

Modell eines Lernenden Automaten

Von

MANFRED HUBER

Mitteilung aus dem Institut für Nachrichtenvermittlung und Datenverarbeitung
der Technischen Hochschule Stuttgart

Mit 3 Textabbildungen

(Eingegangen am 28. September 1964)

Übersicht: Es werden Prinzip und elektronische Realisierung eines Lernenden Automaten beschrieben, dem die Methode „Lernen durch Erfolg“ zugrunde liegt. Der Lernprozeß wird in seinen Einzelheiten festgelegt und in einem Flußdiagramm dargestellt. Zur Realisierung eines Lernenden Automaten wurden drei Speicher verwendet. Es sind dies ein Verhaltensspeicher, der die verschiedenen Verhaltensweisen des Lernenden Automaten enthält; ferner ein Merkspeicher, in dem die erfolgreiche Verhaltensweise gespeichert wird; schließlich ein Erfahrungsspeicher, der zur Speicherung eines schrittweise sich ändernden Erfahrungsschatzes dient.

Summary: The paper describes the principle and electronic realization of a Learning Machine based on the method of learning by trial and error. The details of a learning process are defined and graphically presented in a flow chart. As to the realization of the Learning Machine three stores have been found necessary. A behaviour store contains all forms of behaviour that can be produced by the Learning Machine. A memory store marks the successful behaviour and an experience store follows up the step by step change of the Learning Machine's experience.

Einleitung

Von den verschiedenen Arten des Lernens, die ZEMANEK [1] beschreibt, sei diejenige des „Lernens durch Erfolg“ herausgegriffen. Dieser Begriff entspricht etwa dem „learning by trial and error“ in der englischsprachigen Literatur. Nimmt man nach STEINBUCH [2] die Unterscheidung in eine Lern- und Kannphase hinzu und stellt sich einen Automaten in Verbindung mit seiner Außenwelt vor, so kann darunter folgendes verstanden werden:

a) In der Lernphase probiert der Automat eine Auswahl von verschiedenen Verhaltensweisen, die ihm technisch möglich sind, systematisch oder in zufälliger Folge durch (trial). Wichtig ist, daß der Automat dabei aufgrund der Bedingungen der Außenwelt eindeutig zwischen „Erfolg“ und „Mißerfolg“ (error) unterscheiden kann. Der Automat merkt sich die erfolgreiche Verhaltensweise.

b) In der Kannphase wendet der Automat unmittelbar, d. h. ohne vorheriges Probieren, die in der Lernphase gemerkte Verhaltensweise an. Im Gegensatz zur nur „gemarkten“ Verhaltensweise in der Lernphase kann in der Kannphase von einer „gelernten“ Verhaltensweise gesprochen werden.

Der verwirklichte Lernprozeß

Für den Bau eines Modells eines Lernenden Automaten (LA) wurden über den beschriebenen Sachverhalt hinaus weitere Festlegungen getroffen, die insbesondere den Übergang zwischen der Lern- und Kannphase beschreiben.

Es wird von der Vorstellung ausgegangen, daß der Lernprozeß sich aus einzelnen Schritten zusammensetzt, bei denen ein Erfahrungsschatz schrittweise aufgebaut werden muß. Das Überschreiten der Schwelle zur Kannphase wird vom Umfang dieser Erfahrung abhängig gemacht. Der LA entscheidet sich daher erst nach mehreren erfolgreichen Versuchen in der Lernphase für die zugehörige Verhaltensweise.

Darüber hinaus wird im LA die bekannte Tatsache berücksichtigt, daß die Beherrschung einer bereits gelernten Verhaltensweise durch Gewöhnung vertiefbar ist. Hierbei wird der

Erfahrungsschatz vermehrt. In der Kannphase tritt beim LA daher immer dann ein sogenannter Vertiefungsschritt auf, wenn die gelernte Verhaltensweise durch Erfolg bestätigt wird.

