Stand: Februar 1998 (1)

Siegfried Wendt
Universität Kaiserslautern

Die Modelle von Moore und Mealy – Klärung einer begrifflichen Konfusion

1. Grobe Charakterisierung des Problems

In der Philosophie ist es selbstverständlich, daß Autoren, die Erkenntnisse früherer Philosophen weitergeben oder kommentieren, die Originalliteratur kennen und sich in ihrer Argumentation explizit auf bestimmte Stellen in den Originaldarstellungen beziehen. In der Technik dagegen ist es allgemein akzeptierte Praxis, daß Autoren von Lehrbüchern, in denen Erkenntnisse früherer Forscher dargestellt oder kommentiert werden, nicht die Originaldarstellungen zugrunde legen, sondern sich mit den Darstellungen in der Sekundärliteratur begnügen. Man denke an die Erkenntnisse von Boole oder Maxwell, die in sehr vielen Lehrbüchern der Digitaltechnik bzw. der theoretischen Elektrotechnik vermittelt werden, ohne daß die Autoren dieser Lehrbücher auf die Originalschriften von Boole oder Maxwell Bezug nehmen. Dagegen wird man wohl kaum ein Buch über Erkenntnisse von Aristoteles oder Kant finden, dessen Autor sich nicht explizit auf bestimmte Stellen in den Schriften dieser Philosophen bezieht.

Es sind inzwischen 40 Jahre vergangen, seit die beiden Autoren Moore und Mealy ihre Aufsätze zu bestimmten Problemen aus dem Bereich der sequentiellen Maschinen veröffentlicht haben. Es gibt wohl kein Buch über Automatentheorie und nur sehr wenige Bücher über Schaltwerkstheorie, in denen die Namen Moore und Mealy nicht genannt werden. Wenn man die Darstellungen in diesen Büchern mit den Originalarbeiten von Moore und Mealy vergleicht, dann muß man vermuten, daß die meisten Autoren der Sekundärliteratur die Originalarbeiten von Moore und Mealy nicht kennen. Mit den Namen Moore und Mealy werden nämlich häufig Erkenntnisse verbunden, die in den Arbeiten von Moore und Mealy nicht präsentiert werden. Wenn man die ursprünglichen Erkenntnisse von Moore und Mealy und die später mit diesen Namen verbundenen Erkenntnisse nicht streng auseinander hält, dann muß es zwangsläufig zu begrifflichen Konfusionen kommen. Deshalb ergeben auch die Darstellungen, die man in den unterschiedlichen Werken der Sekundärliteratur zu den Namen Moore und Mealy findet, kein einheitliches Bild.

Die begriffliche Konfusion äußert sich in dem Sachverhalt, daß man zu den folgenden Fragen in der Sekundärliteratur nicht immer die gleichen Antworten findet: Ist das Moore–Modell ein Sonderfall des Mealy–Modells? Hat Moore gesagt, daß die jetzige Ausgabe nur vom Zustand *vor* dem Übergang abhängt, oder hat er gesagt, daß sie nur vom Zustand *nach* dem Übergang abhängt, oder hat er sich nicht festgelegt? Soll man mit den Namen Moore und Mealy bestimmte Maschinentypen verbinden oder nur Formeln, die einen formalen Zusammenhang zwischen Symbolfolgen beschreiben? Sind die Modelle von Moore und Mealy einander äquivalent, und falls ja, in welchem Sinne?

Die folgende Darstellung wird diese Fragen eindeutig beantworten. Der Autor hat die Hoffnung, daß seine Ausführungen dazu beitragen können, daß in der zukünftigen Sekundärliteratur mit den Namen Moore und Mealy eine einheitliche und in sich schlüssige Begriffswelt verbunden wird.

2. Der begriffliche Kontext

Die Analyse und Interpretation der Arbeiten von Moore und Mealy fällt umso leichter, je klarer die Begriffswelt ist, in der sich der Analyst und Interpret bewegt. Deshalb soll nun vor dem Beginn der Analyse und Interpretation der begriffliche Kontext dargestellt werden, innerhalb dessen sich der Autor bewegt.

2.1 Maschinen und Modelle

Maschinen sind physikalische Gebilde. Hier interessieren nur solche Maschinen, die als *deterministische kausale Systeme* modelliert werden können (s. Abschnitt 2.2).

Zwei Maschinen sind schnittstellenäquivalent, wenn in dem System, innerhalb dessen die jeweilige Maschine eingebettet betrieben wird, die eine Maschine durch die jeweils andere ersetzt werden kann, ohne daß dies im umgebenden System festgestellt werden kann.

Für den Modellierer stellt die zu modellierende Maschine ein beobachtbares und zum Teil beeinflußbares Gebilde dar. Im ersten Schritt der Modellierung wählt der Modellierer die Orte für die Beobachtung und die Beeinflussung des Systemverhaltens. Anschließend legt er sich auf ein bestimmtes Auflösungsvermögen fest, d.h. er entscheidet sich pro Beobachtungs– bzw. Beeinflussungsort für ein Repertoire der dort potentiell beobachtbaren Sachverhalte. Das besonders interessierende Ergebnis der Modellierung besteht in der Darstellung eines formalen Zusammenhangs zwischen den an den Beobachtungs– und Beeinflussungsorten beobachtbaren Vorgängen.

Dieses Ergebnis einer Modellierung muß man sich als ein aus zwei Abschnitten bestehendes Dokument vorstellen. In dem einen Abschnitt wird das formale Modell angegeben. Für die Beschreibung des formalen Modells genügt eine formale Spezifikation in mathematischer Symbolik, d.h. es müssen die Repertoires und Funktionen formal spezifiziert werden.

Im zweiten Abschnitt des Dokuments, welches das Ergebnis der Modellierung beschreibt, muß dargestellt werden, wie die Elemente der formal spezifizierten Repertoires in Bezug auf die modellierte Maschine zu deuten sind.

Es ist durchaus möglich, mit dem formalen Modell Experimente durchzuführen, ohne zu wissen, was die formalen Elemente bedeuten sollen. Unter einem Experiment mit dem formalen Modell soll dabei der Vorgang verstanden werden, daß der Experimentator diejenigen formalen Beobachtungsergebnisse vorgibt, die seine Beeinflussung darstellen, und daß er dann hieraus über die Modellfunktionen die noch fehlenden Beobachtungsergebnisse bestimmt.

Eine Veranschaulichung des Begriffs des formalen Experiments kann nur anhand eines konkreten formalen Modells erfolgen. Deshalb erfolgt die Veranschaulichung erst im Abschnitt 2.3, nachdem ein bestimmtes formales Modell eingeführt worden ist.

Ein Modell paßt genau dann zu einer Maschine, wenn man die Elemente der formalen Beobachtungsrepertoires derart als beobachtbare Erscheinungen deuten kann, daß jedes Experiment mit dem formalen Modell zu einem entsprechenden Experiment mit der Maschine paßt.

