übersehende Tatsache, dass diese beiden Regeln die Semantik der betreffenden Netz-Klasse nicht vollständig beschreiben, da sie nichts darüber sagen, welche Schaltvorgänge nebenläufig eintreten können.

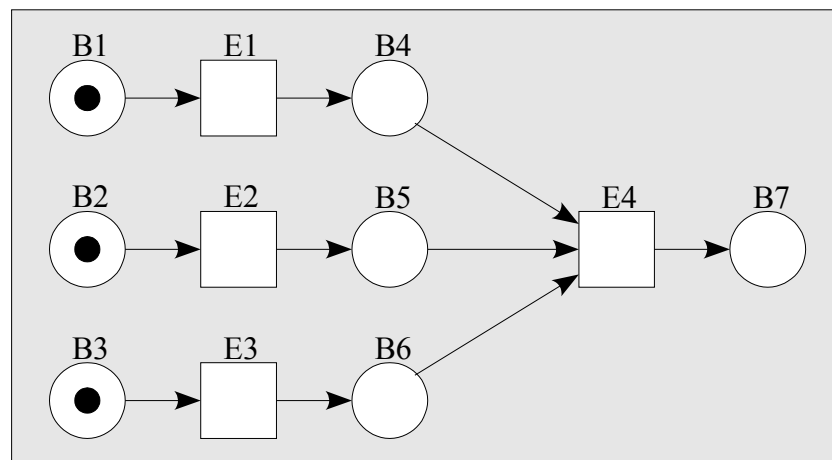


Abbildung 7.1: Ein markiertes B/E-Netz

Die Aktivierungs-Regel für B/E-Netze besagt, dass das Ereignis E1 (in Abb. 7.1) aktiviert ist. Die Schalt-Regel legt fest, dass nach dem Eintreten (Schalten) von E die Bedingung B1 ungültig (ohne Token) und die Bedingung B4 gültig (mit einem Token markiert) ist. Entsprechendes gilt für E2 und E3.

Es widerspricht aber weder der Aktivierungs- noch der Schalt-Regel, wenn man die zusätzliche Regel einführt: Ereignisse müssen stets nacheinander eintreten. Diese Regel schließt jegliche Nebenläufigkeit aus. Sie könnte z. B. dadurch motiviert sein, dass man das markierte Netz von einem sequentiellen Prozessor (Mensch oder Maschine) ausführen lassen will.

Hat man dagegen zwei solche Prozessoren, so könnte man festlegen: es dürfen jeweils höchstens zwei Ereignisse nebenläufig eintreten. Auch diese Festlegung widerspricht weder der Aktivierungs- noch der Schaltregel für B/E-Netze.

Der Aspekt der Nebenläufigkeit (der Semantik einer bestimmten Netz-Klasse) wird häufig mit Hilfe der Begriffe **Konflikt** und **Schritt** definiert. Ein Konflikt liegt vor, wenn zwei eckige Knoten *aktiviert* sind, aber *nicht* nebenläufig schalten können.

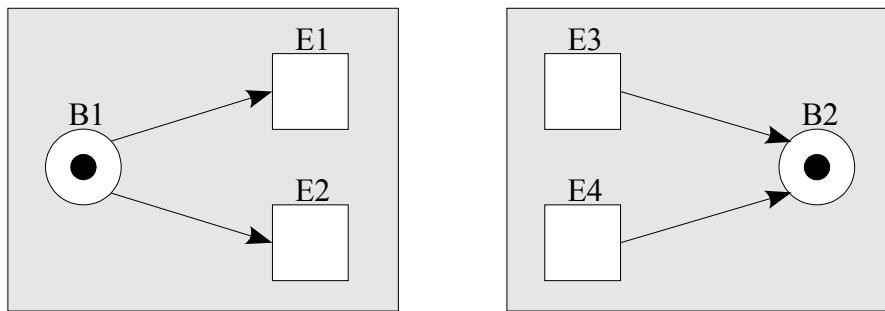


Abbildung 7.2: Zwei typische Konflikte in B/E-Netzen

Ein Schritt ist eine Menge bzw. eine Multimenge von eckigen Knoten (in einer Menge von Knoten kann jeder Knoten höchstens einmal vorkommen, z. B. $M = \{E1, E3\}$). In einer Multimenge von Knoten können Knoten auch mehrfach vorkommen, z. B. $MM = \{E1, E1, E2, E3, E3\}$). Ein Schritt ist aktiviert, wenn kein Knoten des Schritts mit einem anderen Knoten des Schritts in Konflikt steht. In Abb. 7.1 ist z. B. der Schritt $\{E1, E2, E3\}$ aktiviert.

Ein *aktivierter Schritt* umfasst genau solche eckigen Knoten, die nebenläufig zueinander schalten können.

Für Bedingungs-Ereignis-Netze gibt es eine allgemein anerkannte und verwendete Definition des Begriffs "aktivierter Schritt". Für Stellen-Transitions-Netze, farbige Netze und Prädikat-Transitions-Netze gibt es dagegen keine so einheitliche Definition des Nebenläufigkeits-Aspekts. Vielmehr sind für diese Netz-Klassen verschiedene Konflikt- und Schritt-Begriffe möglich und sinnvoll, und man muss in verschiedenen Zusammenhängen herausfinden bzw. festlegen, welcher davon gemeint ist. Im Abschnitt 16 wird noch genauer auf dieses Problem eingegangen.

8 Die Schritt-Semantik ist nur die zweitbeste Semantik für B/E-Netze

Die Aktivierungs-Regel für B/E-Netze gibt an, unter welchen Umständen ein einzelnes Ereignis E aktiviert ist (nämlich, wenn alle Vorbedingungen von E erfüllt und alle Nachbedingungen von E nicht-erfüllt sind).

Die Schalt-Regel für B/E-Netze gibt an, was passiert, wenn ein einzelnes Ereignis E eintritt (die Vorbedingungen werden nicht-erfüllt und die Nachbedingungen werden erfüllt).

Zusammen werden die Aktivierungs- und die Schalt-Regel hier als Schalt-Semantik für B/E-Netze bezeichnet. Diese Semantik ist unvollständig, da sie nicht festlegt, unter welchen Umständen mehrere Ereignisse nebenläufig eintreten können (siehe dazu auch Abschnitt 7).

Im Zusammenhang mit B/E-Netzen ist ein *Schritt* eine Menge von Ereignissen (des betreffenden Netzes). Die Schritt-Semantik ist eine Verallgemeinerung der Schalt-Semantik. Sie besteht aus einer Aktivierungs-Regel für Schritte (Unter welchen Umständen ist ein Schritt aktiviert? Und einer Schalt-Regel für Schritte (was passiert im einzelnen, wenn ein Schritt ausgeführt wird?).

Im Gegensatz zur *Schalt-Semantik* klärt die *Schritt-Semantik* auch den Nebenläufigkeits-Aspekt von markierten Netzen: zwei Ereignisse sind nebenläufig, wenn sie "in einem Schritt" eintreten.

Trotzdem gibt es Gründe dafür, die Schritt-Semantik nicht als die richtige Semantik für B/E-Netze, sondern nur als "zweitbeste Semantik" zu betrachten. Die "beste" Semantik ist die Prozessnetz-Semantik, die einem markierten B/E-Netz ein Prozessnetz (genauer: eine Menge von Prozessnetzen) als Bedeutung zuordnet.

Die Gründe gegen die Schritt-Semantik und für die Prozess-Semantik sind nicht so sehr von harter, mathematischer Art. Sie haben mehr mit den Intuitionen hinter den Begriffen zu tun und mit den Zielen und Absichten, die mit der Theorie der Petri-Netze verfolgt werden.

Wenn man ein markiertes B/E-Netz schrittweise ausführt, dann erhält man eine Folge von "globalen Zuständen" und "globalen Übergängen". Dabei ist ein globaler Zustand eine Menge von Bedingungen (die Bedingungen, die bei in diesem Zustand erfüllt sind) und ein globaler Übergang ist eine Menge von Ereignissen (die Ereignisse, die bei diesem Übergang eintreten). Einzelne Bedingungen und Ereignisse werden in diesem Zusammenhang häufig auch als "lokale Zustände" bzw. als "lokale Übergänge" bezeichnet.

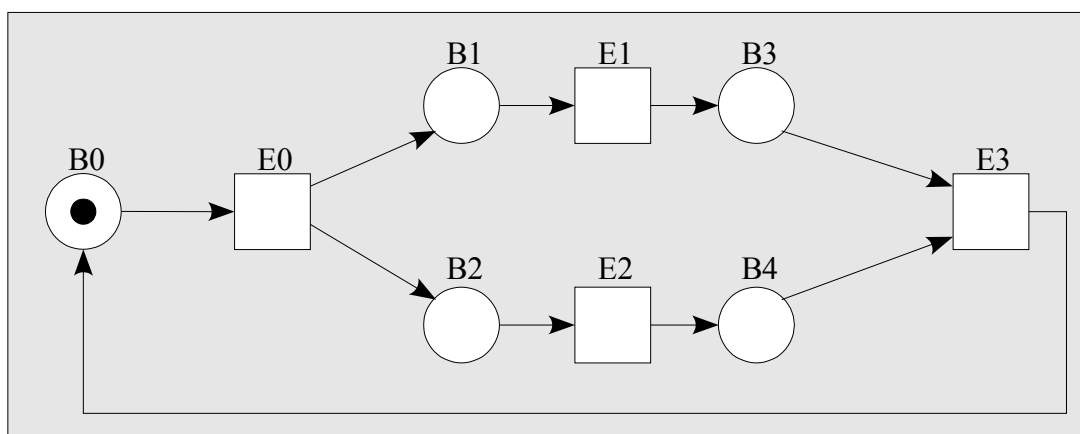


Abbildung 8.1: Ein markiertes B/E-Netz N

Eine schrittweise Ausführung des markierten B/E-Netzes in Abb. 8.1 kann z. B. aus folgenden globalen Zuständen und Übergängen bestehen:

$Z_0 = \{B_0\}$	$\ddot{U}_0 = \{E_0\}$	$Z_1 = \{B_1, B_2\}$	$\ddot{U}_1 = \{E_1, E_2\}$
$Z_2 = \{B_3, B_4\}$	$\ddot{U}_2 = \{E_3\}$	$Z_3 = \{B_0\}$	$\ddot{U}_3 = \{E_0\}$
$Z_4 = \{B_1, B_2\}$	$\ddot{U}_4 = \{E_1\}$	$Z_5 = \{B_3, B_2\}$	$\ddot{U}_5 = \{E_3\}$
$Z_6 = \{B_3, B_4\}$	$\ddot{U}_6 = \{E_3\}$	$Z_7 = \{B_0\}$...

Die Schritt-Semantik macht aus einem markierten B/E-Netz in erster Linie ein sequentielles System (oder: einen endlichen Automaten), denn als Bedeutungen von Netzen werden **Folgen** von Zuständen und Übergängen betrachtet. Der Aspekt der Nebenläufigkeit wird dagegen in der Feinstruktur der (globalen) Zustände und Übergänge versteckt: zwei Ereignisse treten nebenläufig ein, wenn sie zum selben globalen Übergang gehören; zwei Bedingungen sind nebenläufig erfüllt, wenn sie zum selben globalen Zustand gehören.

In vielen Anwendungen ist es möglich und sinnvoll, den sequentiellen Aspekt eines Systems zu betonen und den Aspekt der Nebenläufigkeit zu vernachlässigen. Für solche Anwendungen ist das Konzept eines endlichen Automaten bzw. eines Netzes mit Schritt-Semantik angemessen.

Je stärker aber der Aspekt der Nebenläufigkeit in den Vordergrund tritt, desto unangemessener wird es, den Ablauf eines Systems als Folge von globalen Zuständen und Übergängen zu modellieren.

Als Beispiel stelle man sich ein Telefonsystem vor, z. B. das der Telecom. Selbst von diesem System kann man zumindest annehmen, dass sein Ablauf aus einer Folge von globalen Zuständen und Übergängen besteht. Für diese Zustände und Übergänge gilt dann aber:

1. Man kann sie nicht gezielt beobachten. Man kann z. B. nicht feststellen, ob sich das Telefonsystem in einem bestimmten Moment im "Ruhezustand" befindet, in dem niemand in Deutschland telefoniert.
2. Man kann sie nicht gezielt hervorrufen. Man kann das Telefonsystem kaum in den Ruhezustand (oder in irgend einen anderen genau definierten globalen Zustand) versetzen.

Globale Zustände und Übergänge (von nebenläufigen Systemen) werden manchmal auch als *Gespenster* bezeichnet, da man Gespenster bekanntlich auch nicht gezielt beobachten oder "hervorrufen" kann.

Die Prozessnetz-Semantik für markierte B/E-Netze vermeidet die Annahme von globalen Zuständen und Übergängen. Sie ordnet einem markierten B/E-Netz eine Menge von Prozessnetzen als Bedeutung zu. Wenn bei der Ausführung eines markierten B/E-Netzes N keine Konflikte auftreten können, dann hat N nur ein Prozessnetz PN als Bedeutung. Können dagegen bei der Ausführung von N Konflikte auftreten, dann hat N mehrere (evt. unendlich viele) Prozessnetze als Bedeutung.

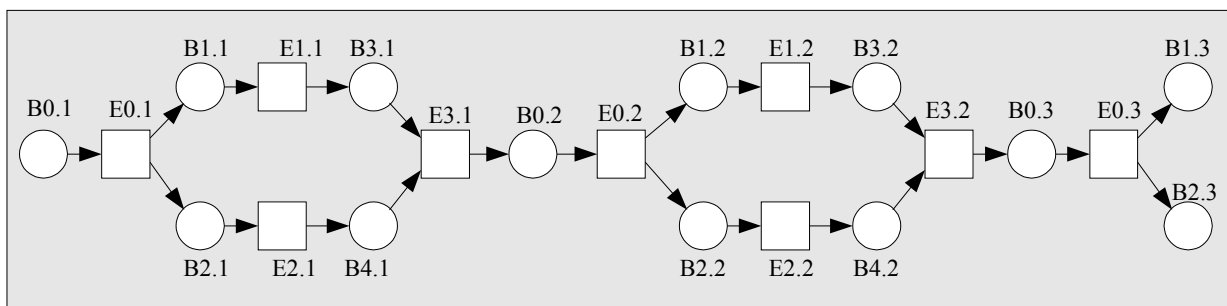


Abbildung 8.2: Das Prozessnetz PN zum Netz N in Abb. 8.1

Im Abschnitt 5 wurden im Zusammenhang mit Prozessnetzen "Momentaufnahmen" (Schnitte) und Beobachtungen (Folgen von Schnitten) erwähnt. Im allgemeinen kann ein Schnitt durch ein Pro-

zess-Netz sowohl (unwiederholbare) Bedingungen als auch (unwiederholbare) Ereignisse enthalten. Ein B-Schnitt (bzw. ein E-Schnitt) ist ein Schnitt, der nur Bedingungen (bzw. nur Ereignisse) enthält.

Sei PN ein Prozessnetz zu einem markierten B/E-Netz N. Dann entspricht jeder B-Schnitt durch PN einem *globalen Zustand* von N und jeder E-Schnitt durch PN entspricht einem *globalen Übergang* (bei der Ausführung) von N.

Beispiel: Die B-Schnitte $C1 = \{B0.1\}$, $C2 = \{B1.1, B2.1\}$ und $C3 = \{B1.1, B4.1\}$ durch das Prozess-Netz PN in Abb. 8.2 entsprechen je einem globalen Zustand des markierten Netzes N in Abb. 8.1. Die E-Schnitte $C4 = \{E0.1\}$ und $C5 = \{E1.1, E2.1\}$ durch PN entsprechen globalen Übergängen von N.

Globale Zustände und Übergänge (eines markierten B/E-Netzes N) sind nichts anderes als spezielle "Momentaufnahmen" (Schnitte durch ein Prozessnetz PN von N). Jede schrittweise Ausführung (eines markierten B/E-Netzes N) entspricht genau einer Beobachtung (eines Prozessnetzes PN von N).

Es gibt aber bestimmte "Momentaufnahmen", die keine globalen Zustände oder Übergänge sind, nämlich Schnitte, die sowohl Bedingungen als auch Ereignisse enthalten). Entsprechend gibt es Beobachtungen, denen keine schrittweise Ausführung entspricht.

Ein Prozessnetz PN beschreibt einen Prozeß unabhängig von speziellen Beobachtungen (siehe Abschnitt 5). Ein markiertes B/E-Netz N unter der Prozessnetz-Semantik beschreibt Prozesse allein anhand von lokalen Zuständen und Übergängen und unabhängig davon, ob es globale Zustände und Übergängen überhaupt gibt oder nicht.

9 Kontakte stören die Prozessnetz-Semantik, aber nur ein bisschen

In einem markierten B/E-Netz hat ein Ereignis E **Kontakt**, wenn alle seine Vorbedingungen erfüllt sind und es trotzdem nicht eintreten kann, weil nämlich auch (mindestens) eine seiner Nachbedingungen erfüllt ist.

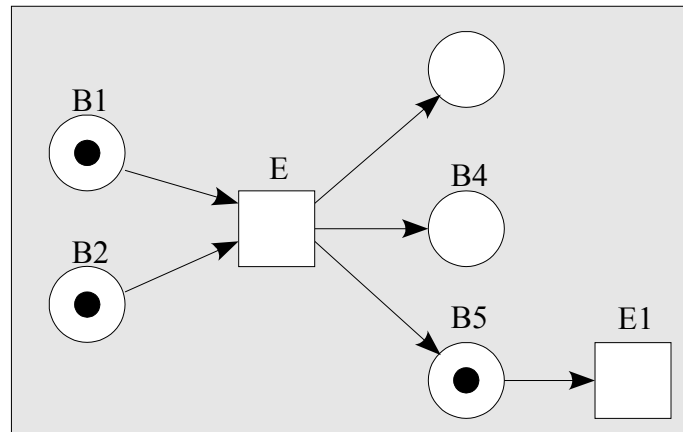


Abbildung 9.1: E hat Kontakt

In Abb. 9.1 gilt: Obwohl B1 und B2 markiert sind, kann E nicht eintreten, weil auch B5 markiert ist. Wenn E1 eintritt, dann wird die Kontakt-Situation beseitigt und E kann eintreten. In diesem Sinne ist E abhängig von E1.

Konstruiert man zu einem markierten B/E-Netz ein Prozessnetz, so werden darin solche Abhängigkeiten (zwischen einem Ereignis mit Kontakt und einem Ereignis, das den Kontakt beseitigt) **nicht** deutlich gemacht. Das folgende Beispiel illustriert diese Tatsache.

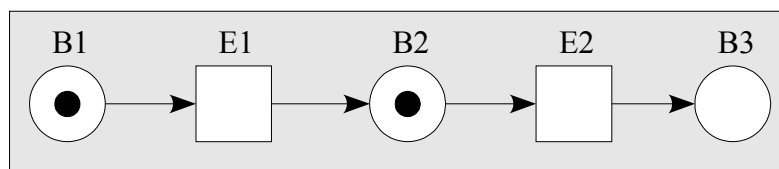


Abbildung 9.2: Ein markiertes B/E-Netz N. E1 hat Kontakt.

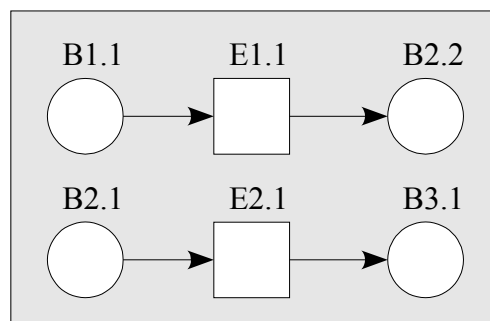


Abbildung 9.3: Das Prozessnetz PN zu N

Wenn man nur das Prozessnetz PN in Abb. 9.3 betrachtet, dann sieht es so aus, als könnten die (unwiederholbaren) Ereignisse E1.1 und E2.1 nebenläufig eintreten, denn PN zeigt keine Abhängigkeit zwischen diesen Ereignissen.

Betrachtet man dagegen das markierte B/E-Netz N in Abb. 9.2, dann wird klar: E1 hat Kontakt. Also muss zuerst E2 eintreten. Erst danach kann E1 eintreten.

Im folgenden Diagramm steht jeder Knoten für einen Schnitt durch das Prozessnetz PN. Schnitte, die man unmittelbar nacheinander beobachten kann, sind durch Pfeile verbunden (siehe auch Abschnitt 5):

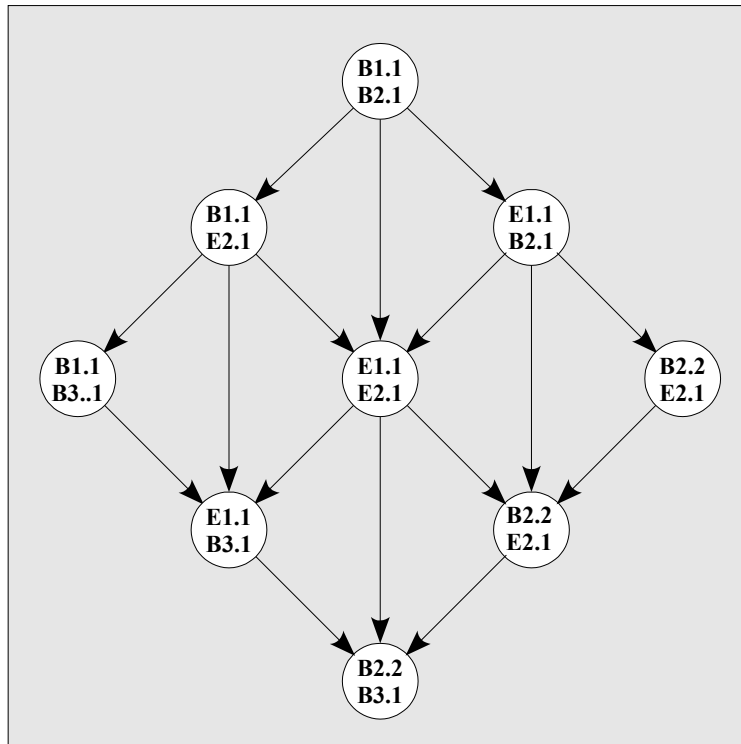


Abbildung 9.4: Der Graph von Schnitten durch PN

Könnte bei einer Ausführung des Netzes N kein Kontakt auftreten, dann würde jeder Weg durch diesen Graphen einer (vollständigen) Beobachtung des dargestellten Prozesses entsprechen. Da das Netz N aber kontaktbehaftet ist, entsprechen einige Wege durch diesen Graphen keiner möglichen Beobachtung.

Von einem der Schnitte durch das Prozessnetz PN kann man unmittelbar erkennen, dass er "unmöglich" ist, nämlich von dem Schnitt $C = \{B2.1, B2.2\}$. Dieser Schnitt enthält zwei Vorkommnisse der selben (wiederholbaren) Bedingung B2. Eine Bedingung in einem B/E-Netz kann aber nicht nebenläufig zu sich selbst erfüllt sein. Oder: wenn die Bedingung B2 zum ersten Mal erfüllt ist (B2.1), dann muss sie zuerst unerfüllt werden, damit sie ein zweites Mal erfüllt werden kann (B2.2). Ganz entsprechendes gilt für Ereignisse. Allgemein ist ein unmöglicher Schnitt einer, in dem eine (wiederholbare) Bedingung oder ein (wiederholbares) Ereignis *mehrfach* vorkommt.

Ersetzt man in einem unmöglichen Schnitt einige Knoten durch ihre Vorgängerknoten und/oder einige Knoten durch ihre Nachfolgerknoten, dann erhält man einen "verbotenen" Schnitt.

Ersetzt man z. B. in dem unmöglichen Schnitt $C = \{B2.2, B2.1\}$ (siehe dazu Abb. 9.3 und 9.4) $B2.2$ durch $E1.1$ und $B2.1$ durch $E2.1$, so erhält man den "verbotenen" Schnitt $C' = \{E1.1, E2.1\}$.