Umgekehrt muß die angesammelte Erfahrung auch wieder abgebaut werden können. Im LA geschieht dies auf zweierlei Weise:

a) Vergessen. In gleichen Zeitabständen trifft jeweils ein Abbauschritt ein, der den Erfahrungsschatz um einen Schritt abbaut. Es handelt sich hier um eine Vergessensfunktion, die unabhängig vom Zustand des Lernprozesses ist. Es wäre auch möglich gewesen, die Vergessenschritte in statistisch verteilten Zeitabständen einfallen zu lassen und die Wahrscheinlichkeitsdichte für das Eintreffen eines Vergessenschrittes abhängig zu machen von dem Umfang der aufsummierten Erfahrung oder von der Zeit, die seit Lernbeginn verstrichen ist.

b) Umlernen. In der Lernphase kann der Fall eintreten, daß die gemerkte Verhaltensweise keinen Erfolg mehr erzielt, weil sich die Bedingungen der Außenwelt geändert haben. Der deshalb erforderliche Umlernvorgang setzt mit einem Abbauschritt ein. Es wurde die Festlegung getroffen, daß alle vorhandenen Lernschritte bezüglich der bisher richtigen Verhaltensweise schrittweise abgebaut werden müssen, bevor erneut Lernschritte für eine jetzt erfolgreiche neue Verhaltensweise einsetzen können.

Auch in der Kannphase sind Abbauschritte für einen Umlernvorgang vorgesehen, wenn die bisher erfolgreich angewandte Verhaltensweise infolge veränderter Außenweltbedingungen zu einem Mißerfolg führt.

Treten die ursprünglich zum Erfolg führenden Bedingungen der Außenwelt wieder auf, bevor ein zwischenzeitlicher Umlernvorgang sämtliche Lern- und Vertiefungsschritte abgebaut hat, so kann der noch vorhandene Erfahrungsschatz wieder verwertet werden.

Flußdiagramm

Eine Übersicht über den Ablauf des soeben beschriebenen Lernprozesses gibt das Flußdiagramm in Bild 1. Es sollen zunächst die Vorgänge in der Lernphase betrachtet werden.

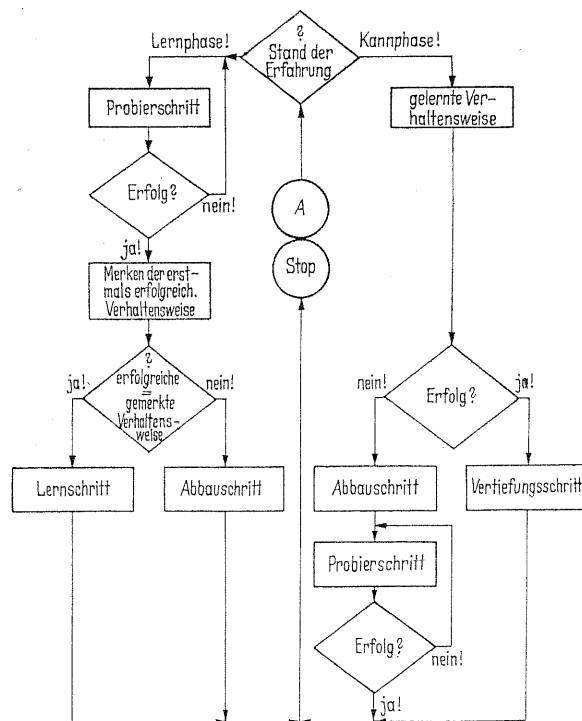


Bild 1. Flußdiagramm.

Nach der Auslösung des Starts (A) und der Entscheidung „Lernphase“ aufgrund des Erfahrungsstandes wendet der LA nacheinander verschiedene Verhaltensweisen an (Probierschritte), und zwar so lange bis „Erfolg“ eintritt. Der LA merkt sich die erfolgreiche Verhaltensweise erstmalig zu Beginn eines Neulernvorganges, wenn noch keinerlei Erfahrung bezüglich irgendeiner Verhaltensweise vorliegt. Bevor der Programmablauf der Lernphase bei „Stop“ endet, wird ein Lernschritt registriert, wenn die gemerkte Verhaltensweise zum Erfolg geführt hat, dagegen ein Abbauschritt, wenn die gemerkte Verhaltensweise nicht mehr mit der erfolgreichen Verhaltensweise übereinstimmt.