2.2 Das Modell des deterministischen kausalen Systems

Solange man ein deterministisches kausales System als sog. *Blackbox* betrachtet, d.h. als Gebilde, in dessen Inneres man nicht hineinschaut, sind die Beobachtungsorte sog. *Schnittstellenorte* am Rande des Gebildes. Die Menge dieser Schnittstellenorte läßt sich eindeutig in zwei disjunkte Teilmengen zerlegen, nämlich in die Menge der Beeinflussungs– oder Eingabeorte einerseits und die Menge der Reaktions– oder Ausgabeorte andererseits. Die Vereinigung der Sachverhalte an allen Eingabeorten zum Zeitpunkt t wird durch den Ausdruck X(t) symbolisiert, und die Vereinigung der Sachverhalte an allen Ausgabeorten zum Zeitpunkt t wird durch den Ausdruck Y(t) symbolisiert. Man definiert die Repertoiremengen repX und repY derart, daß gilt

$$\forall\ t \geq t_0\text{: } X(t) \in \text{repX} \qquad \text{und} \qquad \forall\ t \geq t_0\text{: } Y(t) \in \text{repY}$$

Determinismus und Kausalität können nun formal durch eine einzige Funktion f erfaßt werden:

$$Y(t_1) = f\,[\text{ Verlauf "X(t)" im Intervall } t_0 \leq t \leq t_1\ ]$$

Durch die Funktion f wird eine "zeitliche Fernwirkung" ausgedrückt, denn es wird ja ausgesagt, daß die aktuelle Ausgabe zum Zeitpunkt t_1 durch den gesamten Verlauf "X(t)" der Eingabe vom Zeitpunkt t_0 der Inbetriebnahme des Systems bis zum aktuellen Zeitpunkt t_1 bestimmt wird. Durch die Einführung des Zustandsbegriffs kann man eine Ausgabefunktion formulieren, die eine "zeitliche Nahwirkung" ausdrückt, d.h. bei der die aktuelle Ausgabe zum Zeitpunkt t_1 nur vom Zustand und der Eingabe zum gleichen Zeitpunkt t_1 abhängt:

$$Y(t_1) = \omega\,[\ Z(t_1), X(t_1)\ ]$$

Der Zustand zum Zeitpunkt t_1, symbolisiert durch $Z(t_1)$, muß als Konzentration dessen angesehen werden, was sich das System über den Verlauf "X(t)" im Intervall $t_0 \leq t < t_1$ "gemerkt" hat. Deshalb muß in der Zustandsübergangsfunktion immer noch eine zeitliche Fernwirkung zum Ausdruck kommen:

$$Z(t_1+\Delta t) = \delta\,[\ Z(t_1),\ \text{Verlauf "X(t)" im Intervall } t_1 \leq t < t_1+\Delta t\ ]\ ,$$

Zum Zeitpunkt t_0 seiner Inbetriebnahme ist das System im Zustand $Z(t_0)$. Man definiert die Repertoiremenge repZ derart, daß gilt

$$\forall\ t \geq t_0\text{: } Z(t) \in \text{repZ}$$

Man kann trotz der Verwendung des Zustandsbegriffs die Blackboxvorstellung beibehalten. In diesem Fall ist der Zustand eine rein formale Größe, der kein beobachtbarer Sachverhalt in der Maschine zugeordnet wird. Die Einführung des Zustands kann in diesem Fall nur gerechtfertigt werden, falls die Formulierung der beiden Funktionen ω und δ einfacher ist als die Formulierung der Funktion f.

Man kann aber auch mit der Einführung des Zustandsbegriffs die Blackboxvorstellung verlassen und nun annehmen, daß man in das System hineinschaut. Z(t) symbolisiert in diesem Fall die Vereinigung der Sachverhalte zum Zeitpunkt t an all den Orten im Innern des Systems, auf die man schauen muß, damit man Determinismus und Kausalität vollständig erfaßt.

2.3 Zeitdiskretisierung des Modells des deterministischen kausalen Systems

Die Zeitdiskretisierung des Modells des deterministischen kausalen Systems geschieht einfach dadurch, daß man anstelle der kontinuierlichen Zeitvariablen t den Folgenindex n setzt. Anstelle des Inbetriebnahmezeitpunkts $t=t_0$ setzt man n=1, da es üblich ist, die Zählung der Elemente in einer Folge mit 1 zu beginnen. Die Modellfunktionen lauten dann

$$Y(n_1) = \quad f\ [\ \text{Verlauf "X(n)" im Intervall } 1 \leq n \leq n_1\ ]$$

$$Y(n_1) = \quad \omega\ [\ Z(n_1), X(n_1)\ ]$$

$$Z(n_1+\Delta n) = \delta\ [\ Z(n_1),\ \text{Verlauf "X(n)" im Intervall } n_1 \leq n < n_1+\Delta n\ ]\ .$$

Da der Verlauf "X(n)" eine Folge ist, kann man schreiben

$$Y(n_1) = \quad f\ [\ (\ X(1), X(2), X(3),\ \ldots\ X(n_1-1), X(n_1)\ )\ ]$$

$$Y(n_1) = \quad \omega\ [\ Z(n_1), X(n_1)\ ]$$

$$Z(n_1+\Delta n) = \delta\ [\ Z(n_1),\ (\ X(n_1), X(n_1+1), X(n_1+2),\ \ldots\ X(n_1+\Delta n-2), X(n_1+\Delta n-1)\ )\ ]$$

Da in dieser Schreibweise keine explizite Intervallformulierung mehr vorkommt und deshalb die Variable n für eine andere Verwendung freigeworden ist, kann nun anstelle von n_1 einfach n gesetzt werden. Außerdem ist es üblich, die Zustandsübergangsfunktion mit $\Delta n=1$ zu schreiben:

$$Y(n) = \quad \omega\ [\ Z(n), X(n)\ ]$$

$$Z(n+1) = \quad \delta\ [\ Z(n),\ (\ X(n)\ )\ ]$$

Obwohl hier im Argument von ω und δ die gleiche Information steht, ist doch X(n) formal nicht gleich (X(n)), denn X(n) ist ein Element, wogegen (X(n)) eine Folge ist, die nur ein einziges Element enthält. Allerdings ist dieser formale Unterschied für die Funktionsauswertung irrelevant. Deshalb schreibt man die Modellfunktionen für diskrete deterministische kausale Systeme üblicherweise in der Form

$$\boxed{\begin{aligned} Y(n) &= \quad \omega\ [\ Z(n), X(n)\ ] \\ Z(n+1) &= \quad \delta\ [\ Z(n), X(n)\ ] \end{aligned}}$$

Dieses Modell ist ein sog. *Automatenmodell*. Daß es auch Automatenmodelle anderer Form gibt, kann erst später gezeigt werden, nachdem im folgenden Abschnitt 2.4 der Begriff der *sequentiellen Maschine* vorgestellt wurde. Die Gemeinsamkeit aller Automatenmodelle besteht darin, daß es die drei Repertoiremengen repX, repY und repZ gibt und daß die beobachteten Verläufe immer als Folgen (X(1), X(2), ...), (Y(1), Y(2), ...) und (Z(1), Z(2), ...) behandelt werden. Das Wissen darum, wie die Elemente in den Repertoires physikalisch zu deuten seien, gehört nicht zum Automatenmodell.

Häufig werden die Wörter *Automatenmodell* und *Automat* als Synonyme betrachtet.

Nun wird das im Abschnitt 2.1 bereits angekündigte Beispiel zur Veranschaulichung des Begriffs eines formalen Experiments vorgestellt. Die Veranschaulichung erfolgt anhand des durch Zeitdiskretisierung des Modells des deterministischen kausalen Systems gewonnenen Automatenmodells.

Ein Experiment mit diesem formalen Modell kann man immer in Form einer Tabelle darstellen, wie sie in Bild 1 exemplarisch gezeigt ist. Der Experimentator gibt einen Anfangszustand Z(1) und eine endliche X–Folge (X(1), X(2), ... X(k)) vor und bestimmt anschließend über die Funktionen ω und δ die restlichen Glieder der Z–Folge sowie die gesamte Y–Folge. Die Inhalte der schattierten Felder der Tabelle werden also vorgegeben, und die Inhalte der nicht schattierten Felder werden anschließend bestimmt

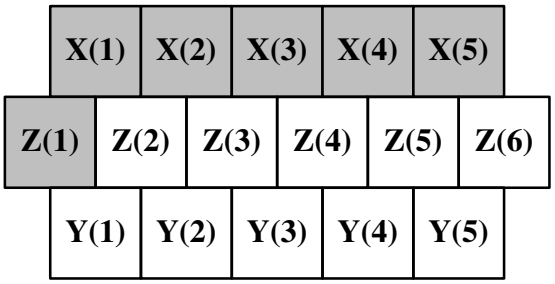


Bild 1 Exemplarische Tabelle als Darstellung eines Experiments mit einem bestimmten formalen Modell

Das Ausfüllen der Felder in der Tabelle stellt nicht notwendigerweise eine Simulation der modellierten Maschine dar, denn bezüglich der Reihenfolge, in der die Tabellenfelder gefüllt werden, bestehen bestimmte Freiheitsgrade, welche die Maschine nicht hat. Beispielsweise kann man zuerst alle noch fehlenden Elemente der Z–Folge bestimmen und sich erst dann um die Felder der Y–Folge kümmern.