Entfernt man aus dem Schnitt-Graphen in Abb. 9.4 den unmöglichen Schnitt $C = \{B2.2, B2.1\}$ und alle verbotenen Schnitte, so erhält man folgenden reduzierten Schnitt-Graphen:

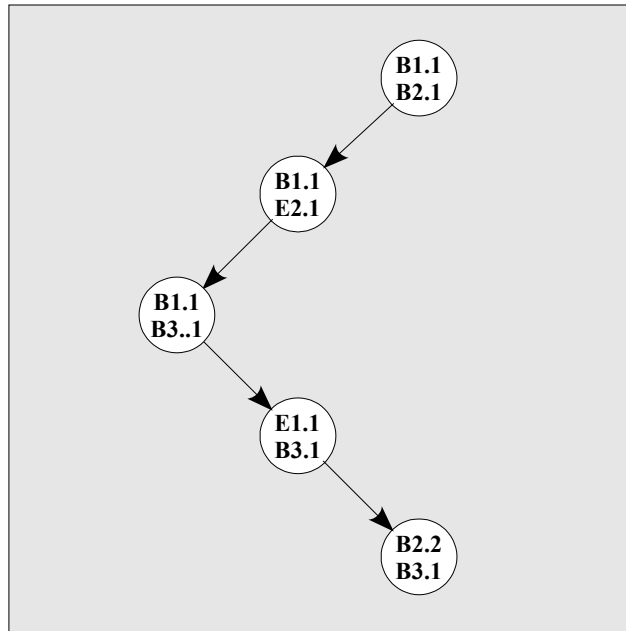


Abbildung 9.5: Der reduzierte Schnittgraph zu PN

Nur die Wege durch den reduzierten Schnitt-Graphen entsprechen möglichen Ausführungen des B/E-Netzes bzw. möglichen Beobachtungen des entsprechenden Prozessnetzes.

Zusammenfassung: Tritt bei der Ausführung eines markierten B/E-Netzes *Kontakt* auf, dann sind im entsprechenden Prozessnetz einige Knoten *unabhängig* voneinander, (d.h. sie liegen gemeinsam auf einem Schnitt) ohne wirklich *nebenläufig* zu sein. In solchen Prozessnetzen sind Knoten nur dann nebenläufig, wenn sie *auf einem Schnitt* liegen, der weder *unmöglich* noch *verboten* ist.

10 Kontakte lassen sich durch Komplemente beseitigen, aber ...

In einem markierten B/E-Netz sind zwei Bedingungen B_1 und B_2 *komplementär* zueinander, wenn Folgendes gilt:

1. Die Menge $\bullet B_1$ der Vorgänger von B_1 ist gleich der Menge B_2^\bullet der Nachfolger von B_2 ist (d.h. es gilt $\bullet B_1 = B_2^\bullet$).
2. Die Menge B_1^\bullet der Nachfolger von B_1 ist gleich der Menge $\bullet B_2$ der Vorgänger von B_2 ist (d.h. es gilt $B_1^\bullet = \bullet B_2$).
3. Genau eine der beiden Bedingungen ist erfüllt (d.h. trägt ein Token) und die andere ist nicht erfüllt.

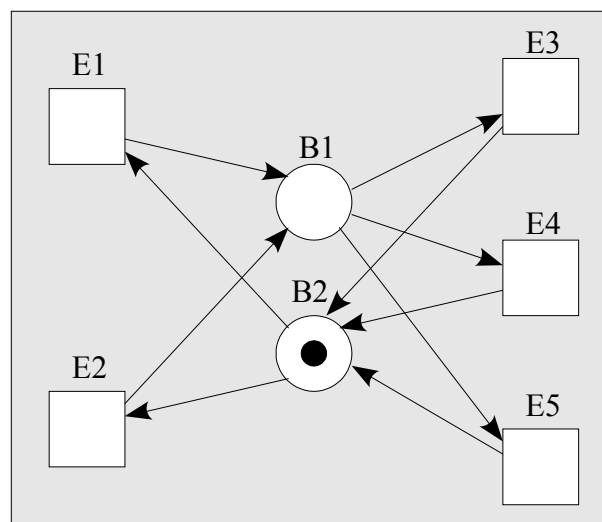


Abbildung 10.1: Zwei komplementäre Bedingungen B_1 und B_2

Im Beispiel in Abb. 10.1 gilt: 1. $\bullet B_1 = \{E_1, E_2\} = B_2^\bullet$ und 2. $B_1^\bullet = \{E_3, E_4, E_5\} = \bullet B_2$ und 3. B_1 ist nicht erfüllt und B_2 ist erfüllt. Also sind B_1 und B_2 komplementär zueinander.

Dass B_1 das Komplement von B_2 (und damit B_2 das Komplement von B_1) ist, bedeutet inhaltlich: Wird B_1 durch ein Ereignis erfüllt (in Abb. 10.1 z. B. durch E_2), so wird B_2 dadurch nicht-erfüllt. Und umgekehrt: Wird B_1 durch ein Ereignis nicht-erfüllt (z. B. durch E_4), so wird B_2 dadurch erfüllt. Die Eigenschaft "Genau eine der beiden Bedingungen B_1 und B_2 ist erfüllt und die andere ist nicht-erfüllt" bleibt also stets erhalten, unabhängig davon, wieviele und welche Ereignisse eintreten.

Nimmt man ein markiertes B/E-Netz N und fügt man zu einigen Bedingungen des Netzes ihr Komplement hinzu, so erhält man ein markiertes Netz N' . Man kann sich anhand von Beispielen davon überzeugen (und allgemein beweisen), dass die beiden Netze N und N' im wesentlichen die gleichen Ausführungen zulassen (man muss nur davon absehen, dass N' ein paar Bedingungen mehr umfasst als N). In diesem Sinne verändert das Hinzufügen von Komplement-Bedingungen ein markiertes Netz nicht.

Hat man ein markiertes B/E-Netz N , bei dessen Ausführung Kontakt-Situationen auftreten (siehe Abschnitt 9.), so kann man durch Hinzufügen von Komplement-Bedingungen diese Kontakt-Situationen beseitigen. Im Zweifelsfall kann man zu jeder Bedingung in N (die noch kein Komplement hat) eine komplementäre Bedingung hinzufügen. In vielen Fällen genügt es aber, ein paar Komplemente an den richtigen Stellen einzubauen.

Abb. 10.2 zeigt das Netz aus Abb. 9.2 ergänzt um eine Bedingung B4, die komplementär zu B2 ist. Das Netz in Abb. 9.2 ist kontaktbehaftet, das Netz in Abb. 10.2 ist kontaktfrei:

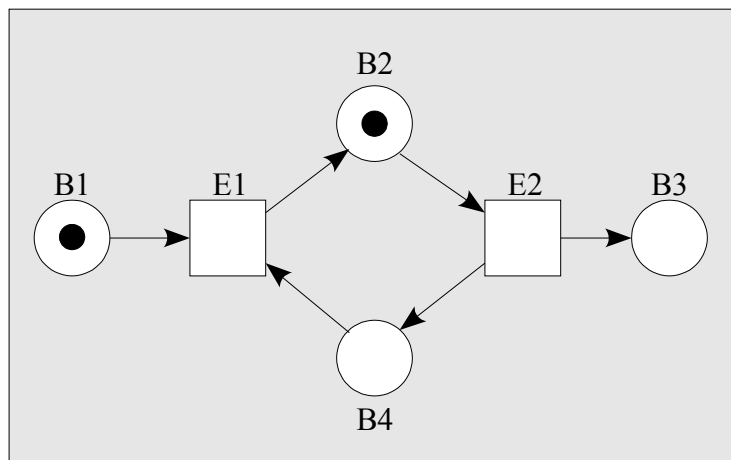


Abbildung 10.2: Das Komplement B4 zu B2 verhindert den Kontakt von E1

Das Prozessnetz zu diesem kontaktfreien Netz sieht so aus:

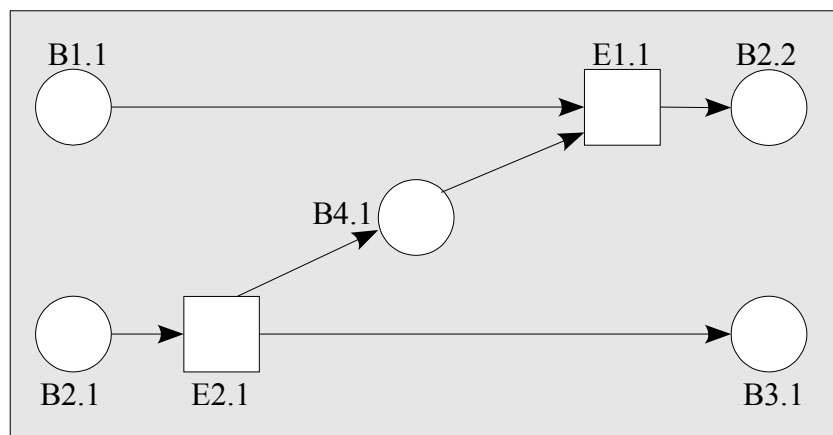


Abbildung 10.3: Das Prozessnetz zur vorigen Abbildung

Der Schnitt-Graph zu diesem Prozessnetz sieht so aus:

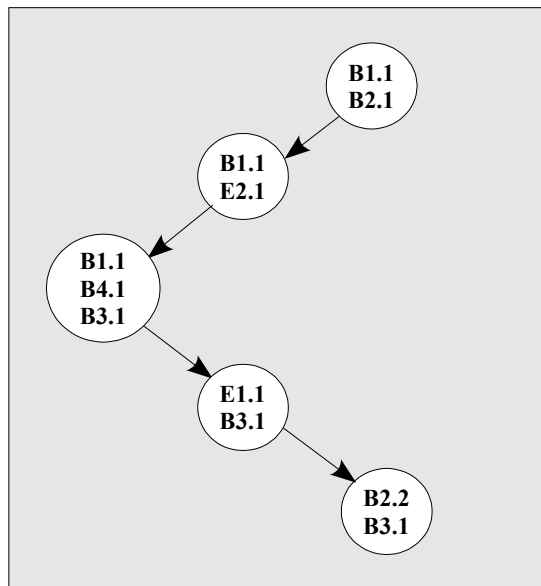


Abbildung 10.4: Der Schnittgraph zur vorigen Abbildung

Die Auswirkungen der Komplement-Bedingung B4 erkennt man besonders deutlich, wenn man die Abbildungen 10.2, 10.3 und 10.4 mit den Abbildungen 9.2, 9.3 und 9.4 bzw. 9.5 vergleicht.

Fügt man zu einem markierten B/E-Netz N eine Komplement-Bedingung hinzu, so ändert man dadurch nicht die Menge der möglichen Ausführungen von N. Und umgekehrt gilt: Ist in einem Netz N eine Komplement-Bedingung enthalten und entfernt man sie, so ändert sich dadurch die Menge der möglichen Ausführungen nicht.

Man kann also z. B. in einem markierten Netz alle Bedingungen durch ihre Komplement-Bedingungen ersetzen. Abb. 10.5 zeigt ein einfaches Beispiel für diese Netz-Transformation.

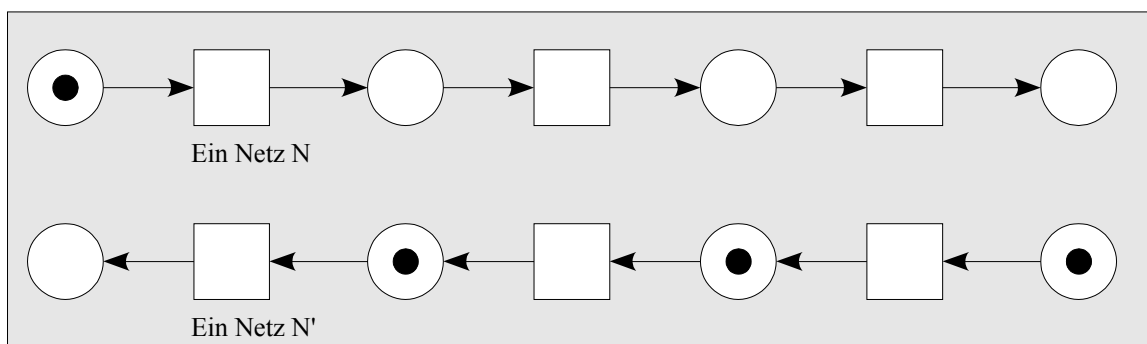


Abbildung 10.5: Ersetzung aller Bedingungen durch ihre Komplemente

Anschaulich gilt: Wenn man das Netz N (in Abb. 10.5) ausführt, dann wandert ein Token von links nach rechts. Wenn man dagegen N ausführt, dann wandert ein "Loch" von links nach rechts, die Token wandern von rechts nach links.

Ein Formalist, der sich nur für den Kalkül der Netze interessiert und nicht für die Vorstellungen und Absichten, die diesen Kalkül motivieren, wird vermutlich dazu neigen, eine Bedingung und ihr Komplement als gleichwertig zu betrachten. Kontakt-Situationen wird er für unwesentlich halten,

da man sie durch Komplementierung beseitigen kann. Und unter geeigneten Umständen wird er das Wandern von Token als äquivalent zum Wandern von Löchern betrachten.

Geht man dagegen von der Vorstellung des Universums der Signale aus (siehe Abschnitt 1.), dann besteht ein inhaltlicher Unterschied zwischen der Ausbreitung eines Signals und der Ausbreitung "der Abwesenheit eines Signals". Ein kontaktbehaftetes Netz N ist dann etwas wesentlich anderes als ein kontaktfreies Netz N' , auch wenn N und N' in einem bestimmten formalen Sinn die gleichen Ausführungen zulassen.

Zur Illustration dieser intuitiven Vorstellung sollen die Netze N und N' in Abb. 10.5 inhaltlich interpretiert werden.

Das Netz N könnte ein Stück Straße mit drei Autos darauf darstellen. Jedes Auto ist durch ein Token repräsentiert und fährt von rechts nach links. Dem Kontakt im Netz entspricht anschaulich die Gefahr eines Auffahrunfalls auf der Straße.

Im Netz N' wird die Bewegung einer Lücke über die gleiche Straße beschrieben. Die Lücke ist durch ein Token repräsentiert und bewegt sich von links nach rechts. In diesem Netz gibt es keinen Kontakt, d.h. der Gefahr eines Auffahrunfalls entspricht keine bekannte Eigenschaft des Netzes.

Die weitere Entwicklung der Netztheorie (zusammen mit weiteren Anwendungen) muss zeigen, ob dieser intuitive Unterschied zwischen dem kontaktbehafteten Netz N' und dem kontaktfreien Netz N sich auch formal erhärten lässt.

Zum Abschluss dieses Abschnitts noch eine Anmerkung zu Kontakten und Komplementen in Stellen-Transitionsnetzen (S/T-Netzen).

In einem S/T-Netz hat eine Transition T Kontakt, wenn alle Vorstellen von T genügend Token tragen, aber (mindestens) eine Nachstelle von T nicht genügend freie Kapazität hat und T deshalb nicht schalten kann.

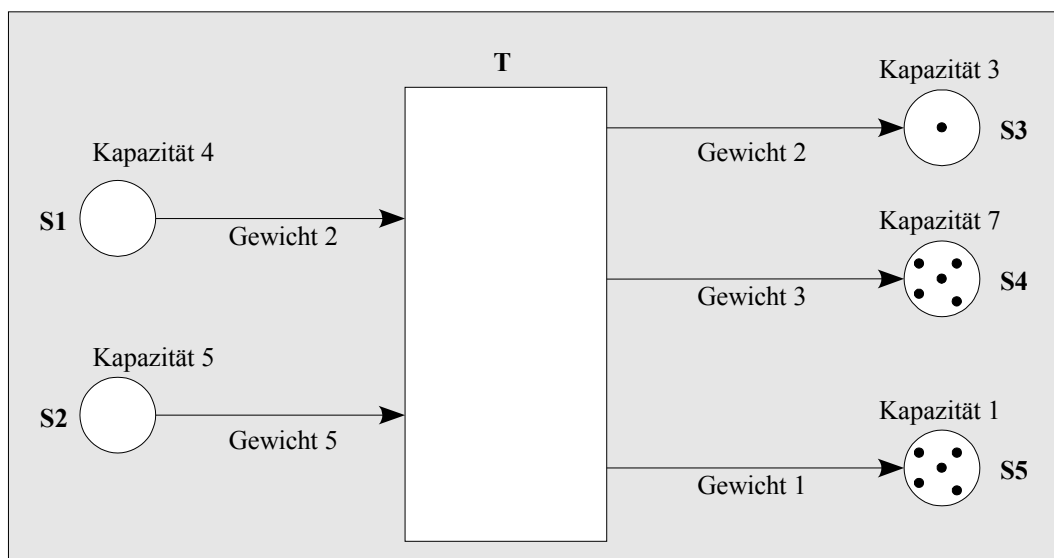


Abbildung 10.6: Die Transition T an der Stelle $S4$ Kontakt

Eine Transition T kann nur an solchen Stellen Kontakt haben, die eine endliche Kapazität haben. Haben alle Nachstellen von T unbegrenzte Kapazität, so kann T nicht in eine Kontakt-Situation geraten.

Auch in S/T-Netzen kann man (wie in B/E-Netzen) Kontakte durch Komplemente beseitigen. Sei $S1$ eine Stelle in einem markierten S/T-Netz und habe $S1$ eine Kapazität von n Token und eine momentane Markierung von m Token (mit $0 \leq m \leq n$). Dann müssen auf der Komplement-Stelle $S2$ von $S1$ genau $n - m$ Token liegen:

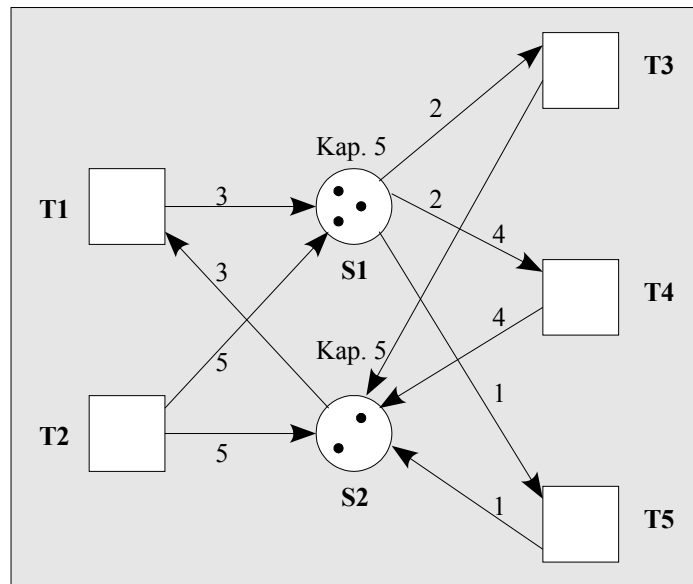


Abbildung 10.7: Komplementäre Stellen $S1$ und $S2$ in einen S/T-Netz

Hat man in einem markierten S/T-Netz zu einer Stelle $S1$ die entsprechende Komplement-Stelle $S2$ eingezeichnet, so kann man die Kapazitätsbeschriftung an den Stellen $S1$ und $S2$ weglassen: Auf $S1$ und $S2$ liegen zusammengenommen n Token (wobei n die ursprüngliche Kapazität von $S1$ ist) und keine Transition wird die Summe der Token auf $S1$ und $S2$ je ändern (siehe dazu auch das Beispiel in Abb. 10.6).

Hat eine Stelle $S1$ (in einem markierten S/T-Netz) dagegen eine unbegrenzte Kapazität, so ist es nicht möglich bzw. nicht sinnvoll, zu $S1$ eine komplementäre Stelle $S2$ einzuzeichnen. Denn $S2$ müsste als Markierung "unbegrenzt viele" Token tragen und diese Markierung könnte durch keine Transition verändert werden. Denn wenn man zu abzählbar-unendlich vielen Token endlich viele hinzufügt oder wenn man von abzählbar-unendlich vielen Token endlich viele wegnimmt, dann erhält man als Resultat doch wieder abzählbar-unendlich viele Token. Normalerweise werden in S/T-Netzen auch nur endlich viele Token auf jeder Stelle zugelassen.

11 Konfusion: In Netzen klar erkennbar

"Konfusion" ist einer der wichtigsten Begriffe der Netztheorie. Konfusion (im netztheoretischen Sinne) ist ein ganz alltägliches und weit verbreitetes Phänomen. Konfusion ist sehr eng mit Nebenläufigkeit verbunden und wurde zum ersten Mal im Rahmen der Netztheorie präzise und umfassend definiert.

Konfusion (im netztheoretischen Sinne) liegt immer dann vor, wenn die Reihenfolge, in der zwei nebenläufige Ereignisse eintreten, für den weiteren Ablauf eines Systems wesentlich ist.

Schon auf dieser informellen Ebene kann man erkennen, dass Konfusion auf einer Art von Widerspruch beruht: Zwei (unwiederholbare) Ereignisse sind nebenläufig, wenn sie unabhängig voneinander sind und somit in beliebiger Reihenfolge eintreten können. Wenn diese "beliebige Reihenfolge" dann doch wichtig ist, liegt Konfusion vor.

Ein Beispiel:

Ereignis E0: Eine Frau bekommt ein Kind K1.

Ereignis E1: Eine zweite Frau (eine Schwester der ersten Frau) bekommt ein Kind K2.

Der Großvater der beiden Kinder (d.h. der Vater der beiden Schwestern) hat in seinem Testament folgendes *bedingtes Ereignis* spezifiziert: Wenn K1 *vor* K2 geboren wird, dann erbt K1 mein ganzes Vermögen.

Unter der plausiblen Annahme, dass die beiden Schwestern ihre Kinder unabhängig voneinander bekommen, (d.h. unter der Annahme, dass E0 und E1 nebenläufig sind) liegt hier Konfusion vor.

Allgemein ist Konfusion aus folgendem Grund problematisch: Anhand eines Vergleichs von zwei *stetigen* Größen (im Beispiel: die Geburtszeitpunkte von K1 und K2) soll eine *diskrete* Entscheidung getroffen werden (K1 erbt oder erbt nicht). Da die stetigen Größen beliebig nahe beieinander liegen können, wird man in manchen Fällen keine richtige, sondern nur eine *willkürliche* Entscheidung treffen können. In diesem Sinne ist das oben erwähnte Testament nicht (in allen Fällen) ausführbar.

Wenn der Bediener einer technischen Anlage aufgrund der Reihenfolge, in der nebenläufige (unabhängige) Ereignisse eintreten, wesentliche Entscheidungen treffen muss, dann ist die Anlage möglicherweise gefährlich und vermutlich fehlerhaft konstruiert.

Sportliche Leistungen werden häufig durch stetige Größen beschrieben (Zeiträume, Längen). Werden zwei sportliche Leistungen nebenläufig erbracht (was besonders deutlich der Fall ist, wenn z. B. zwei Läufer an weit voneinander entfernten Orten miteinander um die Wette laufen), und wird anhand der Messergebnisse eine diskrete Entscheidung getroffen (z. B. "Läufer A ist Weltmeister"), so liegt ebenfalls Konfusion vor. Häufig wird die Willkür der Entscheidung hinter technischen Einrichtungen verborgen. Z. B. tut man so, als könne man bei einem Wettlauf sinnvoll Hundertstel oder sogar tausendstel Sekunden messen. Die Fragwürdigkeit dieses Verfahrens wird deutlich, wenn man es mit der folgenden, alternativen Vorgehensweise vergleicht: Man setzt für jeden Sportler viele Messapparaturen ein (z. B. wird jeder Läufer von 10 oder von 100 Quarzuhren gestoppt). Wenn die Ergebnisse für einen Läufer sich um höchstens *eine Einheit* unterscheiden, wird das Ergebnis (bestehend aus zwei gleichen oder aufeinanderfolgenden Zahlen, z. B. 117/117 Zehntelsekunden oder 117/118 Zehntelsekunden) anerkannt, andernfalls wird das Ergebnis nicht anerkannt und der Lauf wiederholt. Vermutlich würde bei dieser Vorgehensweise eine Tendenz entstehen, auf das so genannte "Messen" von Hundertsteln und Tausendsteln von Sekunden zu verzichten.