Nun werde die Kannphase betrachtet. Hier tritt nach Start (A) und Entscheidung für die Kannphase sofort die gelernte Verhaltensweise ein. Zwei Fälle werden wieder unterschieden:

a) die gelernte Verhaltensweise wird durch die Außenwelt bestätigt („Erfolg ja“, Vertiefungsschritt, Stop);

b) eine andere als die gelernte Verhaltensweise ist erfolgreich („Erfolg? nein!“, Abbauschritt, anschließend Probierschritte, bis die jetzt erfolgreiche neue Verhaltensweise gefunden worden ist, Stop).

Dem ganzen Flußdiagramm überlagert ist der in gleichen Zeitabständen eintreffende Vergessensschritt, der die jeweils aufsummierten Lern- und Vertiefungsschritte abbaut.

Der geschilderte Ablauf des Flußdiagramms beschreibt eine Reihe von täglich vorkommenden Lernprozessen. Man denke beispielsweise an den Fall, daß ein Mensch vor eine verschlossene Tür kommt und der zunächst noch unbekannte Schlüssel aus einem Schlüsselbund ausgewählt werden muß. Die Verbindung zum Flußdiagramm in Bild 1 ergibt sich aus folgender Gegenüberstellung:

Start:	Situation „Mensch vor verschlossener Tür, passender Schlüssel im Schlüsselbund“
Probierschritte:	Durchprobieren der Schlüssel
Erfolg:	Schlüssel S_i schließt das Schloß auf
gemarkte Verhaltensweise:	Merken des Schlüssels S_i , der zum ersten Mal gepaßt hat. Erst beim Eintritt in die Kannphase gelangen die Merkmale des gemerkten Schlüssels vom Unterbewußtsein ins Bewußtsein
gelernte Verhaltensweise:	Passender Schlüssel S_i wird sofort benützt
Mißerfolg in Kannphase:	Schloß ist ausgewechselt worden. Es paßt jetzt ein anderer Schlüssel S_j desselben Schlüsselbunds

Elektronische Realisierung

Die Baugruppen, die zur Realisierung eines LA in digitaler Technik notwendig sind, lassen sich aus dem Flußdiagramm ableiten. Diese Aufgabe wird durch eine besondere Eigenschaft des Flußdiagramms erleichtert. Wie eine Nachprüfung in Bild 1 ergibt, erfolgen zwischen zwei Entscheidungen höchstens zwei Operationen, die voneinander unabhängig sind und daher auch gleichzeitig ausgeführt werden können. Der ganze Prozeß läßt sich daher durch eine Anzahl Speicher beschreiben, deren Inhalte für sich oder deren Vergleich untereinander die Entscheidungen im Verlauf des Flußdiagramms bestimmen. Die Operationen selbst bestehen dann in der Veränderung der Speicherinhalte (Auf- und Abbauschritte) oder im Auslesen eines Speicherinhaltes (gelerntes Verhalten).

In einem Modell sind 3 Speichergruppen verwirklicht worden. Die folgende Zusammenstellung beschreibt die Zuordnung zwischen den Entscheidungen im Flußdiagramm und diesen Speichergruppen, sowie deren Realisierung.

a) „Stand der Erfahrung?“ — Erfahrungsspeicher E.

Vierstufiger, umkehrbarer Dualzähler. Ein Lern- oder Vertiefungsschritt wird als vorwärtszählender, ein Abbau- bzw. Vergessensschritt als abwärtszählender Schritt registriert. Der Zählerstand entscheidet über Lern- oder Kannphase.

b) „Erfolg?“ — Verhaltensspeicher V.

Zweistufiger Binärzähler mit dezimaler Decodierung. Angelegte Weiterschalttakte legen ein Prüfkriterium nacheinander an die vier Ausgänge des Zählers. Dies entspricht dem Durchprobieren der vier Verhaltensweisen, die dem Modell des LA möglich sind.