2.4 Klassifikation der Ein– und der Ausgabekanäle sequentieller Maschinen

Die Zeitdiskretisierung des Modells des deterministischen kausalen Systems im Abschnitt 2.3 erfolgte rein formal ohne jeden Bezug zu einem bestimmten Maschinentyp. Nun aber werden Maschinentypen vorgestellt, für die eine Automatenmodellierung angemessen ist.

Eine Maschine wird dadurch zur *sequentiellen Maschine*, daß für sie ein Automatenmodell angegeben wird. Das bedeutet nicht, daß ein solches Modell eine Maschine sei; es bedeutet lediglich, daß man von einer Maschine nicht sagen solle, sie sei eine sequentielle Maschine, solange man für sie noch kein Automatenmodell erstellt hat.

Eine zu modellierende Maschine hat i.a. nicht nur einen einzigen Eingabekanal und einen einzigen Ausgabekanal. Man denke beispielsweise an eine Fahrkarten verkaufende Maschine. Diese hat neben dem Eingabekanal für die Münzen noch einen Eingabekanal für Geldscheine sowie mehrere Drucktasten, über die der Benutzer seine Wünsche eingeben kann. Neben dem Ausgabekanal für die Fahrkarte kann die Maschine einen separaten Ausgabekanal für das Wechselgeld sowie bestimmte Anzeigekanäle besitzen, auf denen sie dem Benutzer den aktuell noch fehlenden Geldbetrag oder ihre aktuelle Unfähigkeit, Geld zu wechseln, anzeigen kann.

In Bild 2 ist in symbolisierter Form ein Verlauf von Erscheinungen dargestellt, die beim Betrieb einer sequentiellen Maschine auftreten können. Die schraffierten Intervalle symbolisieren diejenigen Zeitabschnitte, in denen sich die Maschine nicht in einem definierten Ruhezustand befindet.

Der erste definierte Zustand Z(1) wird durch die Inbetriebnahme eingestellt. Es ist durchaus möglich, daß eine Maschine bereits anläßlich der Inbetriebnahme ein definiertes Ausgabeereignis produziert; man denke an eine Maschine, die anläßlich ihrer Inbetriebnahme akustisch meldet: "Die Inbetriebnahme ist erfolgreich verlaufen." Eine solche Ausgabe ist in Bild 2 mit $Y_T(0)$ bezeichnet. Der Index T soll die Ereignisausgaben von den Anzeigeausgaben unterscheiden; letztere sind durch den Index D gekennzeichnet. Mit dem Index T kann man das Wort *Token* verbinden, mit dem Index D das Wort *Display.*

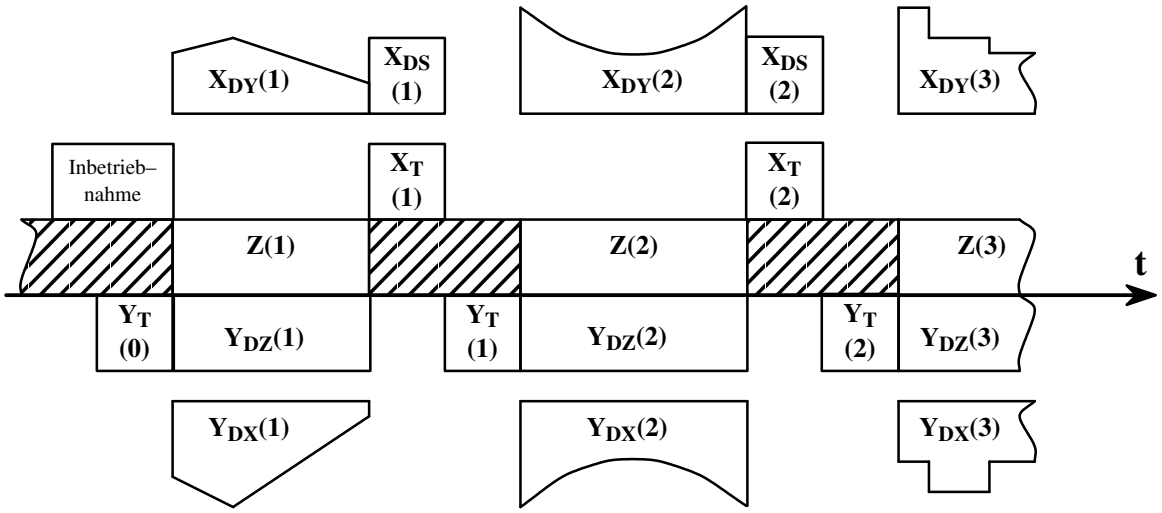


Bild 2 Die verschiedenen Arten von Ein– und Ausgabeerscheinungen beim Betrieb einer sequentiellen Maschine

Da die Zustände in Bild 2 ausschließlich Ruhezustände sein sollen, bedarf es zum Anstoß eines Zustandsübergangs eines externen Eingabeereignisses. Man denke beispielsweise an den Einwurf einer Münze oder an das Auftreten der Vorderflanke eines Taktimpulses. Die Eingabeereignisse, welche jeweils einen Zustandsübergang auslösen können, sind im Bild 2 mit X_T bezeichnet; zu dem Index T kann man wieder das Wort *Token* oder aber hier auch das Wort *Trigger* assoziieren. Die Anzeigeausgaben, die jeweils durch den aktuellen Zustand eindeutig bestimmt sind, sind in Bild 2 mit Y_{DZ} bezeichnet. Sie bleiben jeweils genau so lange konstant wie der Zustand selbst.

Der Index D kommt in Bild 2 nicht nur bei den Ausgaben Y vor, sondern auch bei den Eingaben X. Man stelle sich beispielsweise vor, zum Bedienfeld der Maschine gehöre ein ungerasteter Drehknopf, an dem der Benutzer beliebig herumdrehen kann wie beispielsweise an dem Lautstärkeregler eines Verstärkers. Durch das Drehen an diesem Knopf soll kein Zustandsübergang ausgelöst werden können. Dennoch kann es möglich sein, daß bei unverändertem Zustand der Maschine die Stellung des Drehknopfes ausgabewirksam wird. Eine solche Ausgabe muß zwangsläufig vom Anzeigetyp sein. In Bild 2 sind die Anzeigen, deren Wert unmittelbar durch den Wert von X_{DY} beeinflußt wird, mit Y_{DX} bezeichnet. Solange der Benutzer kein triggerndes Ereignis X_T bringt, kann er beliebig lange den Wert von X_{DY} ver-

ändern und auf diese Weise eine Änderung der Anzeige Y_{DX} bewirken. Wenn er dann irgendwann die triggernde Eingabe X_T bringt, kann es für das Verhalten der Maschine relevant sein, welchen Wert die Eingabe X_D aktuell hat – im Beispiel mit dem Drehknopf kann es also durchaus relevant sein, wie der Drehknopf steht, wenn eine Münze eingeworfen wird. Durch das Ereignis X_T kann also der Eingabekanal X_D abgetastet werden; der Abtastwert ist in Bild 2 mit X_{DS} bezeichnet, wobei zu dem Index S das Wort *Sample* zu assoziieren ist.

Bei einer realen Maschine muß es nicht immer alle in Bild 2 aufgeführten Erscheinungen geben, aber bei den meisten realen Maschinen gibt es Kanäle unterschiedlichen Typs, die man bei der Modellierung auseinanderhalten muß.