Mit Petri-Netzen kann man Konfusion präzise beschreiben und in einem Netz-Modell eines Systems kann man mögliche Konfusionen eindeutig erkennen.

Sei E ein Ereignis in einem markierten B/E-Netz. Die *Konflikt-Menge* von E besteht aus den Ereignissen, mit denen E (unter der momentanen Markierung) in Konflikt steht.

In einem markierten B/E-Netz liegt eine **Konfusion** vor, wenn für zwei Ereignisse E_0 und E_1 gilt:

1. E_0 und E_1 sind beide aktiviert und stehen nicht in Konflikt miteinander (d.h. E_0 und E_1 können nebenläufig zueinander eintreten).
2. Wenn E_0 eintritt, so wird dadurch die Konflikt-Menge von E_1 verändert (d.h. es kommen Konflikt-Ereignisse hinzu, es fallen welche weg oder es geschieht beides).

Die folgenden drei Abbildungen zeigen Beispiele für die drei Varianten von Konfusion.

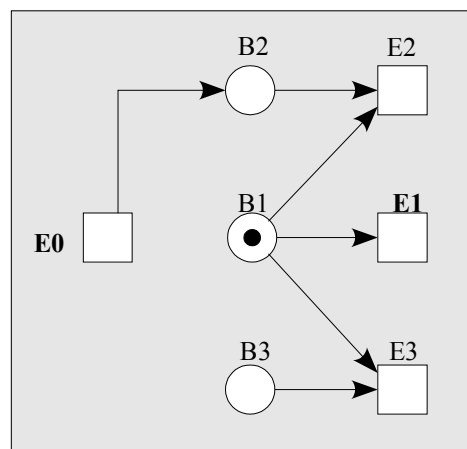


Abbildung 11.1: Wenn E_0 eintritt, wird die Konfliktmenge von E_1 größer

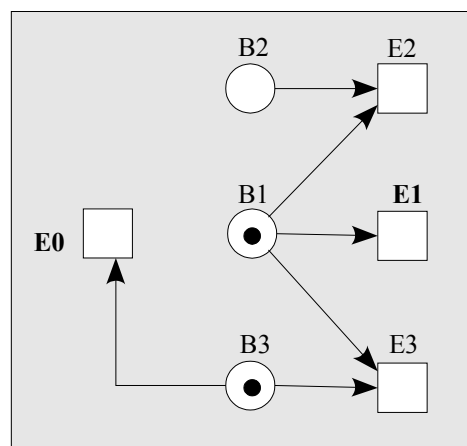


Abbildung 11.2: Wenn E_0 eintritt, wird die Konfliktmenge von E_1 kleiner

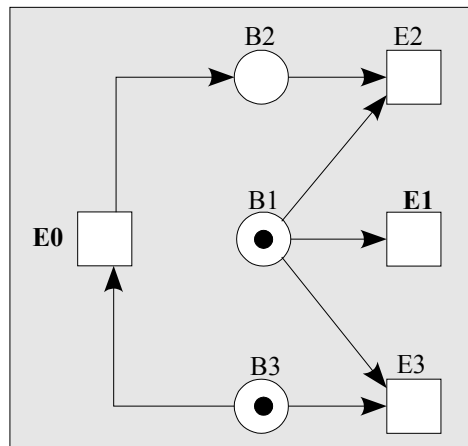


Abbildung 11.3: Wenn E0 eintritt, wird die Konfliktmenge von E1 "kleiner und größer"

12 S/T-Netze, zwei Ideen: Mehrere-Token und unterscheidbare Marken

Das Fundament der Netztheorie bilden die *Prozessnetze* (d.h. die Netze aus einmaligen Bedingungen und Ereignissen). Ein Prozessnetz ist zyklensfrei und umfasst typischerweise unendlich viele Knoten.

Etwas abstrakter sind die B/E-Netze (Netze aus wiederholbaren Bedingungen und Ereignissen). Ein typisches B/E-Netz enthält Zyklen und umfasst nur endlich viele Knoten. Die Bedeutung eines markierten B/E-Netzes N ist ein Prozessnetz (wenn N konfliktfrei ist) bzw. eine Menge von Prozessnetzen (wenn N nicht konfliktfrei ist).

Beim Übergang von einem Prozessnetz zu seiner Darstellung durch ein B/E-Netz werden typischerweise jeweils viele, häufig auch unendlich viele, unwiederholbare Ereignisse zu einem wiederholbaren Ereignis zusammengefasst und entsprechendes gilt für Bedingungen.

Dieser Abstraktionsschritt von *unendlich* vielen *unwiederholbaren* Netz-Elementen zu *endlich* vielen *wiederholbaren* Elementen ist konzeptuell sehr wesentlich, er erfüllt aber noch nicht alle Wünsche, die man bei der Anwendung von Netzen auf praktische Probleme entwickelt. Auch wenn B/E-Netze "nur" endlich groß sind, so tendieren sie doch häufig dazu, zu groß zu werden (größer, als das Papier, auf das man sie zeichnen will und größer als das, was ein Mensch gut überschauen kann).

Man hat deshalb schon relativ früh in der Geschichte der Netztheorie damit begonnen, weitere Netzarten zu definieren und zu untersuchen. Ein Ziel dabei ist es, übersichtlichere, kompaktere, abstraktere, leichter zu verstehende Darstellungen von Systemen zu ermöglichen. Ähnlich wie bei Programmier-Sprachen ist es eher zu leicht als zu schwierig, sich weitere "features" (Konstruktionen) auszudenken, die für eine spezielle Anwendung (oder für viele Anwendungen) nützlich sind. Das Problem besteht darin, möglichst einfache und grundlegende Konzepte und Konstruktionen zu finden und diese nicht nur syntaktisch, sondern auch semantisch sauber zu fundieren. Bei Petri-Netzen heißt fundieren: Auf B/E-Netze oder direkt auf Prozessnetze zurückführen.

Bei der Suche nach Netz-Arten "oberhalb von B/E-Netzen" sind (neben zahllosen "features", auf die hier nicht eingegangen wird) vor allem zwei grundlegende Ideen entwickelt worden, die im Folgenden als die *mehrere-Token-Idee* und die *unterscheidbare-Marken-Idee* bezeichnet werden.

Die mehrere-Token-Idee stammt, das wird kaum überraschen, vom Token-Spiel auf B/E-Netzen. Kurz gesagt, besteht sie darin, überall da, wo bei einem B/E-Netz höchstens *ein* Token erlaubt ist (auf jeder Bedingung darf höchstens ein Token liegen, beim Eintreten eines Ereignisses E "fließt" über jeden Pfeil von bzw. nach E genau ein Token) *mehrere* Token zuzulassen. Die entsprechenden Netze werden meistens *S/T-Netze* (Netze aus Stellen und Transitionen) genannt. Es hat sich herausgestellt, dass die semantische Fundierung dieser Netzart nicht ganz so einfach ist, wie man vielleicht vermutet, wenn man zum ersten Mal der mehrere-Token-Idee begegnet. Auf zwei Probleme von S/T-Netzen wird in den Abschnitten 15. und 16. noch näher eingegangen.

Die Token des Token-Spiels auf S/T-Netzen sind so ununterscheidbar oder gleichwertig wie etwa 10-Cent-Münzen für jemanden, der sich damit etwas kaufen will (und sie nicht z. B. nach Jahrgängen sammelt).

Die unterscheidbare-Marken-Idee besteht darin, anstelle von ununterscheidbaren Token unterscheidbare Marken zu verwenden, allerdings von jeder Markenart jeweils höchstens ein Exemplar.

Es hat sich eingebürgert, solche unterscheidbaren Marken (im Gegensatz zu den "schwarzen" Token) als *farbige Marken* zu bezeichnen. Mit "farbig" meint man dabei: "irgendwie-unterscheidbar-und-wir-wollen-im-Moment-noch-keine-Annahmen-darüber-machen-wie-und-wodurch". In praktischen Anwendungen macht man die Marken natürlich nicht durch wirkliche Farben unterscheidbar,

sondern indem man sie mit Zahlen, Texten, Tabellen etc. beschriftet. Jede mögliche Zahl, jeder mögliche Text und jede mögliche Tabelle etc. gilt dann als eine "Farbe".

Theoretiker neigen (vor allem in einführenden Beispielen) dazu, Marken mit Zeichenkette wie "rot", "grün" etc. zu beschriften und dadurch unterscheidbar zu machen.

In farbigen Netzen nennt man (wie in S/T-Netzen) die *runden Netzknoten* häufig auch *Stellen* und die *eckigen Knoten Transitionen*.

Die unterscheidbare-Marken-Idee erlaubt, wie schon angedeutet, auf einer Stelle höchstens je eine Marke von jeder "Farbe", z. B. eine rote, eine grüne, eine gelbe ... etc. Und wenn eine Transition T schaltet, dann darf über jeden Pfeil, der von T weg oder zu T hinführt höchstens je eine Marke von jeder Farbe fließen, z. B. eine violette, eine rosa etc. Solche Netze werden auch als *strikte farbige Netze* bezeichnet, wobei das "strikt" die Einschränkung "höchstens eine Marke von jeder Farbe" bezeichnet.

Die mehrere-Token-Idee und die unterscheidbare-Marken-Idee lassen sich auch kombinieren zur *mehrere-unterscheidbare-Marken-Idee*. D.h. auf einer Stelle dürfen dann auch z. B. zwei rote, drei grüne, eine gelbe ... etc. Marken liegen und über einen Pfeil dürfen bei einem Schaltvorgang z. B. eine rote, zwei grüne und drei gelbe Marken fließen. Die entsprechenden Netze heißen *farbige Netze* (man lässt einfach die Einschränkung "strikt" fort).

Die so genannten *Prädikat-Transitionsnetze* wurden *vor* den farbigen Netzen erfunden. Sie lassen sich am einfachsten verstehen als eine spezielle Darstellung von (einer Teilklasse der) farbigen Netzen.

13 Kann eine Marke einen Schaltvorgang überdauern?

Wenn zwei Billardkugeln "in einen Zusammenstoß hineingehen", dann kommen aus diesem Zusammenstoß auch zwei Billardkugeln heraus und die herauskommenden Kugeln sind natürlich die selben wie die, die hineingehen.

Wenn in einem B/E-Netz ein Ereignis E mit zwei Vor- und zwei Nachbedingungen eintritt, dann wird von jeder Vorbedingung ein Token entfernt und auf jede Nachbedingung wird ein Token gelegt.

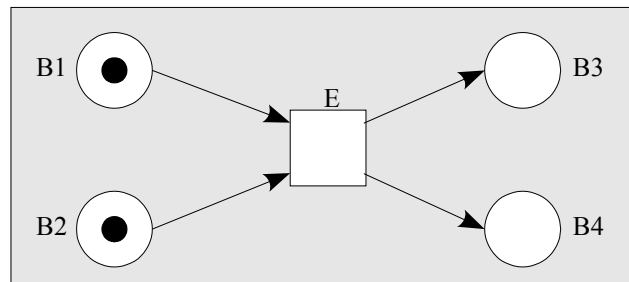


Abbildung 13.1: Ein Ereignis E vor seinem Eintreten

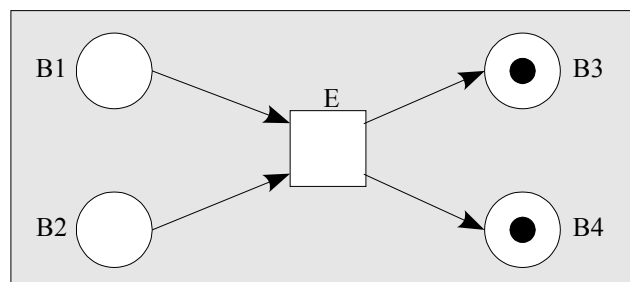


Abbildung 13.2: Das Ereignis E nach seinem Eintreten

Zumindest für Billiard-Spieler ist es nahe liegend anzunehmen, dass die beiden Token, die nach dem Eintreten von E auf B3 und B4 liegen, die *selben* sind wie die, die vor dem Eintreten von E auf B1 und B2 lagen.

Diese Annahme wirft Probleme auf. Wandert das Token auf B1 beim Eintreten von E nach B3 oder B4? Noch deutlicher tritt das Problem auf, wenn E unterschiedlich viele Vor- und Nachbedingungen hat, z. B. so:

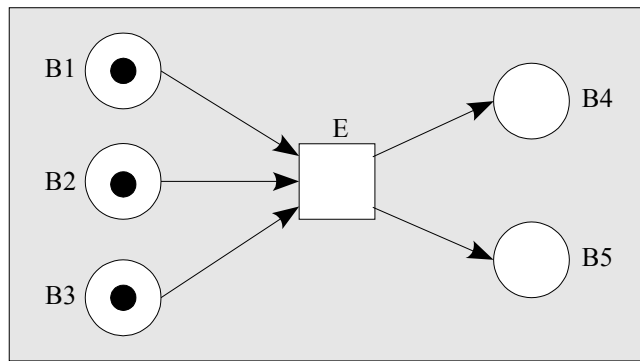


Abbildung 13.3: Ein Ereignis E mit 3 Vor- und 2 Nachbedingungen

Welches der drei Token geht verloren und welches der verbleibenden Token geht wohin?

Man kann diese Probleme grundsätzlich vermeiden, indem man sich vorstellt, dass die Eingabe-Token beim Eintreten von E konsumiert (verbraucht, gelöscht) und die Ausgabe-Token produziert (neu geschaffen) werden.

Die Auffassung, dass beim Eintreten eines Ereignisses Token konsumiert und produziert werden (kurz: die K-P-Auffassung) und Token nicht von einer Vorbedingung zu einer Nachbedingung "durchgereicht" werden, ist in vielen Anwendungsfällen angemessen und sinnvoll, sie sollte aber nicht als ein Dogma der Netztheorie missverstanden werden.

Denn in vielen Anwendungsfällen, insbesondere wenn farbige Netze verwendet werden, ist die K-P-Auffassung unangemessen und die Vorstellung, dass Marken beim Schalten einer Transition erhalten bleiben und durchgereicht werden, ist in diesen Fällen viel natürlicher.

Die Netztheorie geht *nicht* davon aus, dass die Eingabe-Token vernichtet und die Ausgabe-Token neu geschaffen werden. Vielmehr vermeidet sie jede Annahme darüber, ob gewisse Ausgabe-Token mit gewissen Eingabe-Token identisch sind oder nicht. D.h. die Netztheorie verhält sich neutral gegenüber der K-P-Auffassung und der Erhaltungsauffassung.

Im Universum der Signale (siehe Abschnitt 1) ist es nicht sinnvoll, Annahmen über die Identität von Signalen vor einer Interaktion und den Signalen hinter dieser Interaktion zu machen, da es für solche Signale (Elementarteilchen) keine Verfahren geben kann, mit denen man ihre Identität bzw. nicht-Identität feststellen könnte.

Mit Petri-Netzen modelliert man Systeme aber häufig nicht auf der Ebene von Signalen (Elementarteilchen), sondern auf einem viel "höheren", makroskopischen Niveau. In solchen Fällen kann es durchaus sinnvoll sein, zwischen dem Konsumieren und Produzieren von Marken und dem Erhalten von Marken bei einem Schaltvorgang zu unterscheiden. Der Anwender sollte die Möglichkeit nutzen, von Fall zu Fall die ihm angemessener erscheinende Auffassung zugrunde zu legen. Die Netztheorie ist mit beiden Auffassungen verträglich.

14 Schlingen in Netzen, oder: Nebenbedingungen in der Wirklichkeit?

Schlingen in Netzen bzw. Nebenbedingungen in der Wirklichkeit waren schon häufig Gegenstand intensiver und ausgedehnter Diskussionen. Hier sollen die wesentlichen Definitionen, Probleme und Argumente kurz zusammengefasst werden.

Eine *positive Nebenbedingung* NB für ein Ereignis E ist eine Bedingung, die erfüllt sein muss, damit E eintreten kann, die aber durch das Eintreten von E nicht verändert (d.h. nicht-erfüllt) wird. Beispiel: Damit ein Auto an einer Kreuzung (ohne gegen Gesetze zu verstoßen) losfahren kann, muss die entsprechende Ampel grün zeigen. Durch das Losfahren des Autos hört die Ampel aber nicht auf, grün zu zeigen.

Ganz entsprechend ist eine *negative Nebenbedingung* NB für ein Ereignis E eine Bedingung, die nicht erfüllt sein darf, damit E eintreten kann, die aber durch das Eintreten von E nicht verändert (d.h. erfüllt) wird.

Beispiel: Die Bedingung "Die Ampel zeigt rot" ist eine negative Nebenbedingung für das Losfahren eines Autos an einer Kreuzung.

Als Beispiele für Nebenbedingungen werden häufig auch chemische *Katalysatoren* angeführt. Ein Katalysator ist ein Stoff, der anwesend sein muss, damit eine bestimmte chemische Reaktion stattfinden kann, der aber durch diese Reaktion selbst nicht verändert wird. Diese einfache Erklärung des Begriffs "Katalysator" ist vielen noch aus dem Chemie-Unterricht an der Schule geläufig.

Eine *Schlinge* in einem B/E-Netz besteht aus einem Ereignis E und einer Bedingung B, die gleichzeitig Vor- und Nachbedingung von E ist.

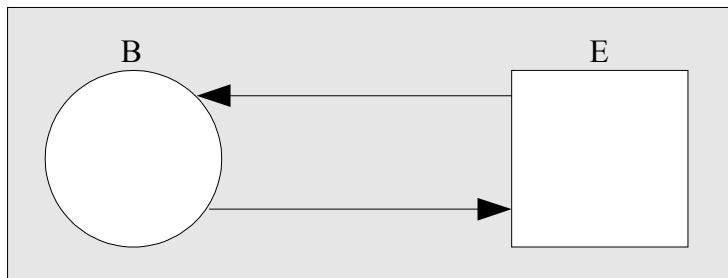


Abbildung 14.1: Eine Schlinge

Aus den (allgemein akzeptierten) Schaltregeln für markierte B/E-Netze folgt, dass das Ereignis E in einer Schlinge nie eintreten kann. Denn die Bedingung B der Schlinge müsste (als Vorbedingung von E) erfüllt und (als Nachbedingung von B) nicht-erfüllt sein. B kann aber unmöglich erfüllt und nicht-erfüllt sein.

In B/E-Netzen gibt es keine Nebenbedingungen für ein Ereignis E, nur Vor- und Nachbedingungen, die durch das Eintreten von E verändert werden. Es wurde aber schon häufig die ursprüngliche Schaltregel so verändert, dass man die Bedingung B einer Schlinge als Nebenbedingung für das Ereignis E der Schlinge verstehen kann.

Meistens wird die Schaltregel folgendermaßen abgeändert: Beim Eintreten eines Ereignisses wird zuerst von jeder Vorbedingung ein Token genommen und *danach* an jede Nachbedingung ein Token gegeben (siehe dazu auch Abschnitt 6.).

Für Schlingen (siehe Abb. 14.1) bedeutet diese veränderte Schaltregel: B ist eine positive Nebenbedingung für E, d.h. B muss erfüllt sein, damit E eintreten kann. Beim Eintreten von E wird zuerst das Token von B genommen und danach wieder ein Token auf B gelegt. Nach dem Eintreten von E liegt also (genau wie vor dem Eintreten von E) wieder ein Token auf B.

Seltener wird die Schaltregel folgendermaßen abgeändert: Beim Eintreten eines Ereignisses wird zuerst an jede Nachbedingung ein Token gegeben und *danach* von jeder Vorbedingung ein Token genommen.

Für Schlingen (siehe Abb. 14.1) bedeutet diese veränderte Schaltregel: B ist eine negative Nebenbedingung für E, d.h. B darf nicht erfüllt sein, damit E eintreten kann. Beim Eintreten von E wird zuerst ein Token auf B gelegt und danach wieder weggenommen. Nach dem Eintreten von E liegt also (genau wie vor dem Eintreten von E) kein Token auf B.

Ein Problem liegt in der Symmetrie dieser beiden Veränderungen der ursprünglichen Schaltregel. Man kann nur *eine* der Veränderungen wählen, es dürfte aber schwer sein, die Wahl ohne Willkür, sondern "wohl begründet" zu treffen.

Statt gleich die Schaltregel zu verändern, kann man sich auch fragen, warum es in B/E-Netzen keine einfache, direkte Möglichkeit gibt, Nebenbedingungen darzustellen. Prof. Petri hat das einmal so begründet: "In letzter Analyse gibt es auch in `der Wirklichkeit` keine Nebenbedingungen".

Diese möglicherweise überraschende Begründung soll erläutert werden, indem für die oben erwähnten Beispiele (Ampel und Katalysator) Analysen skizziert werden, die die Nebenbedingungen zum Verschwinden bringen.

Im Ampel-Beispiel wurde "Die Ampel zeigt grün" als eine Bedingung aufgefasst. Tatsächlich sendet eine grüne Ampel sehr *vielen* "Lichtbündel" aus. Fällt eines dieser Lichtbündel in das Auge eines wartenden Autofahrers, dann ist *dieses* Bündel B die Vorbedingung für das Losfahren dieses Fahrers. Alle anderen Bündel gehen an diesem Fahrer vorbei und sind keine Vorbedingungen für seine Handlungen. Nur indem das Bündel B im Auge des Fahrers "verbraucht" wird, bewirkt es, dass der Fahrer losfährt. Die Ampel produziert zwar weitere Bündel von grünem Licht, die möglicherweise Vorbedingungen für weitere Ereignisse sind, aber die Vorbedingung B wurde durch "ihr" Losfahrereignis von "erfüllt" in "nicht-erfüllt" umgewandelt. Als B/E-Netz könnte man den Vorgang etwa so modellieren:

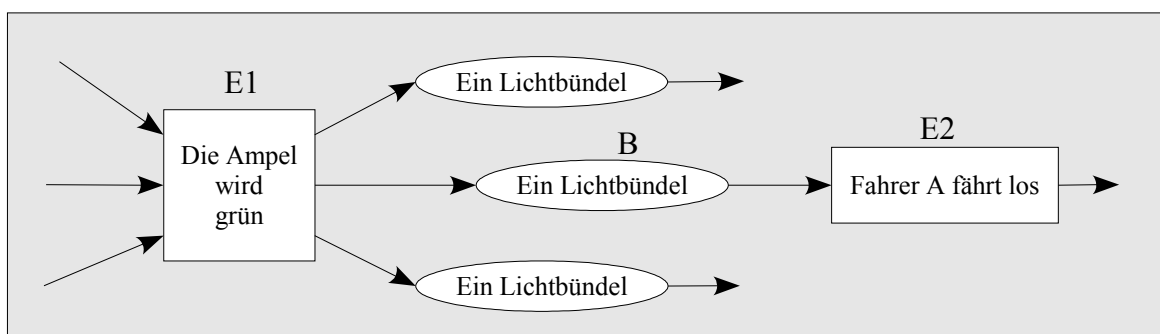


Abbildung 14.2: Ein B/E-Netz für das Ampel-Beispiel

Für das Katalysator-Beispiel wurde oben mit Absicht eine populäre und sehr vereinfachte Erklärung des Begriffs "Katalysator" angeführt. Tatsächlich verändert sich ein Katalysator sehr wohl in der von ihm katalysierten Reaktion, aber nicht nur einmal, sondern mehrmals. Typischerweise durchläuft der Katalysator dabei zyklisch eine Reihe von Zuständen und liegt am Ende wieder in der glei-

chen Form vor wie am Anfang. Das folgende B/E-Netz illustriert einen solchen Vorgang, bei dem zwei Stoffe (Stoff 1 und Stoff 2) mit Hilfe eines Katalysators zu einem Reaktionsergebnis verbunden werden. Die dabei als Zwischenergebnis entstehende "temporäre Verbindung" besteht aus den beiden Stoffen und dem Katalysator:

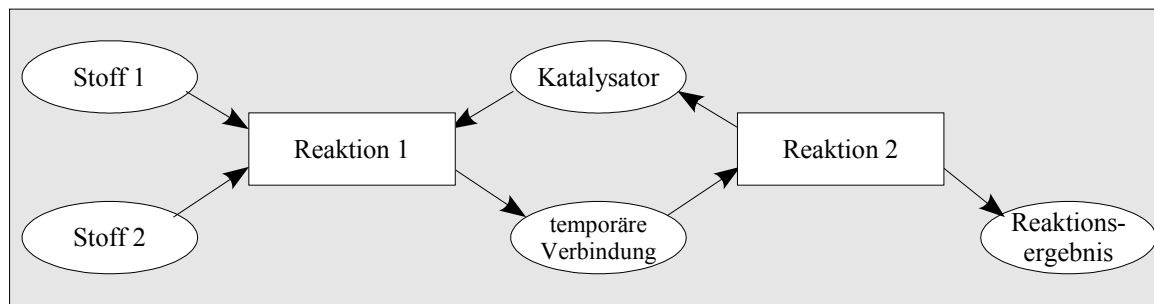


Abbildung 14.3: Ein B/E-Netz für das Katalysator-Beispiel

Dieses B/E-Netz enthält zwar einen Zyklus (Katalysator, Reaktion 1, temporäre Verbindung, Reaktion 2), aber offensichtlich keine Schlingen.