Zur Interpretation des Bildes 2 wurde gesagt, daß die Zustände als Ruhezustände der Maschine zu deuten seien. Nur in diesem Falle bedarf es eines externen Anstoßes für die Zustandsübergänge. Innerhalb der in Bild 2 schraffiert dargestellen sogenannten Übergangsintervalle finden im Innern der Maschine selbstverständlich kontinuierliche Vorgänge statt, von denen man sagen könnte, daß dabei die Maschine ein Kontinuum unterschiedlicher Zustände durchlaufe. Bei bestimmten Maschinenformen kann es durchaus sinnvoll sein, aus diesem Kontinuum von Zustandsübergängen endlich viele Zustände herauszugreifen und diese in das Zustandsrepertoire eines zeitdiskreten Modells aufzunehmen. Insbesondere die Behandlung der sog. asynchronen Schaltwerke in der Schaltwerkstheorie basiert auf einer solchen Betrachtungsweise. Es ist leicht einzusehen, daß die Wahl eines zweckmäßigen Zustandsrepertoires, welches nicht nur die Ruhezustände enthalten soll, bedeutend schwieriger ist, als wenn man sich ausschließlich auf die Ruhezustände beschränkt. In der vorliegenden Arbeit wird dieser Problemkreis nicht weiter verfolgt, weil er für die Betrachtung der Arbeiten von Moore und Mealy irrelevant ist.

3. Die Modellvorstellungen von Moore und Mealy

3.1 Biographische Beziehungen zwischen Caldwell, Shannon, Huffman, Moore und Mealy

Zur Zeit von Moore und Mealy war das formale Modell der sequentiellen Maschine mit einem Eingaberepertoire repX, einem Ausgaberepertoire repY und einem Zustandsrepertoire repZ sowie einer Zustandsübergangsfunktion δ und einer Ausgabefunktion ω bereits etabliert. Im Vorwort zu dem Sammelband [Shannon–56], in dem auch Moore's Aufsatz erschien, sagt Shannon: "*Thus, such a device is characterized by two functions of the current state and input, one function giving the next state and the other the next output. Although seemingly trivial at this level of description, many interesting problems arise in the detailed analysis of such machines.*" Shannon bezeichnet also das formale Maschinenmodell als anscheinend trivial. Es wurde ja bereits 1936 von Turing benutzt [Turing–36], der zwar in seiner Arbeit das Wort *state* nicht verwendet, aber unter der Bezeichnung "*m–configuration*" genau den Zustand eines Automaten mit endlichem repZ meint.

Mealy's Aufsatz [Mealy–55] hat den Titel "A Method for Synthesizing Sequential Circuits". Der Aufsatz beginnt mit dem Satz: "*The theoretical basis of sequential circuit synthesis is developed, with particular reference to the work of D.A. Huffman and E.F. Moore.*" Obwohl Moore's Aufsatz [Moore–56] etwas später erschien als der Aufsatz von Mealy, konnte Mealy beim Schreiben seines Aufsatzes doch schon auf den Aufsatz von Moore Bezug nehmen. Moore und Mealy arbeiteten unter der Leitung von Shannon bei den Bell Telephone Laboratories, und daher kannte Mealy das bereits zur Veröffentlichung eingereichte Manuskript von Moore.

Huffman hatte am MIT unter der Betreuung von S.H. Caldwell, der bereits viel früher auch Shannon bei der Anfertigung seiner Masterarbeit [Shannon–38] beraten hatte, eine Doktorarbeit [Huffman–54] geschrieben, die 1954 unter dem Titel "*The Synthesis of Sequential Switching Circuits*" veröffentlich wurde. Es geht darin um die Synthese von Schaltwerken, wobei nicht nur deren Ruhezustände, sondern auch instabile Zustände betrachtet werden. Huffman führt kein maschinenunabhängiges formales Modell ein, sondern argumentiert immer unter Bezug auf konkrete Charakteristika der realen Schaltungen, die bei ihm zum überwiegenden Teil Relaisschaltungen sind. Dennoch führt er dort, wo er Schaltungsrealisierungen mit Elektronenröhren betrachtet, bereits das heute noch gültige Aufbaumodell für asynchrone Schaltwerke in Form eines rückgekoppelten Schaltnetzes ein. Deshalb wird Huffman mit Recht als Begründer der Theorie asynchroner Schaltwerke betrachtet.

3.2 Die Modellvorstellung von Moore

Moore's Aufsatz trägt den Titel "Gedanken–Experiments on Sequential Machines". Im Unterschied zu Huffman betrachtet Moore keine konkreten Maschinen, sondern er betrachtet tatsächlich ein formales Maschinenmodell. Bezüglich der Eingabe– und Ausgaberepertoires spricht er von Symbolen. Über das Verhalten der Maschine sagt er, daß der gegenwärtige Zustand nur von der vorangegangenen Eingabe und dem vorangegangenen Zustand abhängt und daß die gegenwärtige Ausgabe nur vom gegenwärtigen Zustand abhängt. An keiner Stelle spricht er explizit von einer Zustandsübergangsfunktion oder einer Ausgabefunktion. Diese Funktionen kommen bei ihm stets nur in tabellarischer Form vor.

Obwohl Moore an keiner Stelle seiner Arbeit eine konkrete Maschinenrealisierung betrachtet, weist er doch darauf hin, daß er in der Anschauung mit seinem formalen Modell ein ge-

taktetes Schaltwerk verbindet. Moore hat zweifellos seine Arbeit als Beitrag zur Schaltwerkstheorie betrachtet. Denn immerhin war er Mitarbeiter der Bell Laboratories, wo man sich mit dem Entwurf von Computern und von vermittlungstechnischen Einrichtungen befaßte. Moore erwähnt diese Arbeitsgebiete in seinem Aufsatz. Hinweise auf andere Maschinen – beispielsweise auf Verkaufsmaschinen, in die man Geld hineinwirft und aus denen Ware herausfällt – findet man dagegen in Moore's Arbeit nicht. Man interpretiert Moore sicher korrekt, wenn man behauptet, daß ihm bezüglich seiner Maschinen der in Bild 3 gezeigte Verlauf vor Augen gestanden hat.

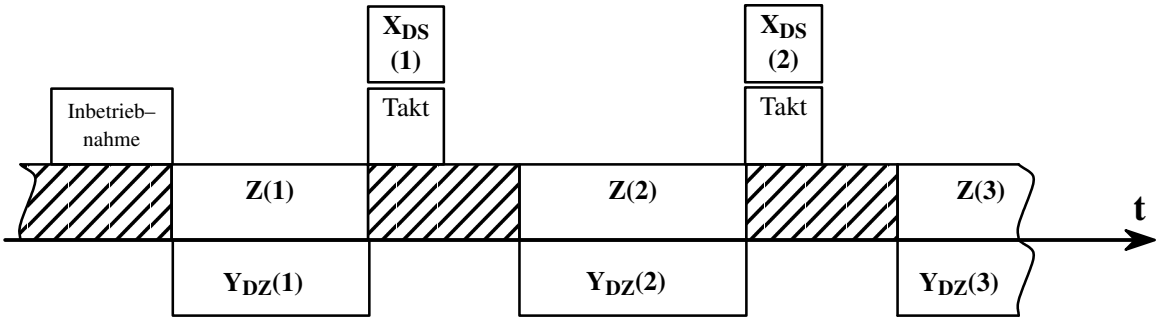


Bild 3 Der Verlauf der Ein– und Ausgabeerscheinungen einer "Moore–Maschine"

In der Tatsache, daß in dem Verlauf in Bild 3 die Y–Folge die gleiche Länge hat wie die Z–Folge, während die X–Folge ein Element weniger hat, liegt bereits eine der Ursachen für die spätere begriffliche Kollision. Möglicherweise hat Moore selbst zu diesem Problem beigetragen, denn in den tabellarischen Verläufen, die er als Beispiele in seiner Arbeit anführt, hat er die drei Folgen für X, Y und Z jeweils gleich lang gemacht, indem er ein letztes X–Element angegeben hat, welches eigentlich irrelevant ist, wenn man den damit festliegenden Folgezustand nicht mehr einträgt.

Zu dem in Bild 3 dargestellten Verlauf gehört das in Bild 4 gezeigte Blockschaltbild eines getakteten Schaltwerks.

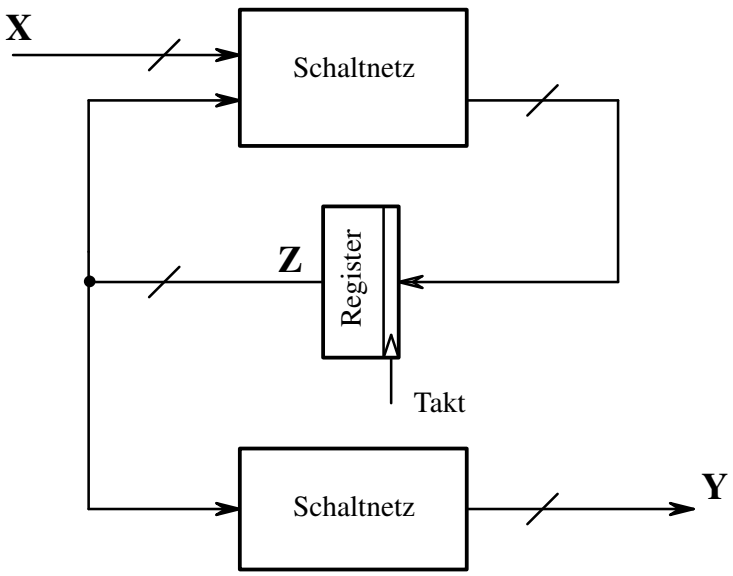


Bild 4 Blockschaltbild eines "Moore–Schaltwerks"

Das formale Modell, welches die Vorstellungen von Moore zum Ausdruck bringt, lautet:

$$Y(n) = \sigma [Z(n)]$$
$$Z(n+1) = \delta [Z(n), X(n)]$$

Dieses Modell ist ein anderes Automatenmodell als dasjenige, welches im Abschnitt 2.3 durch Zeitdiskretisierung des Modells des deterministischen kausalen Systems gewonnen wurde. Man sollte es auch nicht als Sonderfall des dortigen Modells betrachten. Denn wenn X(n) formal im Argument der Ausgabefunktion steht und nur sein Wert irrelevant ist, ist dies nicht dasselbe, wie wenn X(n) im Argument gar nicht vorkommt. Im Fall

$$Y(n) = \omega [Z(n), \text{wertirrelevantes } X(n)]$$

hat man nämlich die Vorstellung, daß das Ausgabeelement Y(n) zwar nicht vom Wert des Eingabeelements X(n) abhängt, aber doch von seinem Auftreten, d.h. man hat die Vorstellung, daß Y(n) erst ausgegeben werden kann, nachdem X(n) eingegeben wurde. Anders liegt der Fall, wenn die Ausgabefunktion die Form

$$Y(n) = \sigma [Z(n)]$$

hat, denn in diesem Fall hat man die Vorstellung, daß Y(n) bereits ausgegeben werden kann, sobald die Maschine den Zustand Z(n) einnimmt. Dies bedeutet unter anderem, daß bereits unmittelbar nach der Inbetriebnahme, also wenn der Zustand Z(1) eingenommen wird, eine Ausgabe erfolgen kann, ohne daß auf die erste Eingabe X(1) gewartet werden muß.

Aus Moore's Aufsatz geht eindeutig hervor, daß es ihm nicht darum ging, erstmalig ein bestimmtes formales Maschinenmodell zu präsentieren, sondern daß ihn die Frage interessierte, was man experimentell über eine sequentielle Maschine herausfinden könne, wenn man nicht in ihr Inneres schauen darf. Moore ist ebenso wie Huffman und Mealy primär an der methodischen Synthese von Schaltwerken interessiert, wobei die Frage nach dem erforderlichen Zustandsrepertoire von zentraler Bedeutung ist.

3.3 Die Modellvorstellung von Mealy

In der Einleitung seines Aufsatzes sagt Mealy, daß es ihm darum geht, die Zustandsreduktionsmethoden von Huffman und Moore zu verbessern, indem er die Erkenntnisse seiner beiden Vorgänger in geeigneter Weise zusammenbringt.

Obwohl sein Modell – ebenso wie das von Moore – auf andere Maschinen übertragbar ist, hatte Mealy ebenso wie Moore nur Schaltwerke vor Augen. Im Unterschied zu Moore, der sich auf die Vorstellung getakteter Schaltwerke beschränkt, betrachtet Mealy sowohl getaktete als auch ungetaktete Schaltwerke. Er stellt sein formales Modell anhand der getakteten Schaltwerke vor und überträgt es anschließend auch auf die ungetakteten Schaltwerke. Für den Vergleich mit Moore genügt es, nach Mealy's Modellierung der getakteten Schaltwerke zu fragen.

Bei getakteten Schaltwerken stellt es ein Problem dar, wenn man verlangt, daß die gegenwärtige Ausgabe von der gegenwärtigen Eingabe abhängen soll. Denn im Falle der Schaltwerke ist die Ausgabe vom Anzeigetyp; es kann sich nach Bild 2 also nur um die Ausgaben Y_D

handeln. Die geforderte Endlichkeit von repX macht es nötig, dafür zu sorgen, daß X_{DY} im Bereich zwischen zwei Zustandsübergangsintervallen jeweils einen konstanten Wert hat. Deshalb verlangt Mealy, daß die Eingaben synchronisiert angeboten werden. Das bedeutet aber, daß er außerhalb seiner Maschine einen Synchronisationsmechanismus annehmen muß, der mit dem gleichen Takt versorgt wird wie die Maschine selbst. Zu der Argumentation von Mealy gehört also das Blockschaltbild eines getakteten Schaltwerks, wie es in Bild 5 gezeigt ist.

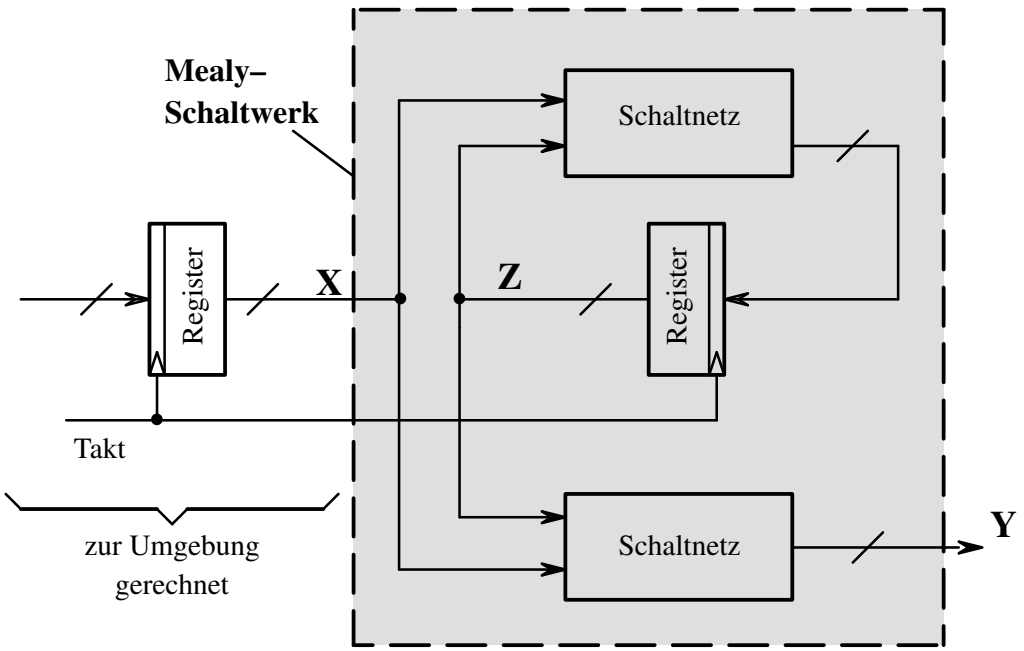


Bild 5 Blockschaltbild des "Mealy–Schaltwerks"

Beim Vergleich dieses Blockschaltbilds mit der Struktur in Bild 4 erkennt man leicht, daß es sich im Grunde in beiden Fällen um dieselbe Art von System handelt; Moore hat lediglich das Register, welches außerhalb des Mealy–Schaltwerks liegt, in seine Maschine hineintransformiert. Eine solche Transformation wird zwangsläufig manchmal dazu führen, daß das Repertoire der Zustände bei Moore etwas größer sein muß als bei Mealy.