Die beiden Beispiele (Ampel und Katalysator) sind in gewisser Weise repräsentativ, d.h. es gibt genau zwei Arten von positiven Nebenbedingungen: bei der einen Art ist die Bedingung "zu groß und umfassend" gewählt (alles grüne Licht der Ampel als eine Bedingung). Wenn man eine solche Bedingung in ihre wirksamen Teile zerlegt (gegebenenfalls bis hinunter zu einzelnen Wirkungsquanten), dann verschwindet der Charakter einer Nebenbedingung. Bei der anderen Art von Nebenbedingung wurden mehrere verschiedene Bedingungen identifiziert, die zyklisch erfüllt werden (die verschiedenen Zustände eines Katalysators). Hebt man die Identifizierung auf, so verschwindet auch hier der Charakter der Nebenbedingung. Für negative Nebenbedingungen gilt ganz entsprechendes. In letzter Analyse gibt es in "der Wirklichkeit" also keine Nebenbedingungen. Trotzdem gibt es in der Wirklichkeit ganz offensichtlich Nebenbedingungen, d.h. das Konzept "Nebenbedingung" ist auf vielen Abstraktionsebenen, auf denen wir die Wirklichkeit erfahren und analysieren, sinnvoll und nützlich (nur auf der "untersten Ebene" nicht).

B/E-Netze sind ihrer Herkunft und dem mit ihnen verbundenen Anspruch nach Teil einer Grundlagen-Theorie. Deshalb enthalten sie Nebenbedingungen nicht als Basiskonzept. Sie ermöglichen es aber, verschiedene Arten von Nebenbedingungen als zusammengesetzte Konstrukte darzustellen (wie für die Beispiele Ampel und Katalysator in den Abb. 14.2 und 14.3 angedeutet).

Für die praktische Anwendung von Netzen ist es natürlich sinnvoll, Notationen für verschiedene Arten von Nebenbedingungen einzuführen. Die genaue Bedeutung solcher Nebenbedingungen hängt stark von der jeweiligen Anwendung ab und kann durch B/E-Netze (ohne Nebenbedingungen) beschrieben werden.

In Stellen-Transitionsnetzen (S/T-Netzen) werden Schlingen (und somit Nebenbedingungen) von manchen Autoren zugelassen und von anderen ausgeschlossen.

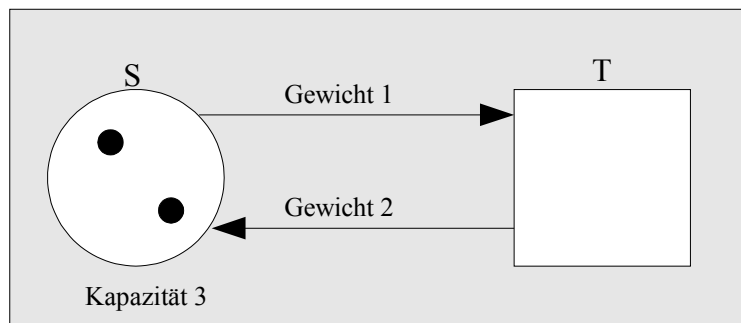


Abbildung 14.4: Eine Schlinge in einem S/T-Netz

Legt man eine Schaltregel zugrunde, die das Schalten einer Transition als Einheit behandelt (und nicht als eine Folge von zwei Teilvorgängen nimm-Token und gib-Token), dann kann die Transition T in Abb. 14.4 nicht schalten, da momentan auf der Stelle S nicht genug Platz ist, um zwei weitere Token darauf zu legen.

Die Schlinge in Abb. 14.4 ist also nicht gleichbedeutend mit dem folgenden Netz:

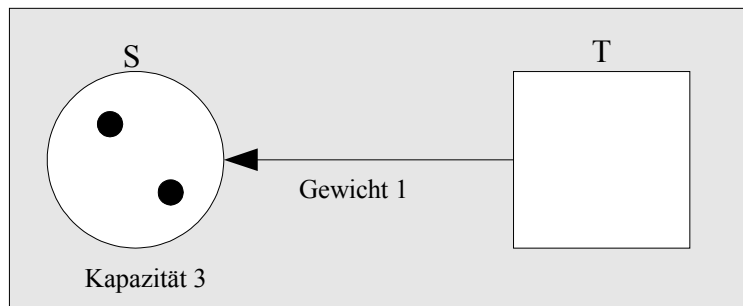


Abbildung 14.5: Kein Ersatz für Abb. 14.4

obwohl in beiden Netzen ein Schalten der Transition T als "Netto-Effekt" die Anzahl der Token auf der Stelle S um 1 erhöht.

Ein S/T-Netz **ohne** Schlingen kann man durch **eine** Matrix darstellen. Die Matrix enthält für jede Transition T_i eine Zeile und für jede Stelle S_j eine Spalte. In der Matrix stehen ganze Zahlen (positive und negative). Die Zahl a_{ij} in der Matrix bedeutet: wenn T_i schaltet wird die Anzahl der Token auf S_j um a_{ij} verändert (erhöht, falls a_{ij} positiv ist oder erniedrigt, falls a_{ij} negativ ist). Ist a_{ij} positiv, so bedeutet das graphisch: von T_i läuft ein Pfeil nach S_j . Ist a_{ij} negativ, so bedeutet das graphisch: von S_j läuft ein Pfeil nach T_i .

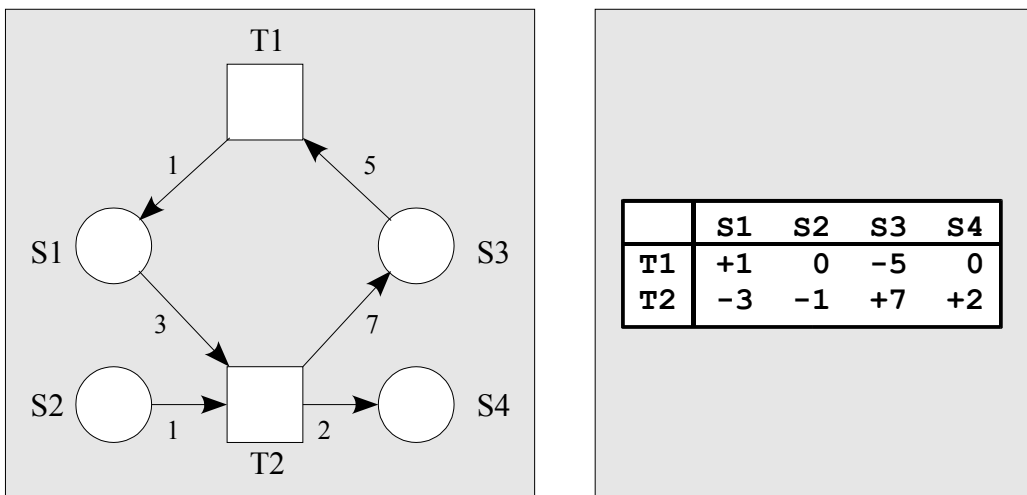


Abbildung 14.6: Ein S/T-Netz als Graph und als Matrix dargestellt

Man beachte: In der Matrix kann man nur entweder ein Pfeil von T_i nach S_j darstellen (durch ein positives a_{ij}) oder einen Pfeil von S_j nach T_i (durch ein negatives a_{ij}), aber nicht beide Pfeile.

Enthält ein S/T-Netz Schlingen, so braucht man deshalb **zwei** Matrizen, um es darzustellen. Die eine Matrix enthält nur positive Zahlen und die andere nur negative (und damit kann man die Vorzeichen auch weglassen und implizit voraussetzen).

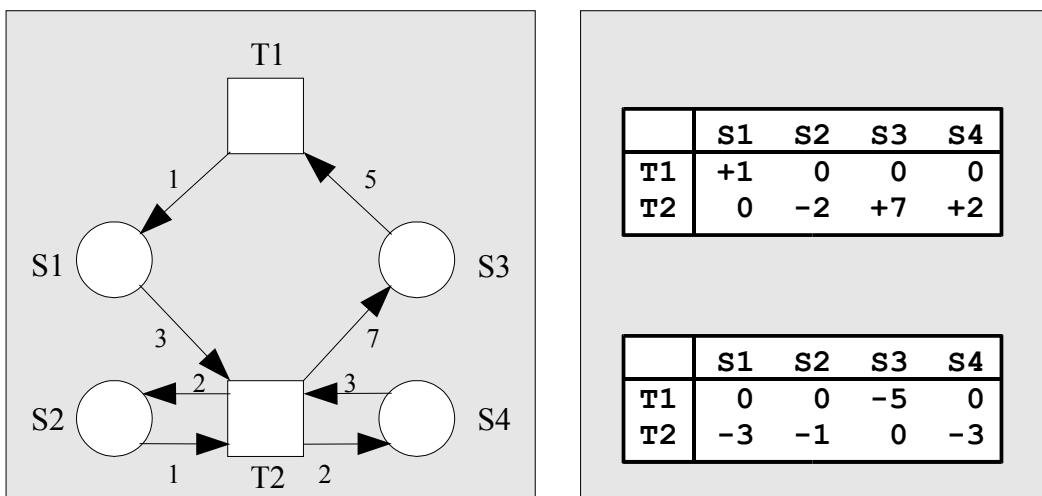


Abbildung 14.7: Ein S/T-Netz mit Schlingen als Graph und durch 2 Matrizen dargestellt

In farbigen Netzen sind "graphische Schlingen" immer erlaubt:

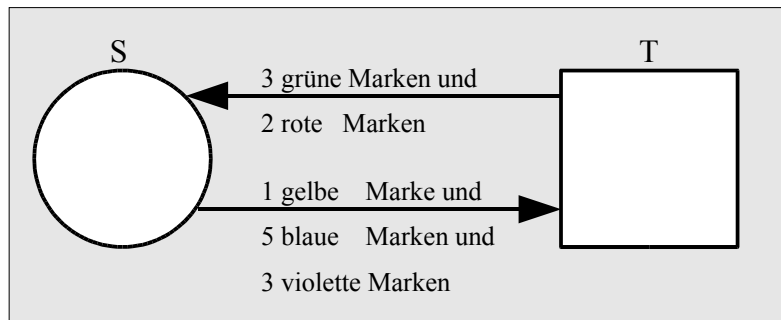


Abbildung 14.8: Eine "graphische" Schlinge in einem farbigen Netz

Bei solchen Schlingen werden in einem Schaltvorgang Marken einer bestimmten Farbe entweder von S genommen oder an S gegeben. Es werden aber nicht z. B. grüne Marken genommen **und** gegeben.

"Semantische Schlingen" werden in theoretischen Zusammenhängen meistens verboten und in praktischen Anwendungen erlaubt:

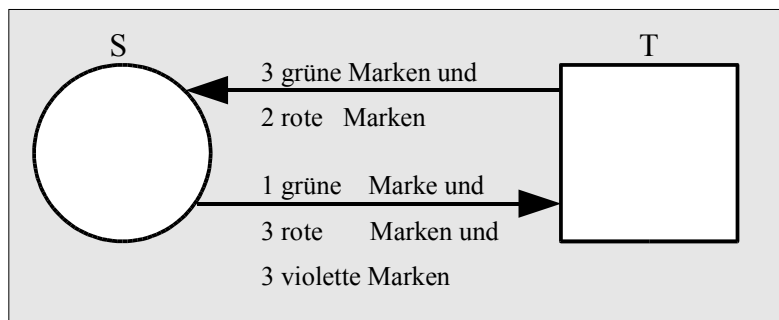


Abbildung 14.9: Eine "semantische" Schlinge in einem farbigen Netz

Bei solchen Schlingen werden in einem Schaltvorgang Marken bestimmter Farben (in Abb. 14.9 etwa grün und rot) sowohl gegeben als auch genommen.

15 S/T-Netz-Problem 1: Prozesse unterscheiden ununterscheidbare Token

Für markierte S/T -Netze gibt es, ähnlich wie für B/E-Netze, eine Prozessnetz-Semantik. Die Prozessnetze zu einem markierten S/T-Netz werden nach folgender Grundidee konstruiert: Trägt eine Stelle S_1 im S/T-Netz n Marken, wird diese Tatsache im Prozessnetz durch n runde Knoten (unwiederholbare Bedingungen) $S_{1.1}, S_{1.2}, \dots, S_{1.n}$ dargestellt:

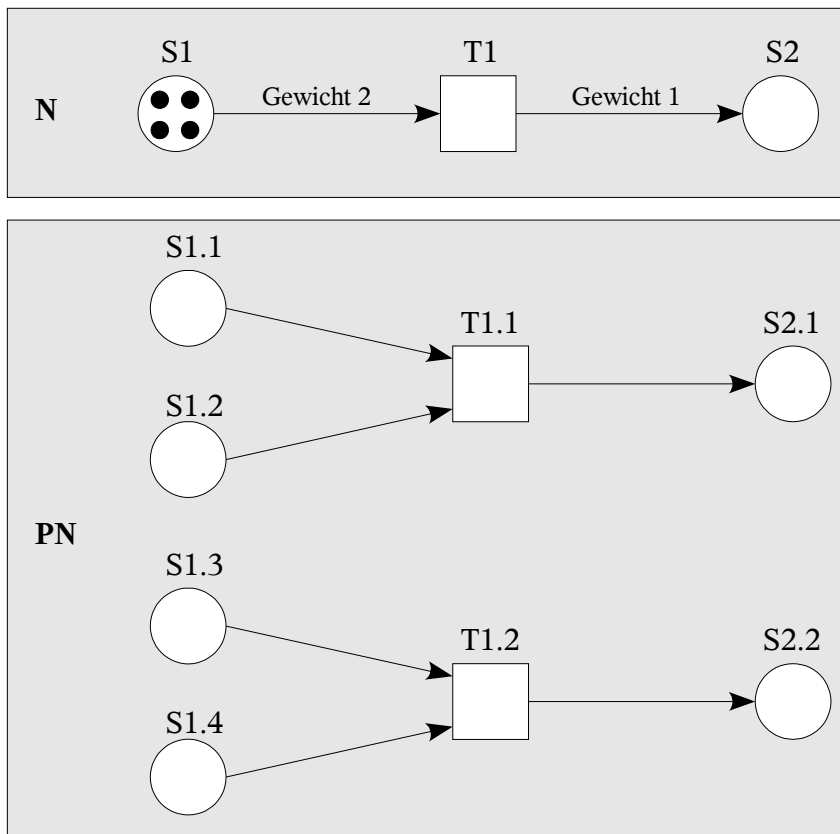


Abbildung 15.1: Ein markiertes S/T-Netz N und das entsprechende Prozessnetz PN

Die Prozessnetz-Semantik für S/T-Netze ist an einem bestimmten Punkt problematisch, wie das folgende Beispiel zeigen soll.

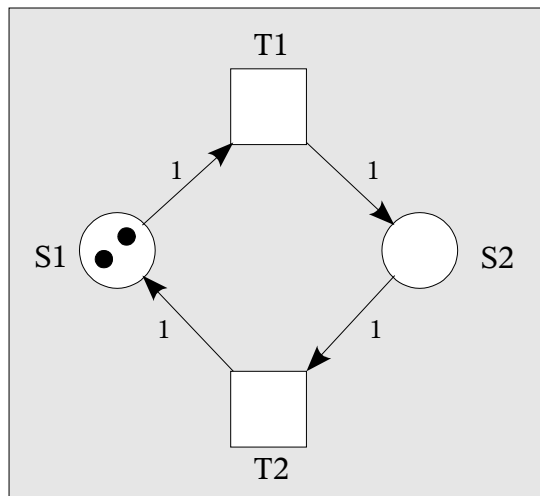


Abbildung 15.2: Ein zyklisches S/T-Netz

In diesem Netz kann man, anschaulich gesprochen, ein Token "im Kreis herum bewegen", indem man zuerst T1 und dann T2 schalten lässt. Ein Anfangsstück eines entsprechenden Prozessnetzes sieht so aus:

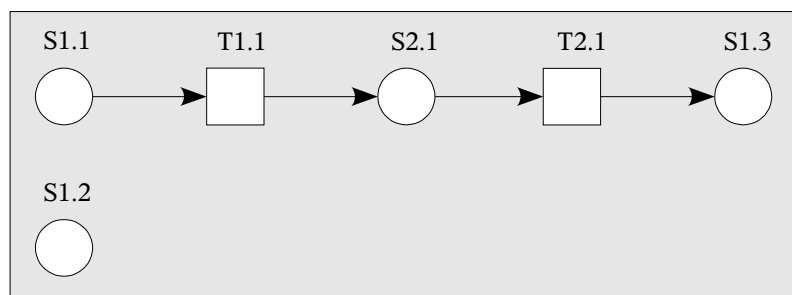


Abbildung 15.3: Anfangsstück eines Prozessnetzes zu Abb. 15.2

Nach einem "Rundlauf" ist bei dem markierten S/T-Netz die Anfangssituation wiederhergestellt: auf der Stelle S1 liegen wieder zwei Token und auf S2 keine Token. Im Prozessnetz werden die beiden Token auf S1 durch die Knoten S 1.2 und S 1.3 repräsentiert. Im S/T-Netz sind die beiden Token auf S1 nicht unterscheidbar. Im Prozessnetz ist es dagegen offenbar so, dass S1.2 das Token repräsentiert, welches von Anfang an auf S1 lag, während S1.3 das Token darstellt, welches durch den ersten Schaltvorgang von T2 (d.h. durch T2.1) auf S1 gelegt wurde.

Spielt man das Token-Spiel auf dem S/T-Netz, so gibt es nach einem Rundlauf (d.h. nachdem T1 und T2 je einmal geschaltet haben) genau eine Fortsetzung: T1 kann ein zweites Mal schalten.

Dieser einen Fortsetzung auf der S/T-Netz-Ebene entsprechen zwei verschiedene Fortsetzungen F1 und F2 auf der Prozessnetz-Ebene:

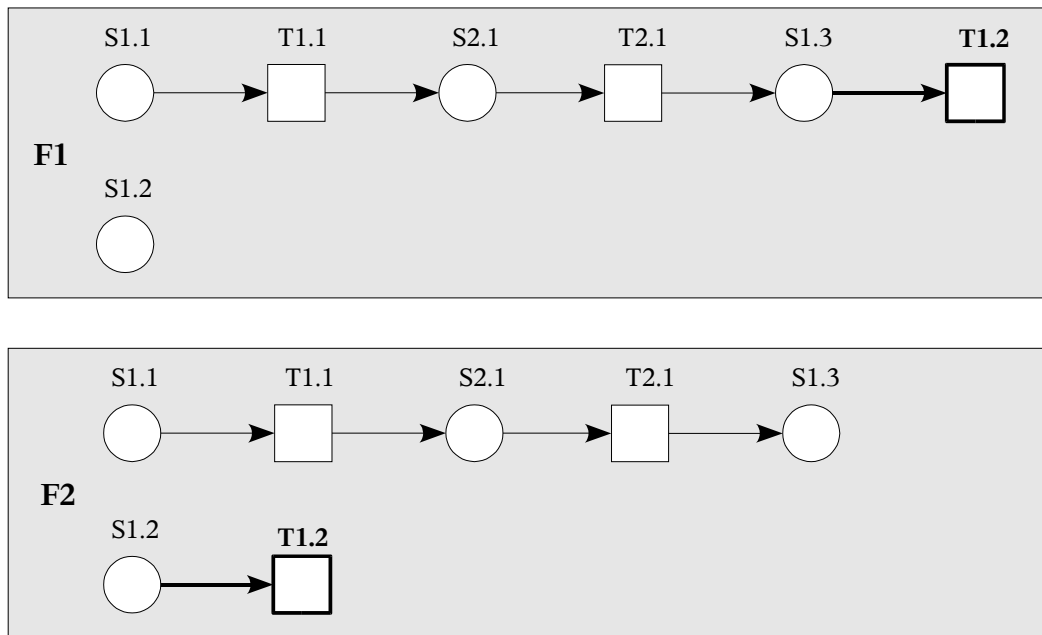


Abbildung 15.4: Zwei Fortsetzungen F1 und F2 zu Abb. 15.3

F1 und F2 unterscheiden sich darin, welches der (ununterscheidbaren) Token auf S1 die Transition T1 wegnimmt.

Prozessnetze als Semantik für markierte S/T-Netze **unterscheiden** also unter bestimmten Umständen zwischen **ununterscheidbaren** Token. Das ist ein ernsthaftes Problem, da die Ununterscheidbarkeit der Token ("nur die Anzahl ist wichtig, nicht die Individualität") ein grundlegendes Charakteristikum von S/T-Netzen ist. Man kann das Problem auch so beschreiben: Wenn man gewisse Ereignisse voneinander unterscheidet, dann ist es schwer, das, was sie konsumieren und produzieren, nicht zu unterscheiden.

In der Literatur gibt es Vorschläge, wie man das Problem der Prozessnetze für markierte S/T-Netze technisch behandeln kann. (z. B. durch Äquivalenz-Relationen auf Prozessnetzen. Entsprechend diesen Relationen sind Prozesse, die "sich nur aufgrund von ununterscheidbare Token unterscheiden" äquivalent). Diese Vorschläge vermögen aber nicht, das Problem in einem tieferen Sinne zu lösen. Es zeichnet sich auch nicht ab, dass eine "bessere" Prozessnetz-Semantik für S/T-Netze gefunden werden könnte. Die Abstraktion von der Individualität (von Marken) zu ihrer bloßen Anzahl (repräsentiert durch ununterscheidbare Token) scheint im Zusammenhang mit Nebenläufigkeit in einem tieferen Sinne problematisch zu sein.

16 S/T-Netz-Problem 2: Wie nebenläufig sind S/T-Netze?

In vielen Veröffentlichungen über Stellen-Transitionsnetze (S/T-Netze) wird eine Aktivierungsregel und eine Schalt-Regel für einzelne Transitionen angegeben ("Unter welchen Umständen kann eine Transition schalten?" und "Was passiert im Einzelnen, wenn eine Transition schaltet?"). Varianten dieser Regeln unterscheiden sich vor allem in den folgenden beiden Punkten:

1. Sind Schlingen erlaubt oder nicht? (siehe Abschnitt 14. und Abschnitt 6.)
2. Werden beim Schalten einer Transition T zuerst Token von den Vorstellen genommen und dann Token an die Nachstellen gegeben oder werden zuerst Token gegeben und dann genommen oder geschieht das Geben und Nehmen koinzident ("im selben Moment")? (siehe Abschnitt 6.)

Der zweite Punkt ist nur dann wichtig, wenn Schlingen erlaubt sind. Ohne Schlingen macht es keinen formal-harten Unterschied, ob man Token erst nimmt und dann gibt, erst gibt und dann nimmt oder koinzident nimmt und gibt.

In vielen Veröffentlichungen über S/T-Netze wird der Leser aber nicht darauf hingewiesen, dass durch die Aktivierungs- und die Schaltregel für einzelne Transitionen noch nichts über den Aspekt der Nebenläufigkeit von Schaltvorgängen gesagt ist (Unter welchen Umständen können Transitionen nebenläufig schalten?).

Tatsächlich gibt es für S/T-Netze *nicht nur eine* Möglichkeit, Nebenläufigkeit zu definieren, sondern ein ganzes Spektrum von Möglichkeiten. Und da Nebenläufigkeit in Netzen eng mit den Begriffen "Konflikt" und "Schritt" zusammenhängt, gibt es für S/T-Netze auch verschiedene Konflikt- und Schritt-Begriffe. Es ist nicht ganz leicht, in der Netz-Literatur Hinweise auf diese Tatsache zu finden.

Auf wie vielfältige Weise man den Aspekt der Nebenläufigkeit in S/T-Netzen präzisieren kann, lässt sich schon an dem folgenden sehr einfachen Beispiel erkennen:

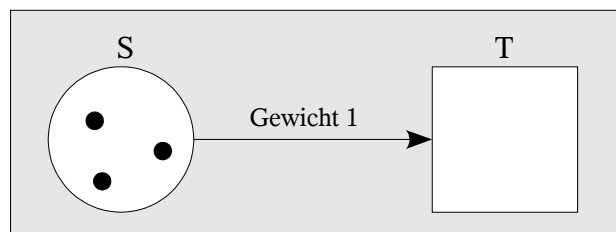


Abbildung 16.1: Wie nebenläufig ist die Transition T?

Bei jedem Schaltvorgang nimmt T ein Token von S (und gibt möglicherweise ein Token irgendwohin weiter, was hier nicht näher dargestellt ist). Die wesentliche Frage ist: Kann T nebenläufig zu sich selbst schalten? Oder, etwas allgemeiner: Wie viele Schaltvorgänge von T können nebenläufig stattfinden? Im Beispiel in Abb. 16.1 gibt es drei Antworten auf diese Frage:

1, 2 und 3.

Um zu sehen, dass alle drei Antworten sinnvoll sind, kann man das Netz in Abb. 16.1 anschaulich z. B. so interpretieren:

Die Stelle S stellt ein Lager dar, in dem Kisten gelagert werden. Jede Kiste wird durch ein Token auf S repräsentiert. Die Transition T stellt die Tür (oder die Auslagerungsvorrichtung) dar, durch die

Kisten aus dem Lager entnommen werden. Ein Schaltvorgang von T entspricht dem Vorgang, dass eine Kiste aus dem Lager entfernt wird.

Je nachdem, wie breit die Tür ist, kann man jeweils nur eine Kiste nach der anderen auslagern oder aber zwei, drei etc. Kisten nebenläufig entnehmen. Im Extremfall kann man alle Kisten, die sich im Lager befinden, nebenläufig zueinander auslagern.

Dieses Beispiel soll illustrieren, dass es bei S/T-Netzen ein ganzes Spektrum von Möglichkeiten gibt, Nebenläufigkeit von Schaltvorgängen zu erlauben bzw. zu verbieten.

Entsprechend vielfältig sind die Möglichkeiten festzulegen, was ein Konflikt in einem markierten S/T-Netz ist. Dazu ein einfaches Beispiel:

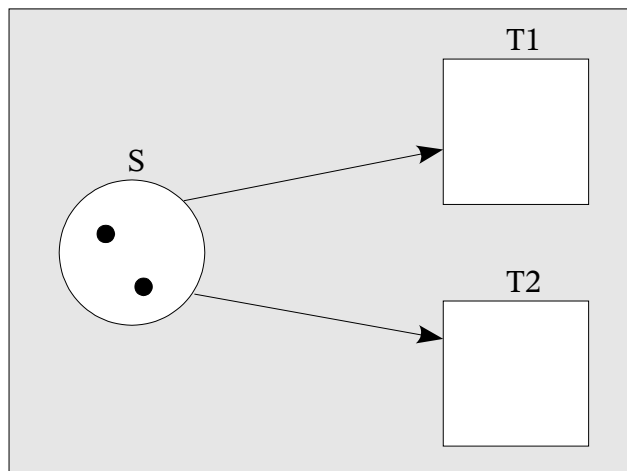


Abbildung 16.2: Stehen T1 und T2 in Konflikt miteinander?

Allgemein einig ist man sich darüber, dass die Transitionen T1 und T2 in Abb. 16.2 nebenläufig zueinander schalten können. Daraus kann man schließen, dass T1 und T2 nicht in Konflikt miteinander stehen.

Man kann aber auch so argumentieren: T1 könnte mit den beiden Token auf S zweimal (nebenläufig zu sich selbst) schalten. Das Gleiche gilt für T2. Also sind insgesamt vier Schaltvorgänge aktiviert, von denen aber nur zwei stattfinden können (da auf S nur zwei Token liegen und jeder Schaltvorgang ein Token verbraucht). Also stehen T1 und T2 in Konflikt miteinander.

Diese Argumentation wird (hoffentlich) noch klarer, wenn man den Konflikt-Begriff folgendermaßen modifiziert: Für B/E-Netze wurden Konflikte zwischen Ereignissen (eckigen Netzknoten) definiert. Stattdessen kann man Konflikt auch zwischen Schritten zu definieren.

Ein Schritt (in einem markierten S/T-Netz) ist eine Multimenge von Transitionen, d.h. jede Transition kann mehrfach zu einem Schritt gehören.

Zwei Schritte stehen in Konflikt miteinander, wenn jeder einzelne von ihnen aktiviert ist, ihre Summe aber nicht aktiviert ist ("weil ein Schritt dem anderen Token oder freie Kapazität auf einer Stelle wegnimmt").

Mit diesem Konflikt-Begriff kann man für Abb. 16.2 so argumentieren: Der Schritt $s_1 = \{T_1, T_1\}$ ist aktiviert. Ebenso ist der Schritt $s_2 = \{T_2, T_2\}$ aktiviert. Der Schritt $s_3 = \{T_1, T_1, T_2, T_2\}$ (die Summe von s_1 und s_2) ist dagegen nicht aktiviert. Also stehen die Schritte s_1 und s_2 in Konflikt

miteinander. Andererseits stehen z. B. die Schritte $s_4 = \{T_1\}$ und $s_5 = \{T_2\}$ nicht in Konflikt, da ihre Summe $s_6 = s_4 + s_5 = \{T_1, T_2\}$ aktiviert ist.

Wie nebenläufig und wie konfliktträchtig man S/T-Netze interpretieren will, kann man auch dadurch präzisieren, dass man eine Übersetzung von (markierten) S/T-Netzen in (markierte) B/E-Netze angibt. Denn für B/E-Netze ist klar, wann Ereignisse nebenläufig eintreten können und wann sie in Konflikt stehen.

Ist die Übersetzung eindeutig, dann erhält man dadurch genau einen Begriff von Nebenläufigkeit bzw. von Konflikt auf der S/T-Ebene. Erlaubt man dagegen, dass ein S/T-Netz in verschiedene B/E-Netze übersetzt wird, dann erhält man für die S/T-Ebene entsprechend viele Definitionen für Nebenläufigkeit und Konflikt.

Im Folgenden soll eine Übersetzung von markierten S/T-Netzen in markierte B/E-Netze skizziert werden, die in starkem Maße mehrdeutig ist. D.h. man kann während der Übersetzung eine Reihe von Entscheidungen treffen und dadurch ausdrücken, ob man das S/T-Netz mehr oder weniger nebenläufig bzw. konfliktträchtig verstanden wissen möchte.

Die Übersetzung eines markierten S/T-Netzes in ein markiertes B/E-Netz wird anhand eines typischen (Teil-) Netzes erläutert.

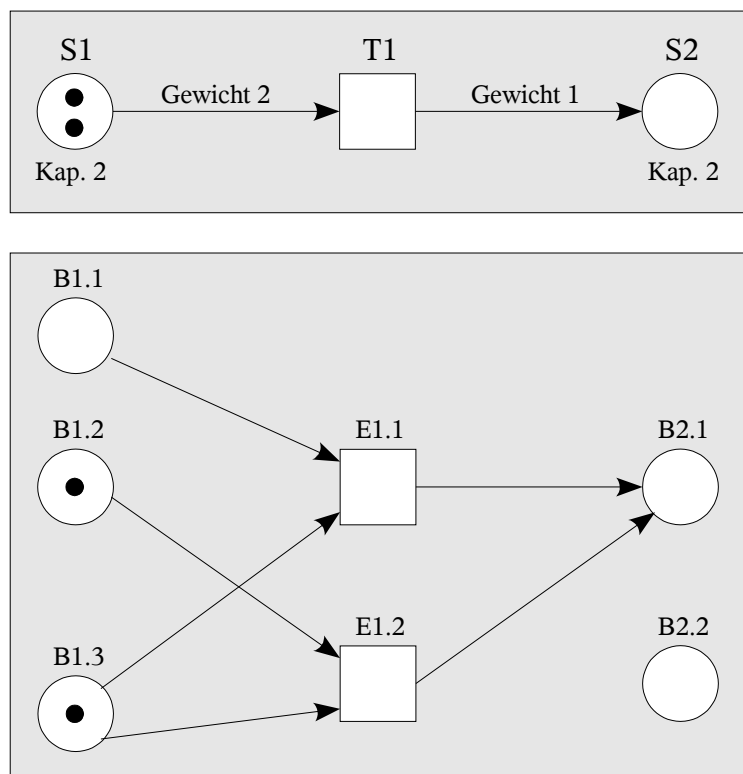


Abbildung 16.3: Ein typisches S/T-(Teil-)Netz und eine seiner Übersetzungen in ein B/E-Netz

1. Jede Stelle mit Kapazität n wird in n Bedingungen übersetzt. Eine Stelle mit unbegrenzter Kapazität wird in abzählbar-unendlich viele Bedingungen übersetzt.

In Abb. 16.3 entsprechen die drei Bedingungen B1.1, B1.2 und B1.3 der Stelle S mit Kapazität 3. Man beachte, dass die Namen "B1.1", "B1.2", ... "E1.1", "E1.2" etc. hier wiederholbare Bedingun-

gen bzw. Ereignisse bezeichnen, und nicht (wie in anderen Abschnitten) unwiederholbare Bedingungen und Ereignisse.

2. Ist eine Stelle mit m Token markiert, dann müssen genau m der entsprechenden Bedingungen (mit je einem Token) markiert werden. Der Übersetzer kann auswählen, welche m Bedingungen er markiert und welche nicht (falls m kleiner als die Kapazität der betreffenden Stelle ist).

In Abb. 16.3 liegen 2 Token auf $S1$. Entsprechend sind zwei der drei Bedingungen $B1.1$, $B1.2$, $B1.3$ markiert (nämlich $B1.2$ und $B1.3$). Genau so gut hätte der Übersetzer auch $B1.1$ und $B1.2$ oder $B1.2$ und $B1.3$ markieren können.

3. Jede Transition wird in ein Ereignis oder in mehrere Ereignisse übersetzt. Der Übersetzer kann (innerhalb gewisser Grenzen) wählen, in wieviele Ereignisse er eine Transition übersetzen will.

In Abb. 16.3 hat der Übersetzer die Transition $T1$ in zwei Ereignisse $E1.1$ und $E2.1$ übersetzt..

4. Ein Pfeil mit Gewicht g muss in g entsprechende Pfeile übersetzt werden. Genauer: Sei $S1$ eine Stelle, $T1$ eine Transition und P ein Pfeil von $S1$ nach $T1$ mit Gewicht g . Seien $B1.1$, $B1.2$... die Bedingungen, in die $S1$ übersetzt wurde und sei $E1.1$ eines der Ereignisse, in die $T1$ übersetzt wurde. Dann muss von genau g vielen der der Bedingungen $B1.1$, $B1.2$, ... je ein Pfeil nach $E1.i$ führen. Entsprechendes gilt, wenn der Pfeil P von $T1$ nach $S1$ führt.

In Abb. 16.3 führt ein Pfeil P mit Gewicht 2 von $S1$ nach $T1$. Diesem Pfeil entsprechen sowohl die beiden Pfeile, die von $B1.1$ und $B1.3$ nach $E1.1$ führen, als auch die beiden Pfeile, die von $B1.2$ und $B1.3$ nach $E1.2$ führen. Von welchen zwei der drei Bedingungen $B1.1$, $B1.2$, $B1.3$ Pfeile nach $E1.1$ bzw. nach $E1.2$ führen, hat der Übersetzer gewählt. Genauso gut hätte er wählen können, zwei Pfeile von $B1.1$ und $B1.2$ nach $E1.1$ zu zeichnen.

5. Ist $T1$ eine Transition und sind $E1.i$ und $E1.j$ zwei der Ereignisse, in die $T1$ übersetzt wurde, dann dürfen $E1.i$ und $E1.j$ nicht alle Vor- und alle Nachbedingungen gemeinsam haben.

In Abb. 16.3 haben die beiden Ereignisse $E1.1$ und $E1.2$ die Vorbedingung $B1.3$ und die Nachbedingung $B2.1$ gemeinsam. Sie unterscheiden sich aber durch die Vorbedingung $B1.1$ (von $E1.1$) bzw. $B1.2$ (von $E1.2$).

Die Forderung 5. begrenzt die Anzahl der Ereignisse, in die eine Transition übersetzt werden darf (falls alle Stellen in der Umgebung der Transition endliche Kapazitäten haben). In Abb. 16.3 könnte man die Transition $T1$ maximal in sechs Ereignisse $E1.1$... $E1.6$ übersetzen, da es sechs Möglichkeiten gibt, unter den drei Vorbedingungen $B1.1$, $B1.2$, $B1.3$ zwei und unter den zwei Nachbedingungen $B2.1$, $B2.2$ eine auszuwählen.

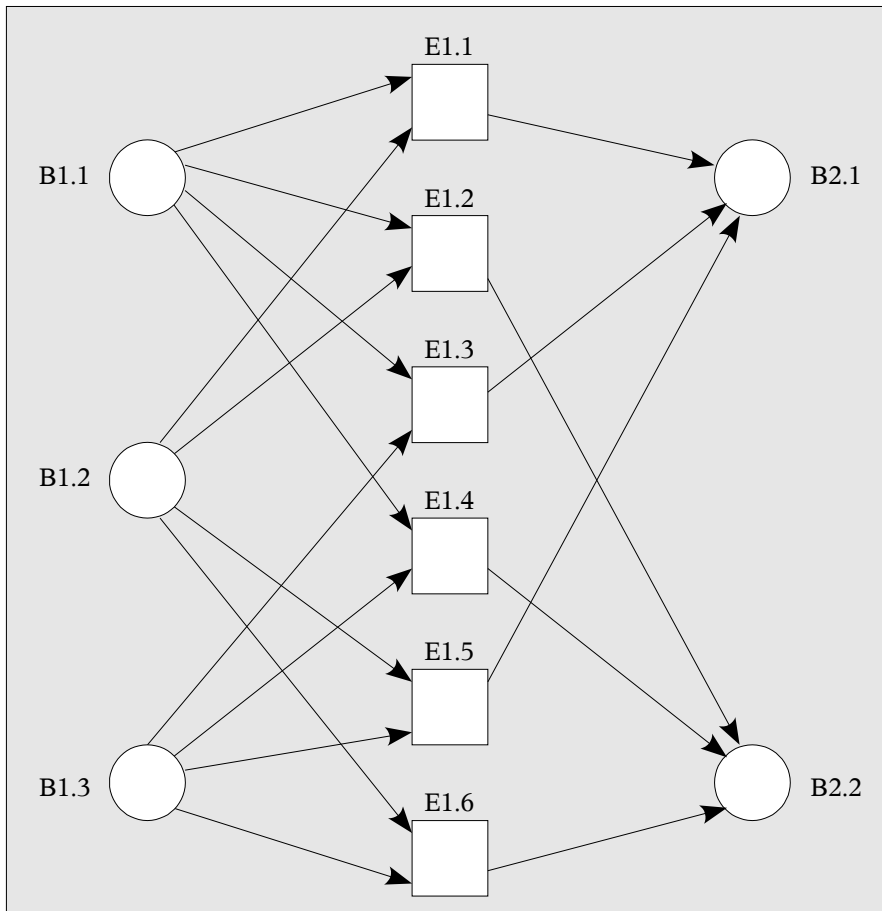


Abbildung 16.4: Die maximale Übersetzung zu Abb. 16.3

Eine minimale Übersetzung liegt dann vor, wenn man jede Transition T in nur ein Ereignis übersetzt. Im Allgemeinen gibt es zu einem S/T-Netz **viele minimale**, aber immer nur genau **eine maximale** Übersetzung.

Im Folgenden sind für ein sehr einfaches S/T-Netz alle (nach dem oben beschriebenen Verfahren möglichen) Übersetzungen in B/E-Netze angegeben. Der Einfachheit halber wurde das S/T-Netz mit der leeren Markierung versehen.

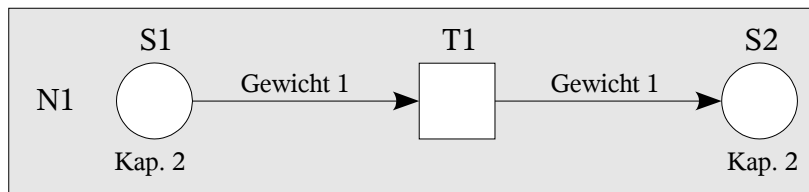


Abbildung 16.5: Ein sehr einfaches S/T-Netz N1 mit leerer Markierung

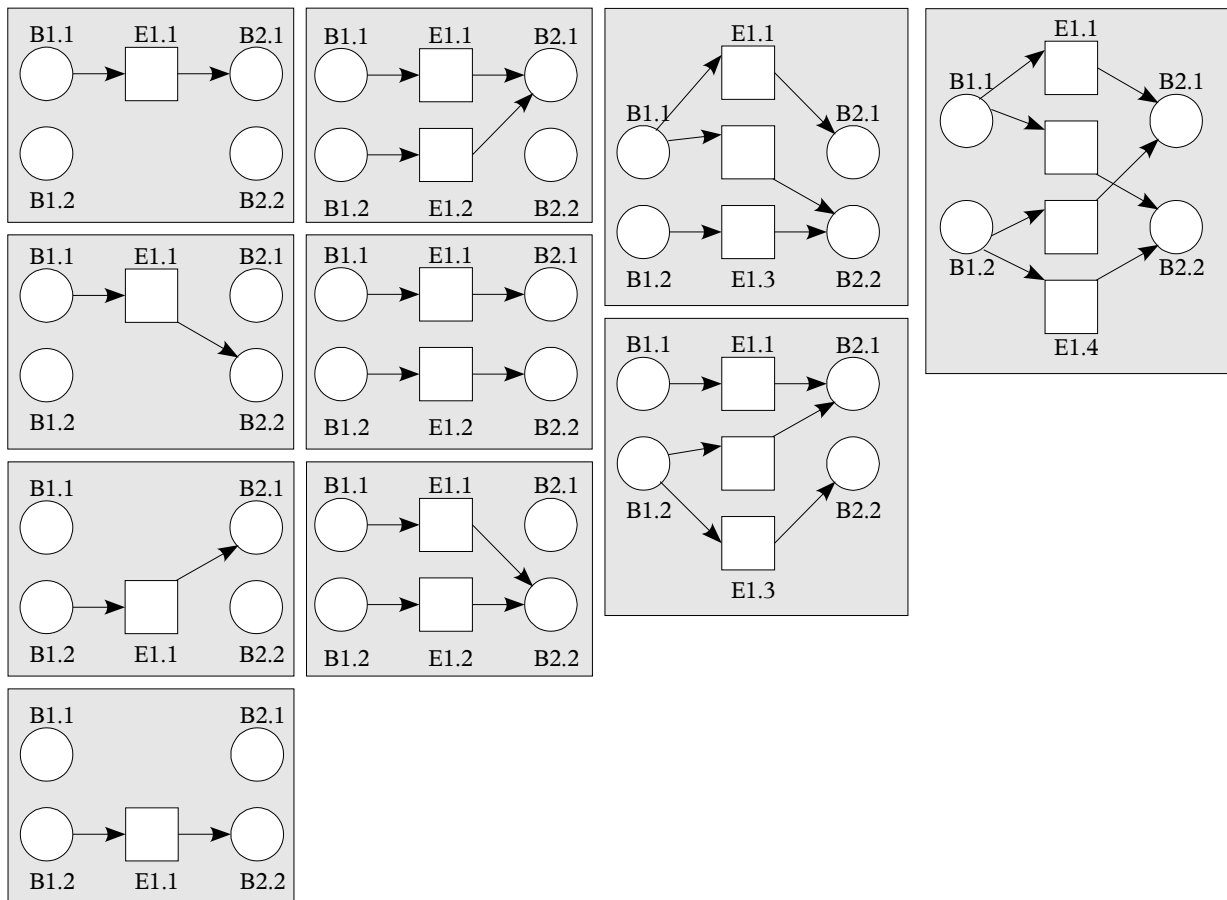


Abbildung 16.6: Alle Übersetzungen des Netzes N1

In der ersten Spalte sind alle Übersetzungen angegeben, bei denen die Transition T1 in ein Ereignis E1.1 übersetzt wurde. In der nächsten Spalte haben alle Übersetzungen zwei Ereignisse und in der nächsten drei. Die maximale Übersetzung steht in der ganz rechten Spalte und umfasst vier Ereignisse.

Sechs der zehn B/E-Netze enthalten isolierte Bedingungen, d.h. Bedingungen, zu denen kein Pfeil hinführt und von denen kein Pfeil wegführt. In praktischen Anwendungen sind solche isolierten Netzknoten in aller Regel nicht sinnvoll und werden meistens ausgeschlossen. Damit bleiben noch vier sinnvolle Übersetzungen.