Die Darstellung in Mealy's Aufsatz ist allerdings keineswegs derart, daß man auf den ersten Blick erkennen könnte, daß Mealy tatsächlich die in Bild 5 gezeigte Struktur vor Augen hatte. Er fordert zwar im Text, daß "*input pulses remain synchronized with the clock*", aber in seinen Schaltbildern kommt der Takt gar nicht vor. Als Schaltbilder verwendet er rückgekoppelte Schaltnetze mit Einheitsverzögerungen, also eine Form, die im Grunde nur für die Betrachtung asynchroner Vorgänge geeignet ist. Er redet zwar von "*pulses*", versteht darunter aber offensichtlich nicht das, was spätere Schaltwerkstheoretiker wie beispielsweise McCluskey darunter verstehen. Mealy's Satz "*Inputs and outputs are in the form of voltage or current pulses which occur synchronously with pulses from the clock*" kann mit den von Mealy angegebenen Schaltbildern nur dadurch in Einklang gebracht werden, daß man ihn wie folgt interpretiert: An den Eingängen und den Ausgängen gibt es Binärsignale in Form von Spannungen oder Strömen, die ihren Wert jeweils nur zu den Taktzeitpunkten ändern können.

Mit dieser vermutlich einzig schlüssigen Interpretation von Mealy's Aussagen zu seiner Maschinenvorstellung gehört der Verlauf in Bild 6. Im Unterschied zu Bild 3 ist hier die Y–Folge um ein Element kürzer als die Z–Folge und hat damit die gleiche Länge wie die X–Folge.

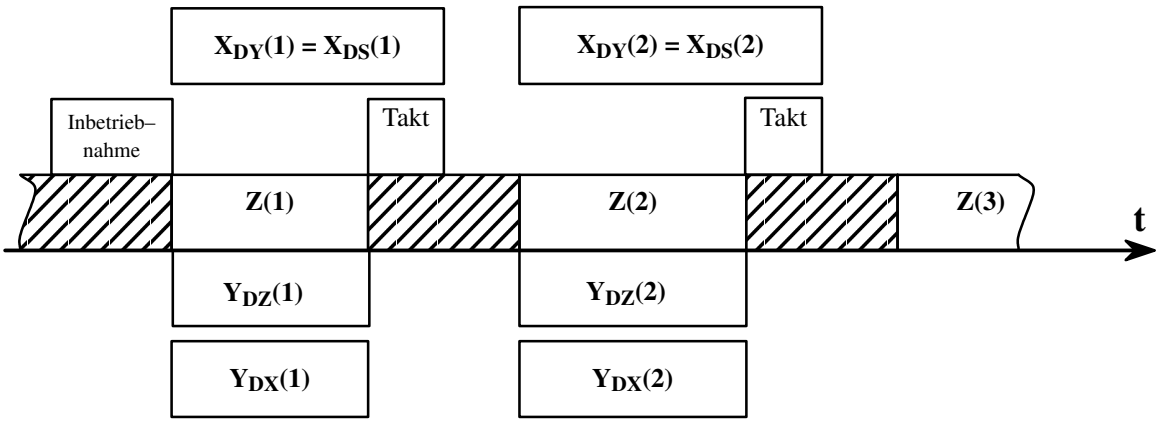


Bild 6 Der Verlauf der Ein– und Ausgabeerscheinungen des "Mealy–Schaltwerks"

Das formale Modell, welches die Vorstellungen von Mealy zum Ausdruck bringt, ist also gleich dem Modell, welches im Abschnitt 2.3 durch Zeitdiskretisierung des Modells des deterministischen kausalen Systems gewonnen wurde:

$$Y(n) = \omega\,[\,Z(n), X(n)\,]$$
$$Z(n+1) = \delta\,[\,Z(n), X(n)\,]$$

Obwohl diese Modellvorstellung bereits vor Mealy in natürlichsprachlicher Form bekannt war, zwingt die jahrzehntelange Praxis dazu, dieses Modell auch in Zukunft als Automatenmodell nach Mealy zu bezeichnen – in gleicher Weise, wie der Satz von den Quadraten über den Seiten eines rechtwinkligen Dreiecks unter dem Namen Pythagoras läuft, obwohl er schon vor der Zeit des Pythagoras bekannt war.

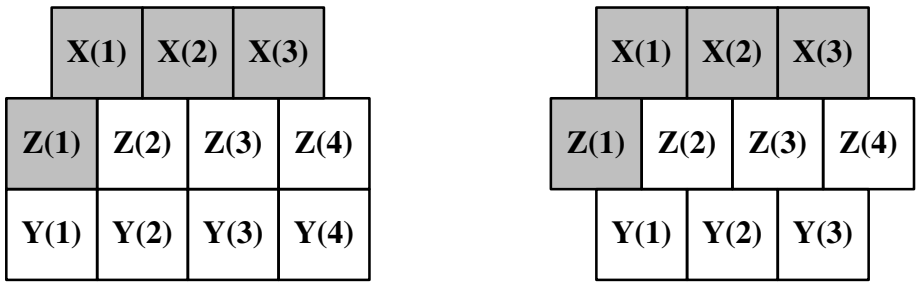
Die Leistung Mealys besteht nicht darin, dieses formale Modell erfunden zu haben, sondern darin, es konsequent auf Schaltwerke angewendet zu haben.

3.4 Die Modelle von Moore und Mealy im Vergleich

Die formalen Automatenmodelle von Moore und Mealy werden bestimmt durch drei diskrete Repertoiremengen repX, repY und repZ, eine Ausgabefunktion, über die das Ausgabeelement Y(n) zu bestimmen ist, sowie eine Zustandsübergangsfunktion, über die der Folgezustand Z(n+1) zu bestimmen ist. Bild 7 zeigt die unterschiedliche Form der Tabellen zur Darstellung von formalen Experimenten mit einem Automatenmodell nach Moore bzw. Mealy.

$Y(n) = \sigma [Z(n)]$ $Z(n+1) = \delta [Z(n), X(n)]$	$Y(n) = \omega [Z(n), X(n)]$ $Z(n+1) = \delta [Z(n), X(n)]$

.



Moore–Modell Mealy–Modell

Bild 7 Form der Tabellen zur Darstellung von formalen Experimenten mit den beiden Automatenmodellen nach Moore und Mealy

Nun können die Fragen beantwortet werden, die im Abschnitt 1 aufgeführt sind.

Frage: Ist das Moore–Modell ein Sonderfall des Mealy–Modells?

Antwort: Nein.

Frage: Hat Moore gesagt, daß die jetzige Ausgabe nur vom Zustand *vor* dem Übergang abhängt, oder hat er gesagt, daß sie nur vom Zustand *nach* dem Übergang abhängt, oder hat er sich nicht festgelegt?

Antwort: Er hat sich festgelegt, und zwar auf den Zustand *vor* dem Übergang.

Frage: Soll man mit den Namen Moore und Mealy bestimmte Maschinentypen verbinden oder nur Formeln, die einen formalen Zusammenhang zwischen Symbolfolgen beschreiben?

Antwort: Mit den Namen Moore und Mealy sollte man nicht nur Formeln verbinden, die einen formalen Zusammenhang zwischen Symbolfolgen beschreiben, sondern auch die Vorstellung unterschiedlicher Maschinentypen.

Eine *Moore–Maschine* ist eine sequentielle Maschine, deren Ausgabeerscheinungen auf den Typ Y_{DZ} beschränkt sind. Eine Moore–Maschine wird durch ein Moore–Automatenmodell angemessen modelliert.

Eine sequentielle Maschine, welche keine Moore–Maschine ist, wird als *Mealy–Maschine* bezeichnet. Eine Mealy–Maschine wird durch ein Mealy–Automatenmodell angemessen modelliert.

Ein *Moore–Schaltwerk* ist ein Sonderfall einer Moore–Maschine. Ein Schaltwerk, welches kein Moore–Schaltwerk ist, wird als *Mealy–Schaltwerk* bezeichnet.

Frage: Sind die Modelle von Moore und Mealy einander äquivalent, und falls ja, in welchem Sinne?

Antwort: Wegen der Längenunterschiede bei den Y–Folgen können ein Moore–Modell und ein Mealy–Modell einander in keinem Sinne äquivalent sein.

Falls eine Äquivalenz vorgelegen hätte, hätte es entweder eine Schnittstellenäquivalenz oder eine Isomorphieäquivalenz sein müssen.

Schnittstellenäquivalenz zweier Automatenmodelle A und B liegt vor, wenn es zu jedem möglichen Anfangszustand $Z_A(1)$ im Modell A einen Anfangszustand $Z_B(1)$ im Modell B gibt und umgekehrt, derart, daß unter der Annahme dieser Anfangszustände die beiden Modelle die gleiche Funktion beschreiben, die jeder endlichen Folge von Elementen aus repX jeweils eine Folge von Elementen aus repY zuordnet.

Isomorphieäquivalenz zweier Automatenmodelle A und B liegt vor, wenn es zu jedem formalen Experiment mit dem Modell A ein formales Experiment mit dem Modell B gibt und umgekehrt, derart, daß die Elemente der Folgen "$X_A(n)$", "$Z_A(n)$" und "$Y_A(n)$" den Elementen der Folgen "$X_B(n)$", "$Z_B(n)$" und "$Y_B(n)$" umkehrbar eindeutig zugeordnet werden können.

4. Abweichende Vorstellungen

Wenn spätere Schaltwerks– oder Automatentheoretiker mit den Namen Moore und Mealy Vorstellungen verbanden, die von den in den Arbeiten von Moore und Mealy zum Ausdruck gebrachten Vorstellungen abweichen,
- dann haben sie entweder gar nicht bei Moore und Mealy nachgeschaut,
- oder sie haben sich durch Mealy's zum Teil recht konfuse Darstellung verwirren lassen,
- oder aber sie waren sich der Abweichungen bewußt und haben die Namen Moore und Mealy nur noch als Markenzeichen verwendet.

Wie es im einzelnen war, kann hier offen bleiben.

4.1 Abweichung bezüglich der Maschinenvorstellungen

Es wurde schon gesagt, daß Mealy in seinem Aufsatz von ”pulses” spricht und daß es Schwierigkeiten bereite, dies mit seinen Schaltbildern in Einklang zu bringen. Dies hat manche Autoren veranlaßt, den Unterschied zwischen Moore und Mealy dadurch zu charakterisieren, daß sie auf einen angeblichen Unterschied in der Art der Ein– und Ausgabeerscheinungen dieser Schaltwerke hinweisen.

W.J. Cadden [Cadden–59] kennzeichnet ein Moore–Schaltwerk als ein Pulse–Input–Level–Output–Schaltwerk, abgekürzt PL. Ein Mealy–Schaltwerk ist seiner Einschätzung nach PP, und ein Huffman–Schaltwerk LL.

Auch E.J. McCluskey [McCluskey–86] sieht den Unterschied zwischen Moore und Mealy durch die Charakterisierung PL bzw. PP erfaßt. Daraus folgert McCluskey, daß ein Schaltwerk nur dann ein Mealy–Schaltwerk sei, wenn das Schaltwerk die in Bild 8 gezeigte Struktur habe und alle Binärkomponenten der Ausgabe Y durch eine Schaltnetzverknüpfung mit dem Taktsignal entstehen. Aufgrund seiner Definition kommt McCluskey konsequenterweise zu dem Schluß, daß es nicht nur Schaltwerke des Mealy– und des Moore–Typs gebe, sondern auch solche, die keinem der beiden Typen zuzuordnen seien.

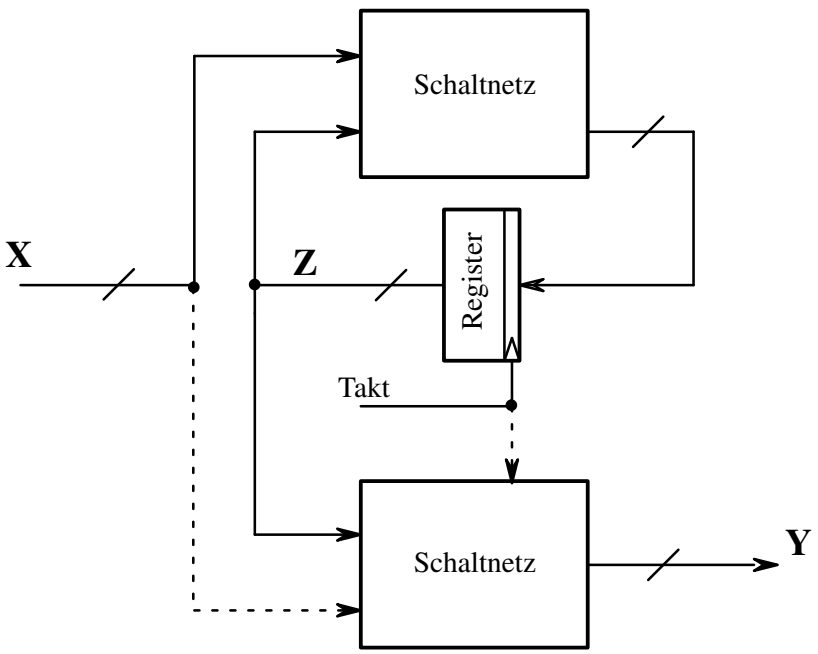


Bild 8 Blockschaltbild eines Schaltwerks, anhand dessen McCluskey den Unterschied zwischen den Schaltwerkstypen von Moore und Mealy definiert

4.2 Abweichungen bezüglich des formalen Modells

Die Zustandsübergangsfunktion

$$Z(n+1) = \delta [Z(n), X(n)] ,$$

die ja ohnehin bei Moore und Mealy die gleiche Form hat, bot kein Potential für alternative Interpretationen. Die Ausgabefunktionen ω und σ jedoch sind bis zum heutigen Tage Gegenstand etlicher Debatten unter Schaltwerkstheoretikern geblieben. Die Ausgabefunktionen wurden oft nicht in der Form

$$Y(n) = \omega [X(n), Z(n)] \qquad \text{bzw.} \qquad Y(n) = \sigma [Z(n)],$$

angegeben, sondern in einer Form ohne den Folgenindex n:

$$X \times Z \rightarrow Y \text{ (Mealy)} \qquad \text{bzw.} \qquad Z \rightarrow Y \text{ (Moore).}$$

Damit war die Möglichkeit gegeben, daß jemand auf die Idee kommen konnte, es müsse bei Moore

$$Y(n) = \mu [Z(n+1)]$$

heißen. Allerdings wird diese Festlegung meist nicht in dieser Form hingeschrieben, sondern es wird einfach eine von Moore abweichende Zählung festgelegt. Indem die Eingabefolge "X(n)" der Länge 0 aus der Betrachtung ausgeschlossen wird, wird implizit festgelegt, daß das Ausgabeelement, welches im Moore–Modell als Y(2) bezeichnet wird, nun das erste der Ausgabefolge sein soll. Das Y(1) aus dem Moore–Modell fällt dann aus der Betrachtung heraus [Mikolajczak–91]. Man erhält also auf diese Weise ein Automatenmodell, bei dem die Ausgabefolge um ein Element kürzer ist als beim Moore–Modell – und das ist ein Mealy–Modell, wie man in Bild 7 sofort erkennt.

Für die Frage, ob ein Modell ein Mealy–Modell sei oder nicht, ist es selbstverständlich völlig unerheblich, wie man in einer graphischen Darstellung der X–, Y– und Z–Folgen die Lagerelation zwischen der Y– und der Z–Folge wählt (s. Bild 9).