Das folgende S/T-Netz hat insgesamt neun Übersetzungen, davon sind sieben ohne isolierte Knoten:

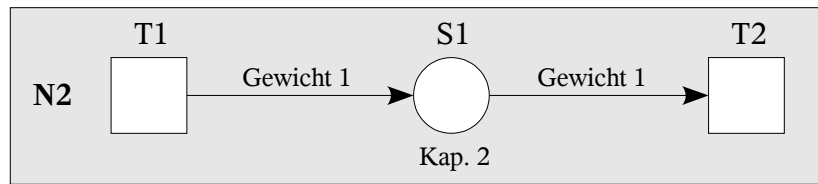


Abbildung 16.7: Noch ein einfaches S/T-Netz N2 mit der leeren Markierung

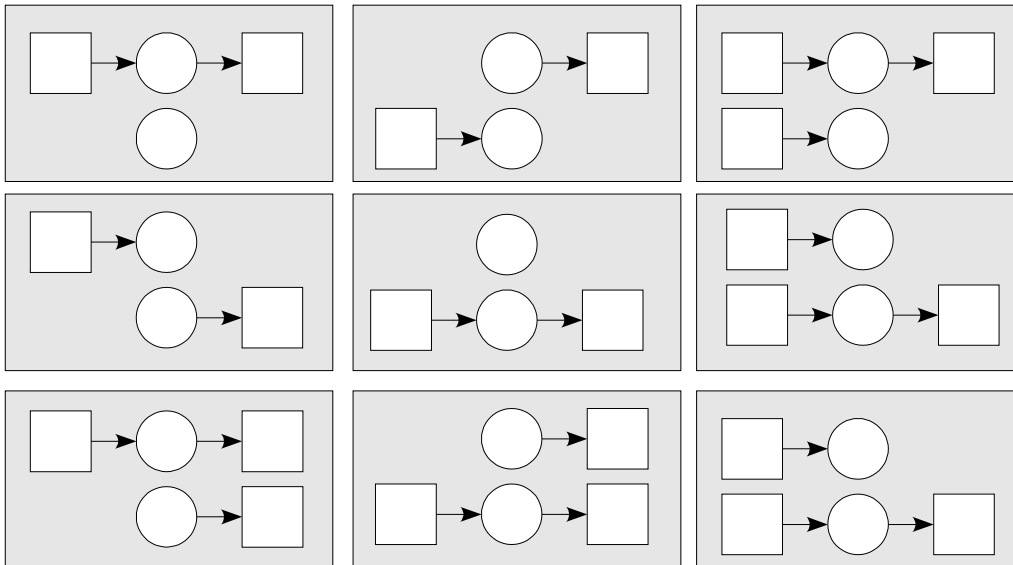


Abbildung 16.8: Alle Übersetzungen des Netzes N2

Jeder Schritt in einem B/E-Netz N_i , welches eine Übersetzung eines S/T-Netzes N ist, entspricht genau einem Schritt im Netz N . Umgekehrt entsprechen im Allgemeinen jedem Schritt s_1 im S/T-Netz N in jedem seiner Übersetzungen N_i mehrere Schritte $s_{1.1}, s_{1.2}, \dots$. Ob zwei S/T-Schritte s_1 und s_2 in Konflikt stehen oder nebenläufig sind, kann man jetzt definitorisch auf verschiedene Weise davon abhängig machen, ob die entsprechenden B/E-Schritte $s_{1.1}, s_{1.2}, \dots$ und $s_{2.1}, s_{2.2}, \dots$ in Konflikt stehen oder nebenläufig sind, z. B. so:

Maximale Nebenläufigkeit: s_1 und s_2 sind nebenläufig, wenn in mindestens einer Übersetzung N_i einer der Schritte $s_{1.1}, s_{1.2}, \dots$ nebenläufig ist zu einem der Schritte $s_{2.1}, s_{2.2}, \dots$.

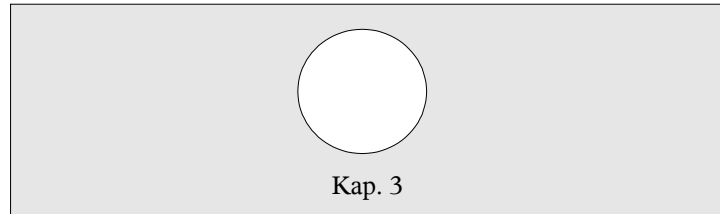
Minimaler Konflikt: s_1 und s_2 stehen in Konflikt miteinander, wenn in jeder Übersetzung N_i jeder der Schritte $s_{1.1}, s_{1.2}, \dots$ mit jedem der Schritte $s_{2.1}, s_{2.2}, \dots$ in Konflikt steht.

Minimale Nebenläufigkeit: s_1 und s_2 sind nebenläufig, wenn in jeder Übersetzung N_i jeder der Schritte $s_{1.1}, s_{1.2}, \dots$ nebenläufig ist zu jedem der Schritte $s_{2.1}, s_{2.2}, \dots$.

Maximaler Konflikt: s_1 und s_2 stehen in Konflikt zueinander, wenn in mindestens einer Übersetzung N_i einer der Schritte $s_{1.1}, s_{1.2}, \dots$ mit einem der Schritte $s_{2.1}, s_{2.2}, \dots$ in Konflikt steht.

Maximale Nebenläufigkeit und minimaler Konflikt gehören zusammen. Ebenso gehören minimale Nebenläufigkeit und maximaler Konflikt zusammen. Zwischen diesen extremen Kombinationen gibt es ein ganzes Spektrum von Möglichkeiten, Nebenläufigkeit und Konflikt in markierten Str-Netzen mit Hilfe der Übersetzungen in markierte B/E-Netze zu definieren.

Der oben skizzierte Algorithmus zur Übersetzung von markierten S/T-Netzen in markierte B/E-Netze ist dadurch charakterisiert, dass Stellen nur in Bedingungen und Transitionen nur in Ereignisse (d.h. runde Knoten nur in runde Knoten und eckige Knoten nur in eckige Knoten) übersetzt wurden. Lässt man diese Einschränkung fallen, dann erhält man noch viel mehr "mögliche Bedeutungen" für ein S/T-Netz. Z. B. kann eine Stelle der Kapazität 3 dann drei Bedingungen bedeuten, die auf verschiedene Arten durch Ereignisse miteinander verbunden (bzw. nicht verbunden) sind, z. B. so:



Dieses S/T-Netz kann unter anderem bedeuten:

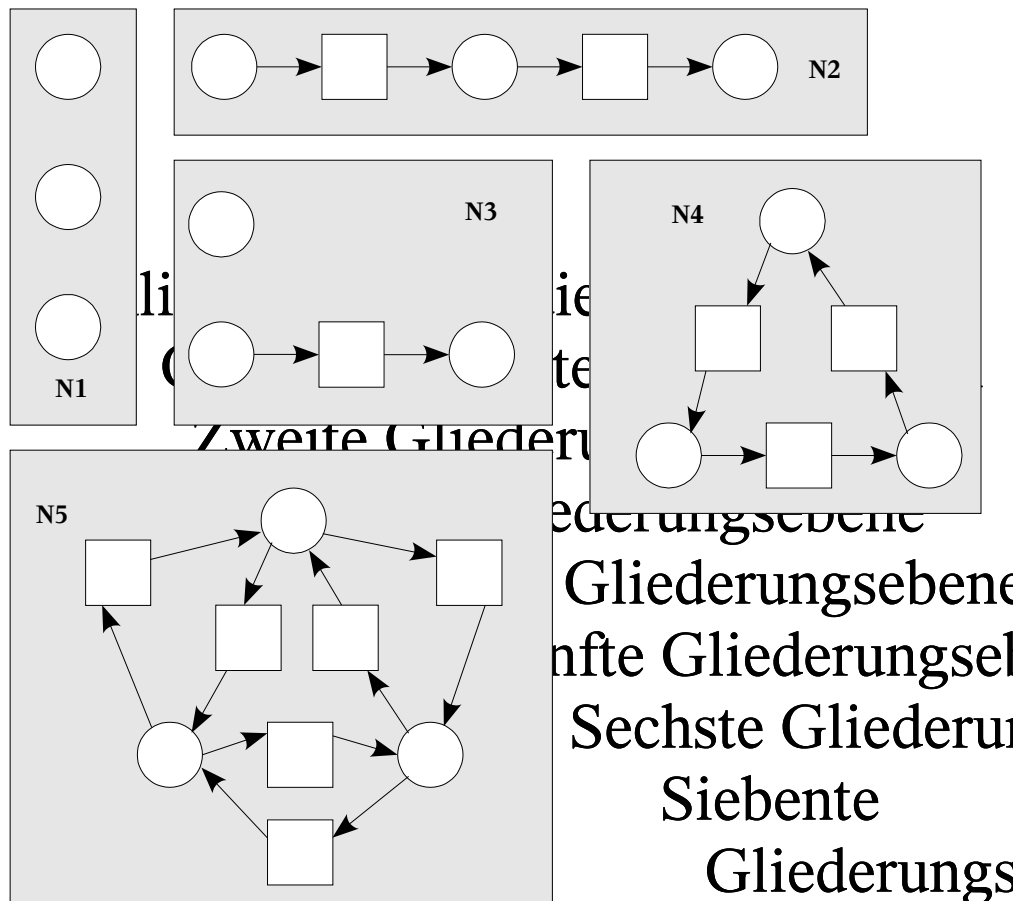


Abbildung 16.9: Verschiedene Übersetzungen einer Stelle mit Kapazität 3

Zusammenfassung: Die Aktivierungsregel und die Schalt-Regel für einzelne Transitionen in markierten S/T-Netzen werden in der Literatur weitgehend einheitlich festgelegt (mit Varianten bezüglich der Behandlung von Schlingen). Dadurch ist z. B. festgelegt, welche Markierungen von einer bestimmten Markierung aus erreichbar sind, ob ein Netz lebendig ist oder nicht etc. Dagegen wird über den Aspekt der Nebenläufigkeit und des Konflikts in S/T-Netzen in theoretischen Arbeiten

meistens gar nichts ausgesagt. Nur anhand von theoretisch-formalen Argumenten kann man wohl kaum einen ganz bestimmten Nebenläufigkeitsbegriff als den Nebenläufigkeitsbegriff für S/T-Netze auszeichnen. Nur im Zusammenhang mit einer konkreten Anwendung (oder einer Klasse von Anwendungen) lässt sich eine sinnvolle Übersetzung von markierten S/T-Netzen in markierte B/E-Netze (und damit eine Präzisierung des Nebenläufigkeitsaspekts) auswählen.

17 Netze in Matrix-Darstellung: Von den B/E-Netzen zu den farbigen Netzen

Netze kann man nicht nur graphisch darstellen, sondern z. B. auch durch bestimmte Matrizen. Die Matrix-Darstellung macht den Zusammenhang zwischen den vier wichtigsten Netz-Klassen (Bedingungs-Ereignisnetzen, Stellen-Transitionsnetzen, strikten farbigen Netzen und farbigen Netzen) besonders klar und durchsichtig.

Das folgende Beispiel zeigt, wie man ein B/E-Netz durch eine Matrix darstellen kann:

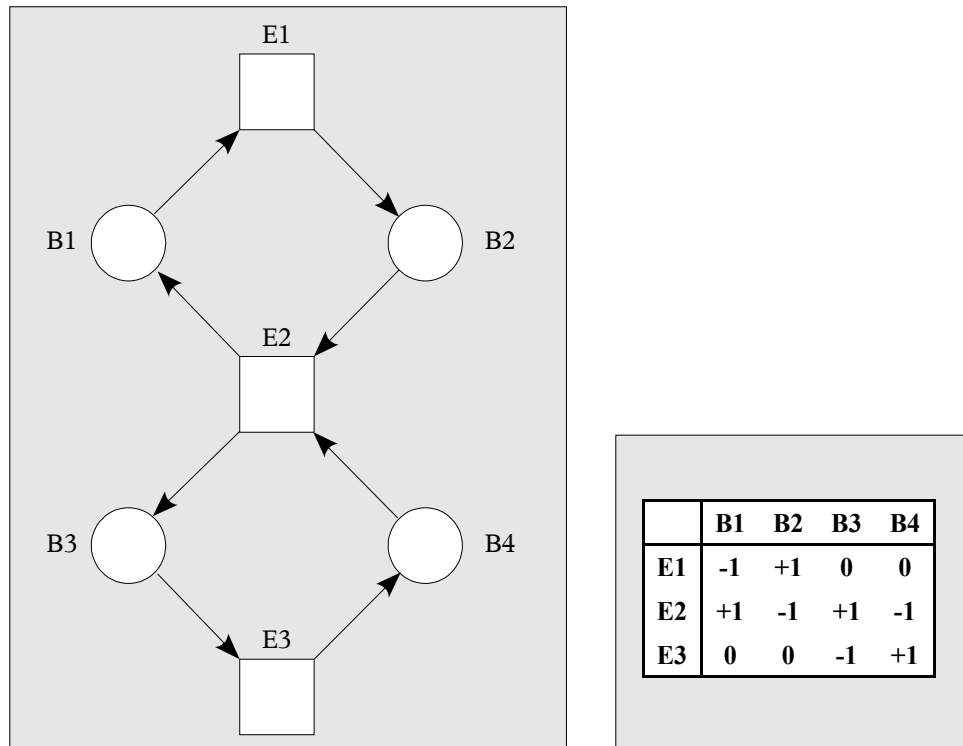


Abbildung 17.1: Ein B/E-Netz in graphischer Darstellung und durch eine Matrix dargestellt

Die Matrix enthält für jedes Ereignis eine Zeile und für jede Bedingung eine Spalte. Steht an der Kreuzung der E_i -Zeile mit der B_j -Spalte eine -1 , dann bedeutet das: Wenn das Ereignis E_i eintritt, wird ein Token von der Bedingung B_j entfernt. Eine $+1$ bedeutet, dass beim Eintreten von E_i ein Token auf B_j gelegt wird. Sind E_i und B_j nicht mit einem Pfeil verbunden, so wird das in der Matrix durch eine 0 ausgedrückt.

Mit einer Matrix kann man nur *einen* Pfeil zwischen zwei Knoten E_i und B_j darstellen: Einen Pfeil von B_j nach E_i durch eine -1 oder einen Pfeil von E_i nach B_j durch eine $+1$. Will man auch Schlingen erlauben (ein Pfeil von B_j nach E_i und einen Pfeil von E_i nach B_j), dann ist es günstiger, ein Netz durch zwei Matrizen darzustellen. In die eine trägt man alle Plus-Einsen ein und in die andere alle Minus-Einsen.

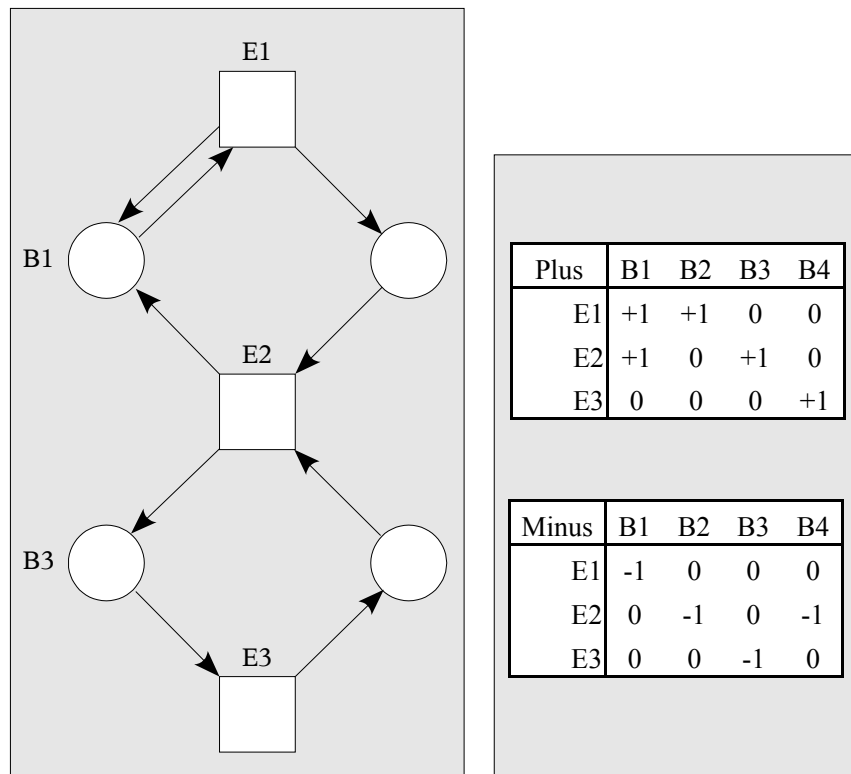


Abbildung 17.2: Ein B/E-Netz mit Schlingen graphisch und durch zwei Matrizen dargestellt

Da in der Plus-Matrix alle Einsen ein Plus-Vorzeichen und in der Minus-Matrix alle Einsen ein Minus-Vorzeichen haben, kann man diese Vorzeichen auch weglassen. Somit kann man ein B/E-Netz mit Schlingen durch zwei Matrizen darstellen, in denen nur Nullen und Einsen vorkommen.

Schlingen sind in B/E-Netzen nicht sinnvoll, da das Ereignis in einer Schlinge nie eintreten kann. Sie wurden hier aber schon erwähnt, um den Übergang zu S/T-Netzen (in denen Schlingen sinnvoll sein können und häufig zugelassen werden) zu vereinfachen.

In einem B/E-Netz fließt über jeden Pfeil jeweils ein Token. Entsprechend kann man ein B/E-Netz durch zwei 0-1-Matrizen darstellen.

In einem S/T-Netz fließen über jeden Pfeil jeweils mehrere Token. Entsprechend kann man ein S/T-Netz durch zwei Matrizen von natürlichen Zahlen darstellen.

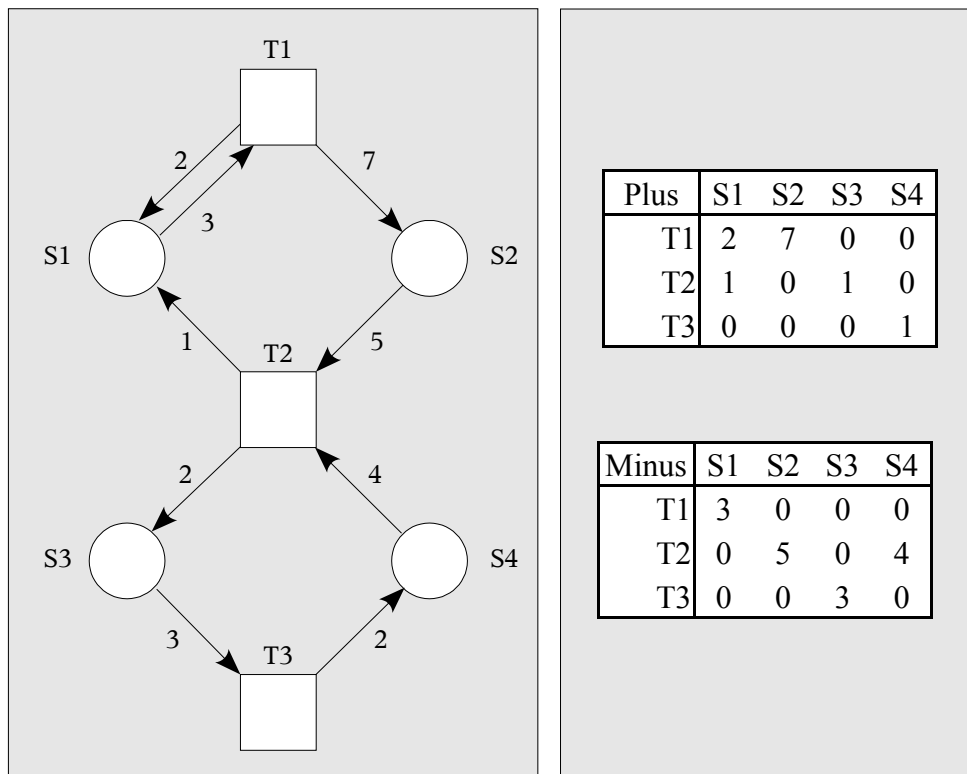


Abbildung 17.3: Ein S/T-Netz, graphisch und durch zwei Matrizen dargestellt

Den Zusammenhang zwischen der graphischen und der Matrix-Darstellung kann man auch so beschreiben: In der Plus-Matrix werden die Pfeile beschrieben, die von einem eckigen Knoten (Transition) zu einem runden Knoten (Stelle) führen. In der Minus-Matrix werden die Pfeile von rund nach eckig beschrieben. Enthält ein Netz keine Schlingen, so kann man die beiden Matrizen zu einer Matrix verschmelzen (indem man die Zahlen in den Matrizen mit ihren Vorzeichen versieht und die beiden Matrizen addiert, siehe Abb. 17.1). Im Folgenden wird aber weiterhin die allgemeinere Darstellung eines Netzes durch zwei Matrizen beibehalten.

Zur vollständigen Darstellung eines S/T-Netzes gehört außer den beiden Matrizen Plus und Minus noch ein Vektor, der die Kapazität der einzelnen Stellen enthält. Für das Netz in Abb. 16.3 könnte dieser Vektor z. B. so aussehen:

	S1	S2	S3	S4
Kapazität	3	14	ω	12

Abbildung 17.4: Ein Kapazitäts-Vektor für Abb. 17.3

In diesem Beispiel hat die Stelle S3 eine unbegrenzte Kapazität (" ω ").

Stellen-Transitions-Netze und farbige Netze scheinen, wenn man die üblichen formalen Definitionen anschaut, sehr verschiedene Arten von Netzen zu sein. Legt man dagegen die Matrix-Darstellung zugrunde, dann wird deutlich, dass ein **farbiges Netz** eigentlich ein **S/T-Netz** mit ein

wenig zusätzlicher Struktur ist. Und entsprechend ist ein **strikttes farbiges Netz** nichts weiter als ein **B/E-Netz** mit ein wenig zusätzlicher Struktur.

Der Übergang von einem S/T-Netz zu einem entsprechenden farbigen Netz besteht im Wesentlichen darin, dass man die Stellen und die Transitionen des S/T-Netzes beliebig **zu Gruppen zusammenfasst**. Jede Gruppe S_1, S_2, \dots, S_n von Stellen des S/T-Netzes wird zu einer Stelle S des farbigen Netzes. Die S/T-Stellen S_1, S_2, \dots, S_n werden dann als **die Farben der Stelle S** bezeichnet. Entsprechend wird jede Gruppe T_1, T_2, \dots, T_m von Transitionen des S/T-Netzes zu einer Transition T des farbigen Netzes zusammengefasst und die S/T-Transitionen T_1, T_2, \dots, T_m werden als **die Farben der Transition T** bezeichnet.

Die Aktivierungs-Regel und die Schalt-Regel für farbige Netze stimmen im Grunde genommen mit denen für S/T-Netze überein. Unterschiedlich sind eigentlich nur die Redeweisen, die man üblicherweise auf der S/T-Ebene bzw. auf der farbigen Ebene verwendet.

Statt auf der S/T-Ebene zu sagen: "Die Transition T_3 ist aktiviert", sagt man auf der farbigen Ebene: "Die Transition T ist mit ihrer Farbe T_3 aktiviert". Und statt auf der S/T-Ebene zu sagen: "Auf der Stelle S_5 liegen 3 Token", sagt man auf der farbigen Ebene: "Auf der Stelle S liegen drei Marken der Farbe S_5 ".

Praktisch lässt sich das Gruppieren der Stellen und Transitionen eines S/T-Netzes N zu den Stellen und Transitionen eines farbigen Netzes z. B. so durchführen:

1. Man vertauscht in den beiden Matrizen (Plus und Minus) von N solange Zeilen miteinander, bis die Transitionen, die man gern zusammenfassen möchte, jeweils unmittelbar untereinander stehen. Die Zeilenvertauschungen müssen in beiden Matrizen "gleichmäßig" vorgenommen werden (wenn man in der einen Matrix die 3. mit der 1. Zeile vertauscht, dann muss man das auch in der anderen Matrix tun, etc.).
2. Entsprechend vertauscht man (in beiden Matrizen gleichmäßig) Spalten, bis die Stellen, die man zusammenfassen möchte, unmittelbar nebeneinander stehen.
3. Dann macht man die Gruppierung der Zeilen und Spalten irgendwie deutlich, z. B. durch waagerechte und senkrechte "Trennlinien" durch die beiden Matrizen.

Als Beispiel wird das S/T-Netz aus Abb. 17.3 in ein entsprechendes farbiges Netz umgewandelt. Dabei sollen die Transitionen T_1 und T_3 zu einer Transition **T_1** zusammengefasst werden. T_2 soll einer Einergruppe **T_2** werden. Die Stellen S_1 und S_3 sollen zu **S_1** und S_2 und S_4 sollen zu **S_2** gruppiert werden.

Die beiden Matrizen des S/T-Netzes in Abb. 17.3 sehen ursprünglich so aus:

Plus	S1	S2	S3	S4	Minus	S1	S2	S3	S4
T1	2	7	0	0	T1	3	0	0	0
T2	1	0	2	0	T2	0	5	0	4
T3	0	0	0	2	T3	0	0	3	0

Abbildung 17.5: Die Matrizen des S/T-Netzes aus Abb. 17.3

Nach der Vertauschung von Zeile 2 und 3 hat man:

Plus	S1	S2	S3	S4
T1	2	7	0	0
T3	0	0	0	2
T2	1	0	2	0

Minus	S1	S2	S3	S4
T1	3	0	0	0
T3	0	0	3	0
T2	0	5	0	4

Abbildung 17.6: Die Matrizen nach einer Umordnung der Zeilen

Durch Vertauschung von Spalte 2 und 3 ergibt sich:

Plus	S1	S3	S2	S4
T1	2	0	7	0
T3	1	0	0	2
T2	0	2	0	0

Minus	S1	S3	S2	S4
T1	3	0	0	0
T3	0	3	0	0
T2	0	0	5	4

Abbildung 17.7: Die Matrizen nach einer Umordnung der Spalten

Jetzt kann man die Gruppierungen z. B. so deutlich machen:

Plus		S1		S2	
		S1	S3	S2	S4
T1	T1	2	0	7	0
	T3	0	0	0	2
T2	T2	1	2	0	0

Minus		S1		S2	
		S1	S3	S2	S4
T1	T1	3	0	0	0
	T3	0	3	0	0
T2	T2	0	0	5	4

Abbildung 17.8: Zeilen bzw. Spalten werden zu Gruppen zusammengefasst

Das farbige Netz hat zwei Transitionen **T1** und **T2** und zwei Stellen **S1** und **S2**. Die Transition **T1** hat zwei Farben (T1 und T3), **T2** hat eine Farbe (T2). Die Stellen **S1** und **S2** haben je zwei Farben (S1 und S3 bzw. S2 und S4).

Graphisch kann man dieses farbige Netz z. B. auch so darstellen:

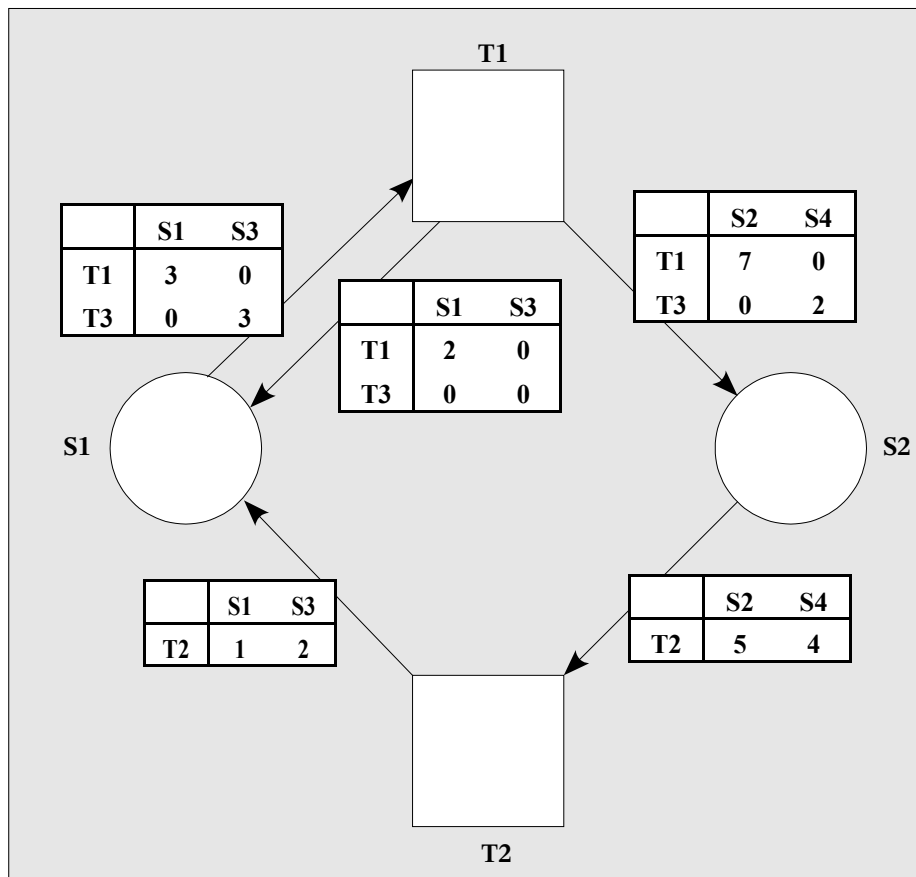


Abbildung 17.9: Das farbige Netz in graphischer Darstellung

Durch die Gruppierung der S/T-Stellen und der S/T-Transitionen zu farbigen Stellen und farbigen Transitionen sind die beiden Matrizen (Plus und Minus) in Teilmatrizen eingeteilt. Jeder Teilmatrix der Plus-Matrix entspricht im farbigen Netz ein Pfeil von einer Transition zu einer Stelle. Jeder Teilmatrix der Minus-Matrix entspricht im farbigen Netz ein Pfeil von einer Stelle zu einer Transition. Enthält eine Teilmatrix nur Nullen, so wird in der Graphik kein entsprechender Pfeil eingezeichnet.

Es bietet sich an, jeden Pfeil im farbigen Netz mit **seiner** Teilmatrix zu beschriften (so wie in Abb. 17.9).

Dann kann man (in der Abb. 17.9) z. B. die Teilmatrix am Pfeil von der farbigen Stelle **S1** zu der farbigen Transition **T1** so lesen:

"Wenn die Transition **T1** mit ihrer Farbe T1 schaltet, dann werden von der Stelle S1 drei Marken der Farbe S1 und 0 Marken der Farbe S3 entfernt. Wenn **T1** dagegen mit der Farbe T3 schaltet, werden null Marken der Farbe S1 und drei Marken der Farbe S3 von **S1** entfernt".

Der Kapazitäts-Vektor des farbigen Netzes ergibt sich auf naheliegende Art (durch "Gruppierung") aus dem Kapazitäts-Vektor des S/T-Netzes. Aus dem Vektor in Abb. 17.4 ergibt sich für das in Abb. 17.8 und 17.9 dargestellte farbige Netz:

	S1		S2	
	S1	S3	S2	S4
Kapazität	3	ω	14	12

Abbildung 17.10: Ein Kapazitäts-Vektor für Abb. 17.9

Entsprechend diesem Vektor hat die farbige Stelle **S1** Kapazität für drei Marken der Farbe S1 und für unbegrenzt viele Marken der Farbe S3. Auf **S2** können maximal 14 Marken der Farbe S2 und 12 Marken der Farbe S4 liegen.

Formal gesehen unterscheiden sich S/T-Netze und farbige Netze kaum voneinander. In praktischen Anwendungen kann die zusätzliche Struktur eines farbigen Netzes trotzdem sehr wichtig sein für die übersichtliche und verständliche Gliederung eines Systemmodells. Insofern sind S/T-Netze und farbige Netze sehr verschieden voneinander.

Geht man von einem **B/E-Netz** aus (statt von einem S/T-Netz) und gruppiert man die Bedingungen und Ereignisse wie oben (für Stellen und Transitionen) erläutert, so erhält man ein **striktes farbiges Netz**. Ein solches Netz kann man ganz einfach auch daran erkennen, dass in seinen Matrizen (bzw. in den Teilmatrizen an den Pfeilen in einer graphischen Darstellung) nur Nullen und Einsen vorkommen statt beliebige natürliche Zahlen. Außerdem gehört zu einem strikten farbigen Netz immer ein Kapazitäts-Vektor, der **nur Einsen** enthält (weshalb er meistens nicht angegeben wird).

Das B/E-Netz in Abb. 17.2 kann man z. B. zu folgendem strikten farbigen Netz gruppieren:

Plus		S1		S2	
		B1	B3	B2	B4
T1	E1	1	0	1	0
	E3	0	0	0	1
T2	E2	1	1	0	0

Minus		S1		S2	
		B1	B3	B2	B4
T1	E1	1	0	0	0
	E3	0	1	0	0
T2	E2	0	0	1	1

Abbildung 17.11: Ein striktes farbiges Netz zu Abb. 17.2

Dieses strikte farbige Netz kann man graphisch z. B. so darstellen:

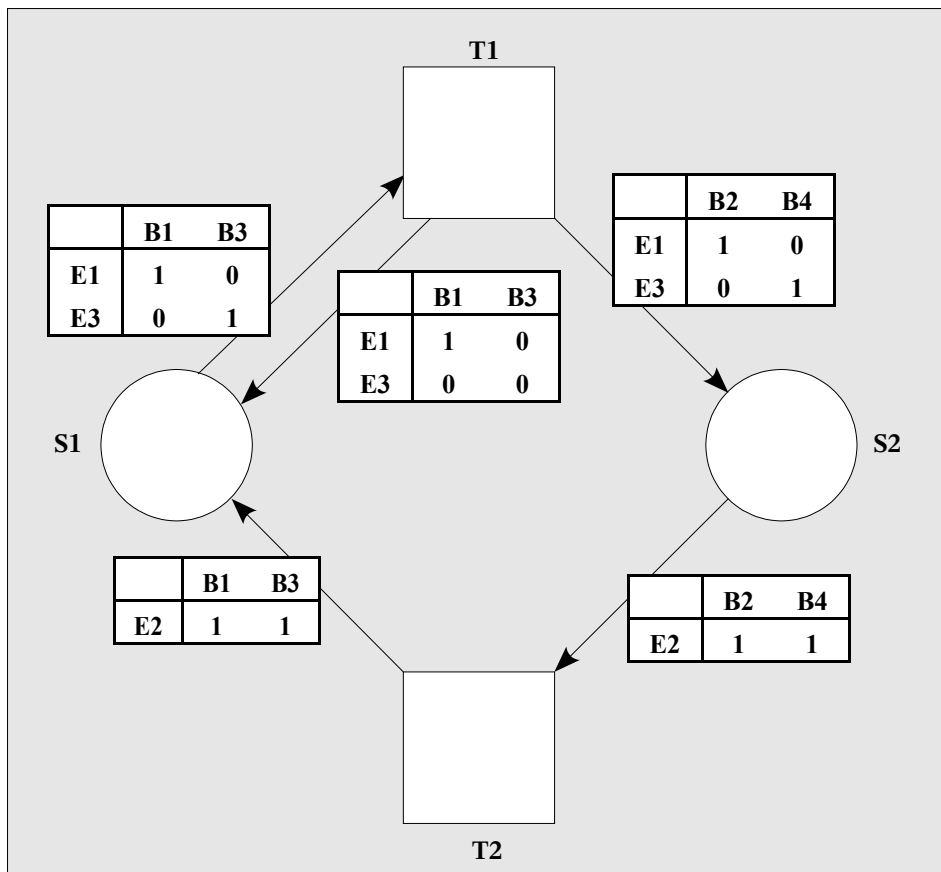


Abbildung 17.12: Graphische Darstellung des strikten farbigen Netzes in Abb. 17.11

Vergleicht man die graphische Darstellung des B/E-Netzes in Abb. 17.2 mit der graphischen Darstellung eines entsprechenden strikten farbigen Netzes in Abb. 17.11 (oder das S/T-Netz in Abb. 17.3 mit dem farbigen Netz in Abb. 17.8) so erkennt man: Informationen wurden aus der Netzstruktur in die Kantenbeschriftungen (Teilmatrizen) verlagert. Man kann sich den Übergang von einem S/T-Netz zu einem farbigen Netz auch so vorstellen: Das S/T-Netz wird so gefaltet, dass gewisse Stellen übereinanderliegen und gewisse Transitionen übereinanderliegen. Dadurch wird die Netzstruktur einfacher. Informationen über "die einzelnen Schichten" der Faltung gibt man getrennt von der graphischen Netzstruktur an, z. B. in Form von Teilmatrizen an den Kanten. Besonders günstig ist es, wenn man gleichförmige oder zumindest sehr ähnliche Teile des S/T-Netzes übereinanderfalten kann. Z. B. enthält das S/T-Netz in Abb. 17.3 zwei Zyklen (S1, T1, S2, T2 und S3, T3, S4, T2). Beim Übergang zum farbigen Netz in Abb. 17.9 wurden diese beiden Zyklen übereinandergefaltet. Allerdings sind die beiden Zyklen nicht völlig gleichförmig, sondern nur ähnlich, da der erste Zyklus eine Schlinge enthält (S1, T1) und der zweite Zyklus nichts Entsprechendes besitzt. Trotzdem lässt sich die Faltung formal durchführen. Ob sie auch inhaltlich sinnvoll ist, hängt von der inhaltlichen Bedeutung des Netzes ab.

Für die Verlagerung von Information aus der Netzstruktur heraus in irgendwelche Beschriftungen hinein (z. B. in die Teilmatrizen an den Kanten) gibt es nur eine formal-harte Grenze: Wenn man **alle** Ereignisse eines B/E-Netzes (bzw. alle Transitionen eines S/T-Netzes) zu **einer** farbigen Transition und **alle** Bedingungen des B/E-Netzes (bzw. alle Stellen des S/T-Netzes) zu **einer** farbigen Stelle zusammenfasst. Im Allgemeinen erhält man dadurch eine Schlinge. An einem Pfeil der Schlinge steht die Plus-Matrix, am anderen Pfeil die Minus-Matrix.

Für das B/E-Netz in Abb. 17.2 sieht dieses extreme strikte farbige Netz z. B. so aus:

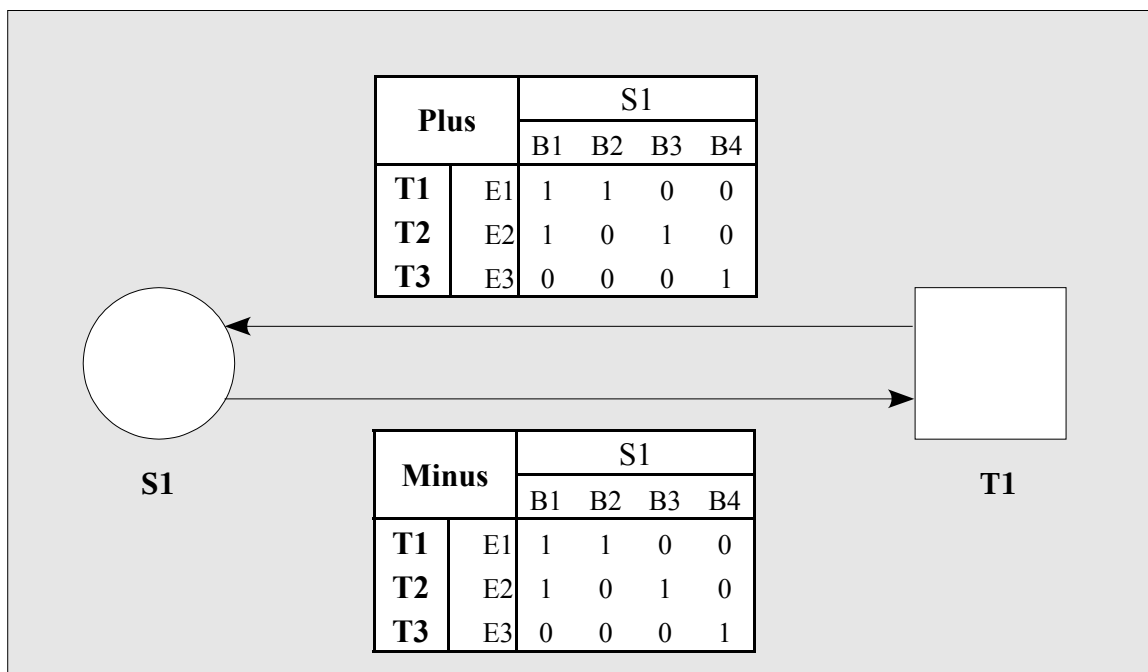


Abbildung 17.13: Das extreme farbige Netz zu dem B/E-Netz in Abb. 17.2

Um einem Mißverständnis vorzubeugen: Mit dem Begriff des farbigen Netzes sollte man nicht eine bestimmte Darstellung solcher Netze verbinden. Statt ein farbiges Netz durch zwei Matrizen oder einen Graphen mit Teilmatrizen an den Pfeilen zu beschreiben, könnte man es auch durch Angabe einiger Mengen und einiger Relationen, oder durch einen Graphen mit bestimmten Computerprogrammen an den Pfeilen oder auf eine von vielen weiteren Arten und Weisen darstellen. Der Begriff des farbigen Netzes sollte "rein semantisch" verstanden werden und unabhängig von irgendeiner Syntax zur Darstellung von Netzen.

Dies erscheint vor allem deswegen erwähnenswert, weil es sich bei den Prädikat-Transitions-Netzen (Pr/T-Netzen) etwas anders verhält. Siehe dazu den folgenden Abschnitt.

18 Farbige Netze und Pr/T-Netze: wie verhalten sich diese Netzarten zueinander?

Prädikat-Transitions-Netze (Pr/T-Netze) wurden vor den farbigen Netzen eingeführt (siehe [3] und [5]). Trotzdem (oder gerade deshalb) kann man den Pr/T-Formalismus besonders leicht als eine **Darstellungsform** für farbige Netze verstehen. Stark vereinfacht gesagt ist ein Pr/T-Netz ein farbiges Netz, welches mit Hilfe bestimmter logischer Formeln dargestellt wurde. Ändert man diese Darstellungsform erheblich (indem man z. B. die logischen Formeln durch Matrizen ersetzt, die das gleiche ausdrücken) dann ist das Netz kein Pr/T-Netz mehr, es bleibt aber ein farbiges Netz (siehe dazu auch die Anmerkung am Schluss des vorigen Abschnitts 17.).

Dazu ein Vergleich aus der Mathematik: Natürliche Zahlen kann man sehr verschieden darstellen, etwa als römische Zahlen oder positional (z. B. dezimal). Möglicherweise sind gewisse Einsichten in das Wesen der natürlichen Zahlen erst auf der Grundlage der positionalen Darstellung möglich geworden. Sicherlich ist der effiziente Umgang mit natürlichen Zahlen abhängig von einer "guten Darstellung" (die positionale Darstellung scheint in dieser Hinsicht viel besser zu sein als die römische).

In diesem Vergleich sollen die farbigen Netze den (abstrakten, darstellungsunabhängigen) natürlichen Zahlen, die Pr/T-Netze den Dezimal-Zahlen entsprechen. Indem man (farbige) Netze mit Hilfe von logischen Formeln beschreibt, versucht man, das tiefe und umfassende Wissen über formale Logik für die Netztheorie nutzbar zu machen bzw. Erkenntnisse der Logik und der Netztheorie zu vereinen. Die Klasse der Pr/T-Netze ist nicht nur semantisch, sondern auch (mehr oder weniger) syntaktisch bestimmt. Die Klasse der farbigen Netze ist nur semantisch bestimmt. Eine Syntax für die Darstellung farbiger Netze kann (bzw. muss) vom Anwender gewählt werden.

Häufig taucht die Frage auf, ob man jedes farbiges Netz auch als Pr/T-Netz darstellen kann. Von der folgenden Antwort hoffe ich, dass sie die Zustimmung von Hartmann Genrich, Kurt Lautenbach und Kurt Jensen finden kann: Es gibt einen Kern des Pr/T-Formalismus, mit dem man nicht alle farbigen Netze darstellen kann. Das Problem liegt darin, dass in einem "Kern-Pr/T-Netz" über einen Pfeil P zwar mehrere Marken fließen können, aber bei jedem Schaltvorgang (an dem P beteiligt ist) die gleiche Anzahl von Marken fließt. Jeder Pfeil P hat also ein festes Gewicht. Dagegen können in einem farbigen Netz bei einem Schaltvorgang über einen Pfeil P z. B. drei Marken der Farbe S7 und bei einem anderen Schaltvorgang über den selben Pfeil P fünf Marken der Farbe S7 fließen.

Möglicherweise gibt es Zusatzmechanismen, mit denen man den Kern des Pr/T-Formalismus so erweitern kann, dass sich jedes farbiges Netz auch als Pr/T-Netz darstellen lässt. Diese Zusatzmechanismen sind aber, soweit mir bekannt, noch nicht in allen Einzelheiten ausgearbeitet und tiefergehend untersucht worden. Für die meisten praktischen Anwendungen kann man sich auf die (farbigen) Netze beschränken, die man auch als Pr/T-Netze darstellen kann.

Anhand eines Beispiels soll der Zusammenhang zwischen Pr/T-Netzen und farbigen Netzen etwas näher erläutert werden.

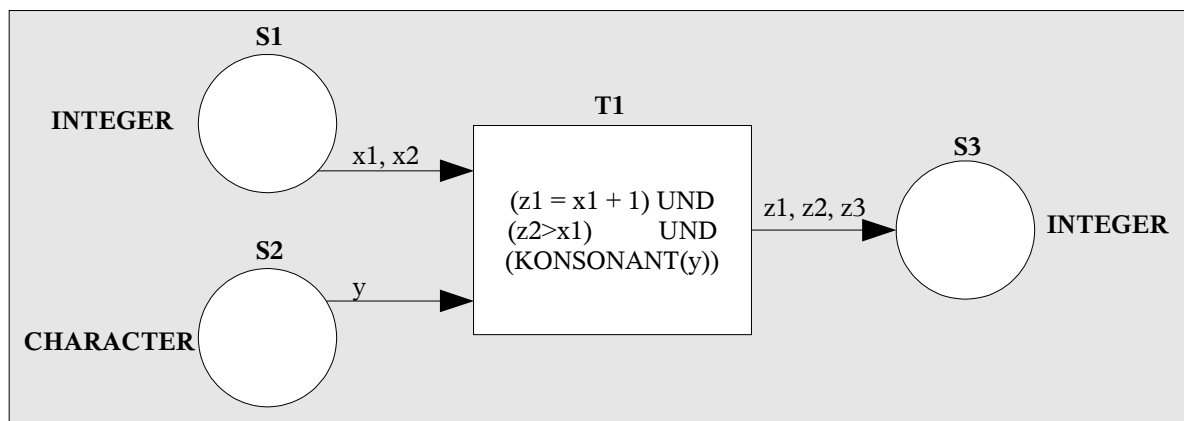


Abbildung 18.1: Ein Prädikat-Transitions-Netz (Pr/T-Netz)

In einem (Kern-) Pr/T-Netz muss jeder Pfeil mit einem oder mehreren Variablen-Namen beschriftet sein. Die Variablen an den zu einer Transition T hinführenden Pfeilen heißen auch Eingabe-Variablen von T. Die Variablen an den von einer Transition wegführenden Pfeilen heißen auch Ausgabe-Variablen von T. Jede Transition T muss mit einer logischen Formel ("einem booleschen Ausdruck") beschriftet sein. In dieser Formel dürfen die Namen der Eingabe-Variablen und der Ausgabe-Variablen von T vorkommen.

Im Beispiel in Abb. 18.1 hat die Transition T drei Eingabe-Variablen (x_1 , x_2 und y) und drei Ausgabe-Variablen (z_1 , z_2 und z_3). In der Formel von T kommen x_1 , z_1 , z_2 und y vor, x_2 und z_3 kommen nicht darin vor.

In praktischen Anwendungen ordnet man jeder Stelle S in einem Pr/T-Netz einen Datentyp zu. Auf S dürfen dann nur Werte dieses Typs als Marken liegen. Alle Variablen in der Umgebung von S (d.h. an Pfeilen, die von S weg- bzw. zu S hinführen) "erben" den Typ von S.

Im Beispiel in Abb. 18.1 sind x_1 und x_2 vom Typ INTEGER, weil S1 den Typ INTEGER hat. Die Variable y erbt den Typ CHARACTER von der Stelle S2 und z_1 , z_2 , und z_3 sind entsprechend vom Typ INTEGER.

Eine Variablen-Belegung für eine Transition T ordnet jeder Eingabe-Variablen und jeder Ausgabe-Variablen von T einen Wert (des entsprechenden Typs) zu.