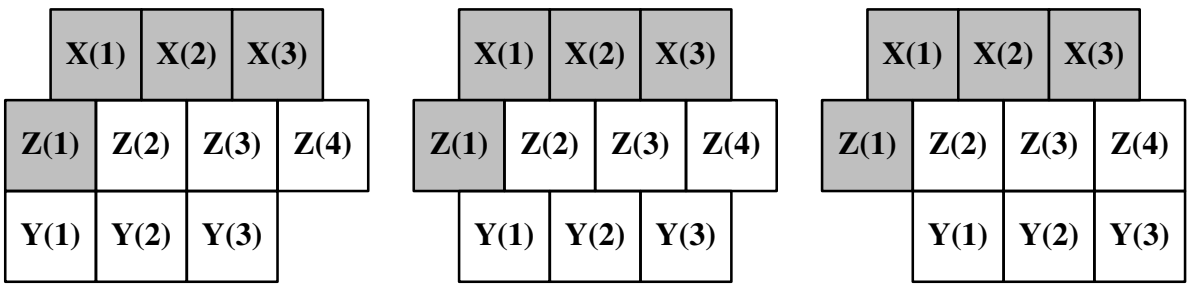


Bild 9 Alternative graphische Darstellungen der Folgen in einem Mealy–Modell

In der Literatur wird das Modell

$$\begin{aligned} Y(n) &= \mu [Z(n+1)] \\ Z(n+1) &= \delta [Z(n), X(n)] \end{aligned}$$

von vielen Autoren als "Moore"–Modell bezeichnet, obwohl es doch etwas anderes ist als das von Moore benutzte Modell. Es ist nämlich nichts anderes als ein Sonderfall eines

Mealy–Modells. Wenn die beiden Funktionen μ und δ gegeben sind, ist nämlich auch die Funktion ω eindeutig bestimmt, da man das Argument von μ durch das Ergebnis von δ ausdrücken kann:

$$\omega [X(n),\ Z(n)] = \mu (\delta [X(n),\ Z(n)]) .$$

Das Modell paßt auf Mealy–Maschinen, die sich jeweils in ihrem aktuellen Zustand Z(n) für n>1 noch daran erinnern, was die letzte Ausgabe Y(n–1) war. Das Modell für solche Maschinen soll *Pseudo–Moore–Modell* genannt werden.

Im Abschnitt 3.4 wurde bereits gesagt, daß ein Mealy–Modell und ein authentisches Moore–Modell, d.h. ein Modell mit der Funktion σ, in keinem Sinne äquivalent sein können, da sie sich hinsichtlich der Längen ihrer Y–Folgen unterscheiden. Dagegen kann zu einem Pseudo–Moore–Modell immer ein schnittstellenäquivalentes Mealy–Modell angegeben werden und umgekehrt [Gill–60].

Die Transformation eines Pseudo–Moore–Modells in ein schnittstellenäquivalentes Mealy–Modell ist trivial, weil es selbst nur ein Sonderfall eines Mealy–Modells ist.

Die umgekehrte Transformation eines gegebenen Mealy–Modells in ein schnittstellenäquivalentes Pseudo–Moore–Modell kann es erforderlich machen, anstelle des vorgegebenen $repZ_{Mealy}$ ein mächtigeres $repZ_{Pseudo–Moore}$ einzuführen. Dies ist immer dann erforderlich, wenn sich das gegebene Mealy–Modell nicht in jedem Fall an die jeweils letzte Ausgabe erinnert. Es wird hier darauf verzichtet, das formale Transformationsverfahren anzugeben.

Der grundsätzliche Unterschied zwischen dem Moore–Modell und dem Pseudo–Moore–Modell äußert sich auch in unterschiedlichen Bedingungen, die an die Definitionsbereiche der Ausgabefunktionen σ und μ gestellt werden. Während der Definitionsbereich der Funktion σ eines Moore–Modells in jedem Falle gleich repZ sein muß, kann der Definitionsbereich der Funktion μ eines Pseudo–Moore–Modells eine echte Teilmenge von repZ sein. Das Beispiel in Bild 10 veranschaulicht diesen Sachverhalt. In diesem Beispiel kommt der Anfangszustand I nie als Folgezustand vor, und deshalb umfaßt hier der Definitionsbereich der Funktion μ nur die beiden Zustände II und III.

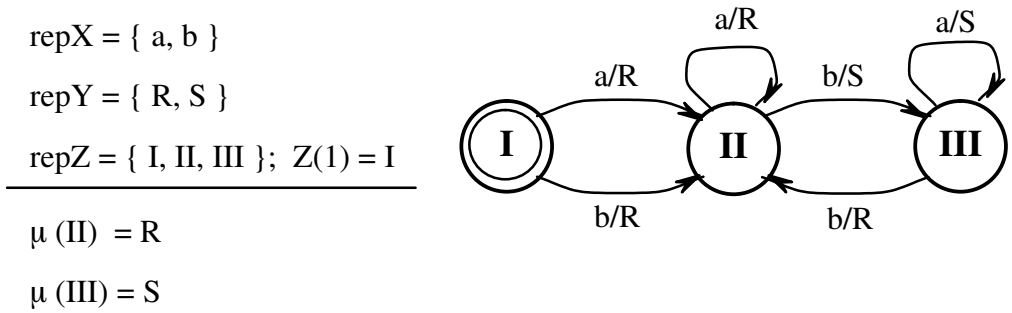


Bild 10 Beispiel eines Pseudo–Moore–Automatenmodells, bei dem der Definitionsbereich der Funktion μ nicht gleich repZ ist

Während das Pseudo–Moore–Modell in der Automatentheorie eine große Rolle spielt, ist es in der Schaltwerkstheorie praktisch bedeutungslos. Hinsichtlich des Moore–Automatenmodells liegen die Verhältnisse gerade umgekehrt: In der Schaltwerkstheorie hat dieses Modell eine große Bedeutung, während es für die Automatentheorie völlig entbehrlich ist.

5. Danksagung

Zu den in diesem Aufsatz dargestellten Erkenntnissen und Vorstellungen hat mein Kollege an der Universität Kaiserslautern, Dr. Jochen Beister, in großem Umfang beigetragen. Da wir beide unter der Begriffskonfusion, die mit den Namen Moore und Mealy verbunden ist, zu leiden hatten, schlug ich ihm vor, das Problem an der Wurzel zu packen und eine konsistente Sprachregelung in einem gemeinsamen Aufsatz darzustellen. Während er in der Anfangsphase noch sehr viel Zeit in das Projekt investieren konnte und wir in der Diskussion zu einer gemeinsamen Sicht fanden, fand er in der Schlußphase doch keine Zeit mehr und mußte es mir alleine überlassen, den geplanten Aufsatz zu schreiben.

6. Literatur

[Cadden–59] W. J. Cadden: Equivalent Sequential Circuits.
IRE Transactions on Circuit Theory, vol. CT–6, pp. 30–34; 1959.

[Gill–60] A. Gill: Comparison of Finite–State Models.
IRE Transactions on Circuit Theory, vol. CT–7, pp. 178–179; 1960.

[Huffman–54] D. A. Huffman: The Synthesis of Sequential Switching Circuits.
Journal of The Franklin Institute, vol. 257, pp. 161–190, 275–303; 1954.

[McCluskey–86] Edward J. McCluskey: Logic Design Principles.
Prentice Hall, Englewood Cliffs, N.J., 1986.

[Mealy–55] George H. Mealy: A Method for Synthesizing Sequentials Circuits.
Bell System Technical Journal, vol. 34, pp. 1045–1079; 1955.

[Mikolajczak–91] B. Mikolajczak (Editor): Algebraic and Structural Automata Theory.
North–Holland, Amsterdam; 1991.

[Moore–56] Edward F. Moore: Gedanken–Experiments on Sequential Machines.
in Automata Studies, Princeton University Press, pp. 129–153; 1956.

[Shannon–38] C. E. Shannon: A Symbolic Analysis of Relay and Switching Circuits.
Transactions of the American IEE, vol. 57, pp. 713–723; 1938.

[Shannon–56] C. E. Shannon (Editor): Automata Studies.
Princeton University Press; 1956.

[Turing–36] A. M. Turing: On Computable Numbers, with an Application to the Entscheidungsproblem.
Proceedings of the London Mathematical Society, vol. 42; 1936.