Die folgenden drei Variablen-Belegungen sind Variablen-Belegungen für die Transition T im Beispiel in Abb. 18.1:

$$VB1 = (x_1: +3, x_2: +17, y: M, \quad z_1: +4, z_2: +4, z_3: +37)$$

$$VB2 = (x_1: +4, x_2: +17, y: A, \quad z_1: +5, z_2: +7, z_3: +37)$$

$$VB3 = (x_1: +4, x_2: +17, y: M, \quad z_1: +5, z_2: +7, z_3: +37)$$

Eine Transition T kann mit einer Variablen-Belegung VB schalten, wenn folgende beiden Bedingungen erfüllt sind:

1. Die Werte, die VB den Eingabe-Variablen von T zuordnet, müssen als Marken auf den entsprechenden Vorgängerstellen von T liegen.
2. Die Formel von T muss durch die Variablen-Belegung VB erfüllt sein.

Angenommen, das Pr/T-Netz in Abb. 18.1 ist folgendermaßen markiert:

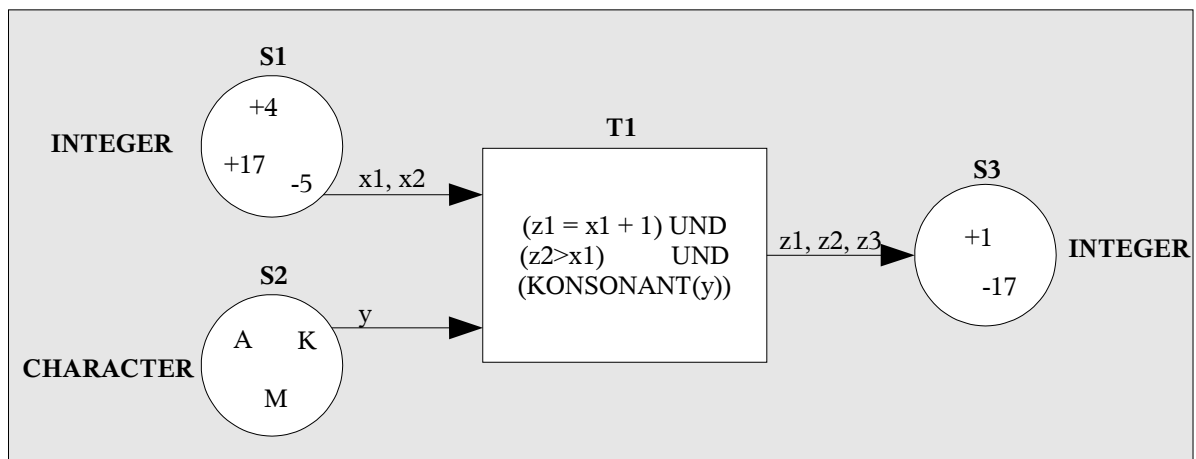


Abbildung 18.2: Das Pr/T-Netz aus Abb. 18.1 mit einer Markierung

Dann kann T mit der Variablen-Belegung

$$VB1 = (x1: +3, x2: +17, y: M, z1: +4, z2: +4, z3: +37)$$

nicht schalten, weil der Wert +3 für die Eingabe-Variable $x1$ nicht auf der Stelle S1 liegt. Mit $VB2 = (x1: +4, x2: +17, y: A, z1: +5, z2: +7, z3: +37)$ kann T nicht schalten, weil die Formel von T nicht erfüllt ist (Der Buchstabe A ist kein Konsonant). Mit $VB3 = (x1: +4, x2: +17, y: M, z1: +5, z2: +7, z3: +37)$ kann T schalten.

Man beachte: In diesem Beispiel ist der Wert der Ausgabe-Variablen $z1$ eindeutig durch die Werte der Eingabe-Variablen bestimmt (durch den Teil " $z1 = x1 + 1$ " der Formel von T).

Der Wert von $z2$ ist dagegen nicht eindeutig durch die Werte der Eingabe-Variablen bestimmt. Er muss aber so gewählt (oder geraten) werden, dass $(z2 > x1)$ gilt.

Der Wert für die Ausgabe-Variable $z3$ kann (unter allen Werten vom Typ INTEGER) frei gewählt werden, da $z3$ in der Formel von T nicht vorkommt.

Die Eingabe-Variable $x2$ kommt auch nicht in der Formel von T vor. Somit kann sie mit einem beliebigen Wert belegt werden, der als Marke auf der Stelle S1 liegt.

Jedes Pr/T-Netz stellt ein farbiges Netz dar. Die Pr/T-Terminologie ist dabei folgendermaßen in die farbige Terminologie zu übersetzen:

Jede **Stelle** (des Pr/T -Netzes) hat (als Stelle des entsprechenden farbigen Netzes betrachtet) so viele **Farben**, wie der Datentyp der Pr/T-Stelle Werte hat.

Im Beispiel in Abb. 18.1 hat die Stelle S2 den Datentyp CHARACTER. Angenommen, dieser Datentyp umfasst 80 verschiedene Werte (A, B.....Z, a, b, ..., z, 0, 1, ..., 9, ?, !, =, %, \$, ...). Dann hat die Stelle S2 (als Stelle eines farbigen Netzes betrachtet) 80 verschiedene Farben (die man z. B. mit A, B, ..., Z, a, b.....z, 0, 1, ..., 9, ?, !, =, %, \$, ... bezeichnen könnte).

Die Stelle S 1 im Beispiel hat den Datentyp INTEGER. In theoretischen Zusammenhängen umfasst dieser Datentyp abzählbar-unendlich viele Werte (0, +1, -1, +2, -2, +3, ...). Also hat die Stelle S1 als farbige Stelle abzählbar-unendlich viele Farben (die man z. B. mit 0, +1, -1, +2, -2, +3, bezeichnen könnte).

Die **Farben** einer **Transition** T (des Pr/T-Netzes) erhält man so: Man nimmt die Menge aller Variablen-Belegungen für T und entfernt daraus alle Belegungen, die die Formel von T nicht erfüllen. Jede übrigbleibende, die Formel erfüllende Variablen-Belegung, entspricht einer Farbe von T.

Im Beispiel in Abb. 18.1 hat die Transition T sechs Variablen: x_1 , x_2 , y , z_1 , z_2 , z_3 . Die Menge aller Variablen-Belegungen für T erhält man, indem man das kartesische Produkt der Wertebereiche dieser Variablen bildet:

INTEGER x INTEGER x CHARACTER x INTEGER x INTEGER x INTEGER

Von diesen Variablen-Belegungen erfüllen abzählbar-unendlich viele die Formel $(z_1 = x_1 + 1)$ UND $(z_2 > x_1)$ UND KONSONANT(y) und abzählbar-unendlich viele Belegungen erfüllen diese Formel nicht. Jede erfüllende Belegung, z. B. VB3 = (x_1 : +4, x_2 : +17, y : M, z_1 : +5, z_2 : +7, z_3 : +37), entspricht einer Farbe für T.

Mit diesen "Übersetzungen" zwischen dem Pr/T-Formalismus und dem farbigen Formalismus sind die folgenden beiden Sätze gleichbedeutend:

"Die Transition T (eines Pr/T-Netzes) schaltet mit der Variablen-Belegung VB" und

"Die Transition T (des entsprechenden farbigen Netzes) schaltet mit der Farbe VB".

19 Ist jedes Netz ein Graph oder jeder Graph ein Netz?

Ein gerichteter Graph besteht aus einer Menge K von Knoten, einer Menge P von Kanten (Pfeilen) und zwei Abbildungen $VON: P \rightarrow K$ und $NACH: P \rightarrow K$, die jeder Kante einen Ausgangsknoten und einen Zielknoten zuordnen.

"Zwei Knoten *akzidieren*" heißt: Die beiden Knoten sind durch eine Kante (in der einen und/oder der anderen Richtung) miteinander verbunden. "Ein Knoten **inzidiert** mit einer Kante" heißt: Der Knoten ist Ausgangs- und/oder Zielknoten der Kante.

Man kann gerichtete Graphen nicht nur graphisch darstellen, sondern auch durch ihre sog. *Akzidenz-Matrix*, oder aber durch ihre beiden sogenannten *Inzidenz-Matrizen*.

Die Akzidenz-Matrix eines gerichteten Graphen enthält für jeden Knoten eine Zeile und eine Spalte und als Einträge Nullen und Einsen. Eine 1 an der Stelle a_{ij} der Matrix bedeutet: Vom Knoten K_i führt eine Kante zum Knoten K_j . Eine 0 bedeutet die Abwesenheit einer solchen Kante.

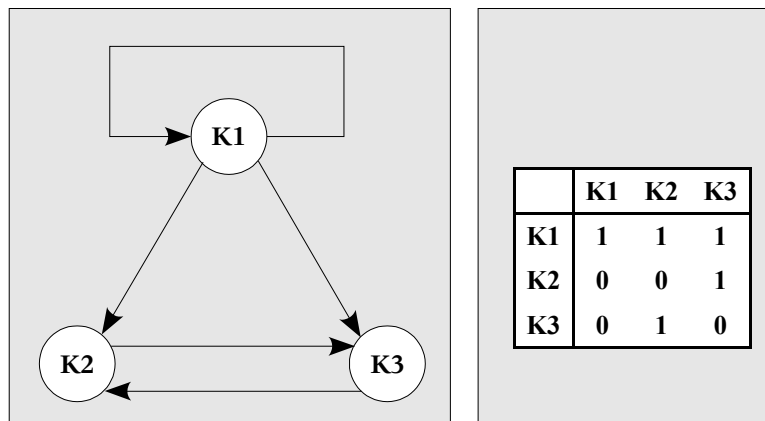


Abbildung 19.1: Ein Graph und seine Akzidenz-Matrix

Die beiden *Inzidenz-Matrizen* eines gerichteten Graphen entsprechen genau den Abbildungen VON und $NACH$ (und werden im Folgenden auch so bezeichnet). Sie enthalten für jeden Knoten K_i eine Zeile, für jede Kante P_j eine Spalte und als Einträge Nullen und Einsen. Eine 1 an der Stelle a_{ij} in der VON -Matrix (bzw. in der $NACH$ -Matrix) bedeutet, dass die Kante P_j beim Knoten K_i beginnt (bzw. endet). Steht sowohl in der VON - als auch in der $NACH$ -Matrix eine 0 an der Stelle a_{ij} , dann inzidieren die Kante P_j und der Knoten K_i nicht.

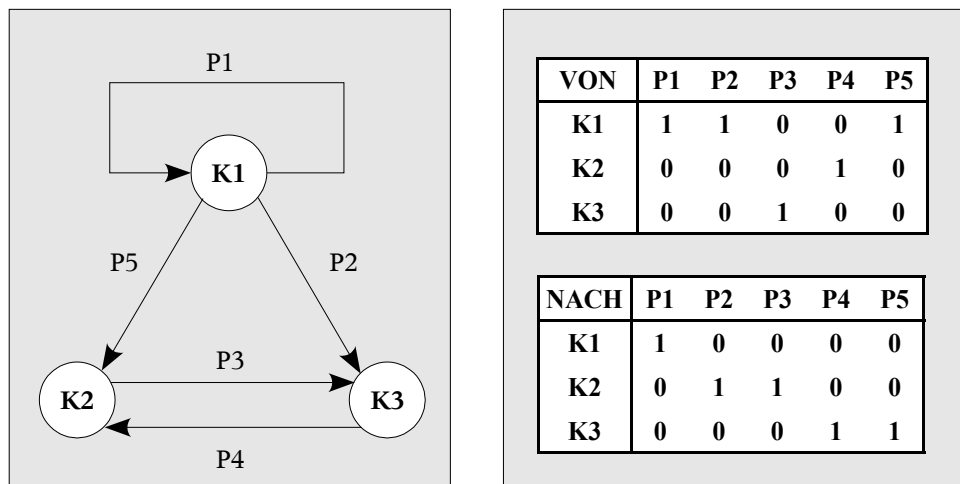


Abbildung 19.2: Ein Graph und seine Inzidenz-Matrizen VON und NACH

Petri-Netze werden häufig als *spezielle gerichtete Graphen* bezeichnet und verstanden, nämlich als Graphen mit zwei Arten von Knoten (bichromatische Graphen) und der Einschränkung, dass Kanten nur zwischen verschiedenartigen Knoten verlaufen dürfen. Bei dieser Auffassung ist offensichtlich jedes Petri-Netz ein gerichteter Graph, aber nicht jeder gerichtete Graph ein Petri-Netz.

Man kann Petri-Netze aber auch als eine *Verallgemeinerung* von gerichteten Graphen verstehen, so dass jeder gerichtete Graph ein Petri-Netz, aber nicht jedes Netz ein gerichteter Graph ist. Die Verallgemeinerung betrifft in erster Linie den Begriff einer Kante.

In einem gerichteten Graph hat jede Kante genau *einen* Ausgangsknoten und *einen* Zielknoten. Entsprechend muss in jeder Spalte der beiden Inzidenz-Matrizen genau eine 1 stehen.

Nun kann man eine "verallgemeinerte Kante" definieren als ein Objekt mit einem oder *mehreren* Ausgangsknoten und einem oder *mehreren* Zielknoten.

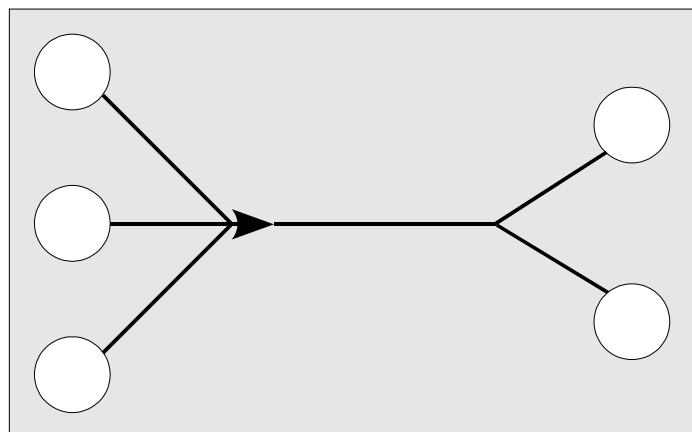


Abbildung 19.3: Eine verallgemeinerte Kante

Ein verallgemeinerter Graph besteht dann aus Knoten und verallgemeinerten Kanten. In den Inzidenz-Matrizen eines verallgemeinerten Graphen können in jeder Spalte mehrere Einsen stehen.

Statt in einem B/E-Netz sowohl die Bedingungen als auch die Ereignisse als *Knoten* (eines gerichteten Graphen) aufzufassen, kann man auch z. B. *die Ereignisse als Knoten* und die *Bedingungen* (zusammen mit den daran hängenden Pfeilen) als *verallgemeinerte Kanten* auffassen.

Im Abschnitt 17. wurde gezeigt, wie man B/E-Netze durch zwei Matrizen (Plus und Minus) darstellen kann. Fasst man ein B/E-Netz als einen verallgemeinerten Graphen auf so, entsprechen Plus und Minus den Inzidenz-Matrizen NACH und VON.

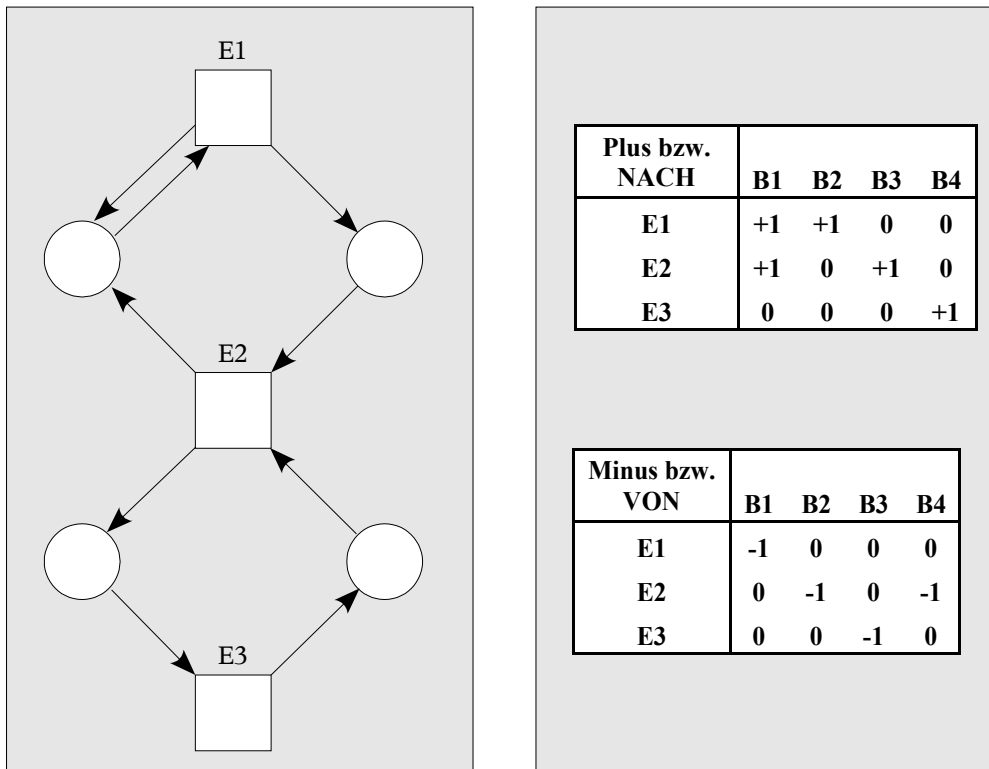


Abbildung 19.4: Ein B/E-Netz und seine beiden Inzidenz-Matrizen

In einem Prozessnetz darf eine (unwiederholbare) Bedingung nur ein Vor- und ein Nach-Ereignis haben. Fasst man ein Prozessnetz als einen verallgemeinerten Graphen auf, dann sind dessen verallgemeinerte Kanten (d.h. die Bedingungen des Prozessnetzes) gerade von dem speziellen Typ, der nur einen Ausgangsknoten und nur einen Zielknoten hat, d.h. "normale" Kanten.

20 Theorie und Anwendungen von Petri-Netzen

Die Theorie der Petri-Netze ist ihrer Herkunft und dem mit ihr verbundenen Anspruch nach eine Grundlagen-Theorie. Ihr Ziel ist es, die fundamentalen Phänomene der Nebenläufigkeit und der Veränderung inhaltlich und formal zu durchdringen.

Formale Theorien müssen von möglichst wenigen und "einfachen" Grundbegriffen ausgehen. Deshalb sind auch Netz-Theoretiker bestrebt, die grundlegenden Definitionen "einfach" und ihre Anzahl gering zu halten. Nur so können sie hoffen, das Wesentliche der Phänomene Nebenläufigkeit und Veränderung herauszuarbeiten oder in einem tieferen Sinne zu verstehen, was Bedingungen und Ereignisse (sozusagen die kleinsten Einheiten der Nebenläufigkeit und der Veränderung) sind, etc.

Wendet man Petri-Netze an, z. B. um damit Systeme zu modellieren, so hat man ganz gegenteilige Interessen. Petri-Netze sind dann Mechanismen, die man entwirft und zusammensetzt, um ein bestimmtes System zu beschreiben. Je mehr die Mechanismen "von sich aus können", desto leichter ist es, damit Systeme zu modellieren. Anwender neigen deshalb fast immer zu Netz-Klassen mit vielen "Zusatz-Einrichtungen", z. B. mit Nebenbedingungen. Sperrpfeilen (inhibitor arcs), Löschpfeilen (flushing arcs), Zeitbeschriftungen, modifizierten oder modifizierbaren Schaltregeln etc.

Ohne solche Zusatz-Einrichtungen ist das Modellieren mit Petri-Netzen manchmal nicht ganz einfach. Mit solchen Zusätzen ist das Modellieren einfacher, aber die Theorie kann dann entsprechend weniger oder nichts über die sich ergebenden Netze aussagen. Dies scheint aber, nach allem was wir heute wissen, kein Charakteristikum des Petri-Netz-Formalismus, sondern ein Zeichen dafür zu sein, dass das Modellieren und Analysieren von nebenläufigen Systemen ein noch sehr neues Gebiet und im Allgemeinen schwierig ist.

Ein erster Schritt aus dem Dilemma zwischen theoretischer Reinheit und Kargheit und praktischer Vielfalt und Unübersichtlichkeit besteht darin, Zusatz-Einrichtungen als Abkürzungen aufzufassen und ihre Bedeutung durch Expansionen (in Netze ohne Zusatz-Einrichtungen) anzugeben.

Dieser Schritt allein genügt aber nicht. Denn häufig sind die expandierten Netze so groß, dass sie kaum noch zu handhaben (auszuführen bzw. zu analysieren) sind. Besonders wünschenswert sind deshalb Methoden, mit denen man ein Netz mit Zusatz-Einrichtungen direkt handhaben (ausführen, analysieren) kann, die dabei aber die gleichen Ergebnisse liefern, als hätte man das Netz zuerst expandiert und dann ausgeführt bzw. analysiert.

Die Theorie der Petri-Netze ist eine Art Atom-Theorie der Nebenläufigkeit. Sie klärt Grundbegriffe, postuliert kleinste Einheiten der Nebenläufigkeit (Bedingungen und Ereignisse), liefert mathematische Modelle (z. B. Halbordnungen von Bedingungen und Ereignissen) und Theoreme über diese Modelle.

In der Praxis arbeitet man aber selten auf der Ebene der Atome. Meistens hat man es mit großen und komplexen Agregaten von Atomen zu tun. Ein wichtiges Ziel ist es, die Eigenschaften der Aggregate aus den Eigenschaften und der Anordnung der Atome herzuleiten. Entsprechend ist es notwendig, auf dem Fundament der Netztheorie ingenieurmäßige Methoden für ihre praktischen Anwendungen zu entwickeln. Diese Entwicklung hat schon begonnen, wird aber sicherlich noch geraume Zeit in Anspruch nehmen.

Netz-Theoretiker werden auch in Zukunft auf wenigen und einfachen Grundkonzepten bestehen. Aber vielleicht können sie den Anwendern verstärkt entgegenkommen, indem sie auch einige Zusatz-Einrichtungen gründlich theoretisch durchdringen.

Netz-Anwender werden vermutlich auch in Zukunft mit jedem neuen Anwendungsgebiet weitere Zusatz-Einrichtungen verlangen. Vielleicht werden sie aber auch ein tieferes Verständnis für die Ziele und Probleme der Netztheorie entwickeln.

21 Literatur-Angaben

- [1] S. Drees, D. Gomm, H. Plünnecke, W. Reisig, R. Walter
"Bibliography of Net-Theory"
Arbeitspapiere der GMD 315, 1988

In dieser Bibliographie sind 2634 Artikel und Bücher zur Netztheorie verzeichnet. Insbesondere findet sich darin eine Liste aller Arbeiten von Prof. Petri.

- [2] W. Reisig
"Petri-Netze - Eine Einführung"
Springer Verlag 1986 (Erstauflage 1982)

Eine sehr lesbare und lesenswerte Einführung. Es werden nur solche Konzepte behandelt, von denen etwas Wesentliches bewiesen werden kann.

- [3] H. J. Genrich, K. Lautenbach
"The Analysis of Distributed Systems by means of Predicate/Transition-Nets"
LNCS Nr. 70, Springer Verlag 1979

Ein früher Artikel über Prädikat/Transitions-Netze. Häufiger zitiert wird die folgende Arbeit:

- [4] H. J. Genrich, K. Lautenbach
"System Modelling with High-Level Petri Nets"
Theoretical Computer Science Nr. 13, North Holland Company 1981

Dieser Artikel motiviert und erläutert Prädikat/Transitions-Netze. Die formale Definition von Pr/T-Netzen in dieser Arbeit erfüllt, was Genauigkeit und Vollständigkeit angeht, nicht alle meine Wünsche.

- [5] K. Jensen
"How to Find Invariants for Coloured Petri Nets"
LNCS Nr. 118, Springer Verlag 1981

Eine Einführung des Begriffs "farbiges Netz". Gut motiviert, präzise und verständlich.

- [6] U. Grude, C. Heuser
"Towards Good Relations Among Net Classes"
Petri Net Newsletter Nr. 24, Gesellschaft für Informatik 1986

Ein Versuch, die vier wichtigsten Netz-Klassen (B/E-Netze, S~T-Netze, strikte farbige Netze und farbige Netze) durch eine einzige, generische Definition zu beschreiben.