Jeweils ein Ausgang von V wird von außen markiert (Bedingung der Außenwelt). „Erfolg“ tritt ein, wenn das Prüfkriterium diesen Ausgang erreicht hat.

c) „erfolgreiche = gemerkte Verhaltensweise?“ — Merkspeicher M.

Jeder der vier möglichen Verhaltensweisen wird eine bistabile Kippstufe zugeordnet. Zu Beginn eines Lernvorgangs wird die erstmals erfolgreiche Verhaltensweise gemerkt, indem die zugehörige Kippstufe gesetzt wird. Steht im Verhaltensspeicher V die erfolgreiche Verhaltensweise und ist gleichzeitig die Kippstufe für diese Verhaltensweise gesetzt, so entspricht die erfolgreiche Verhaltensweise der gemerkten Verhaltensweise.

Die drei beschriebenen Speichergruppen E, V und M erscheinen im Blockschaltbild des LA, siehe Bild 2. Einer der m Eingänge (im Modell $m = 4$) wird von außen markiert („Bedingung der Außenwelt“). Auf den m Ausgängen erscheint das jeweilige Verhalten des LA im (1 aus m)-

Code. Die logische Schaltung L sorgt für das richtige Setzen und Abfragen der drei Speichergruppen und damit für den beschriebenen Ablauf des Lernprozesses. Ein funktionstüchtiges Demonstrationsmodell eines Lernenden Automaten nach den geschilderten Gesichtspunkten ist von LAMPERT am Institut für Nachrichtenvermittlung und Datenverarbeitung der Technischen Hochschule Stuttgart gebaut worden [3].

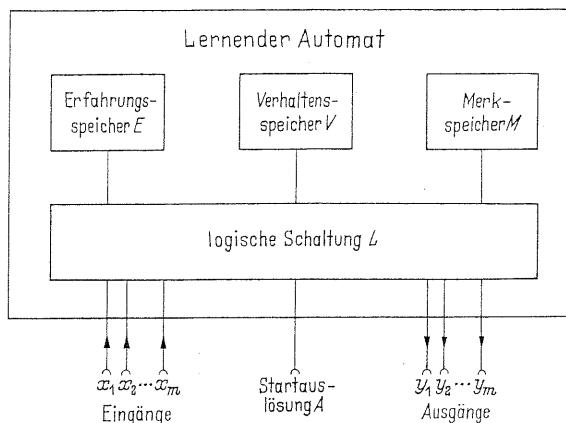


Bild 2. Blockschaltbild.

Anschauliche Darstellung für das Arbeiten des Lernenden Automaten

Für die vier verschiedenen Verhaltensweisen des LA sind in Bild 3 vier Ebenen E_1 bis E_4 gezeichnet, die alle längs der Zeitachse t zusammenstoßen. Auf jeder Ebene möge der Stand der Erfahrung bezüglich der zugeordneten erfolgreichen Verhaltensweise über der Zeit aufgetragen sein. Die Schwelle zwischen der Lern- und Kannphase ist mit a bezeichnet.

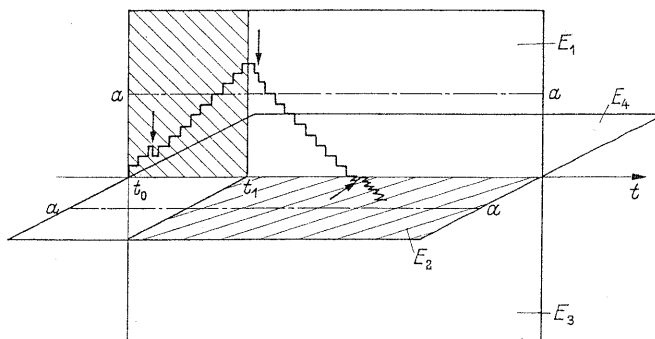


Bild 3. Ausschnitt aus einem Lernprozeß.

Es sei vereinbart, daß die Startauslösung des LA in gleichen Zeitabständen erfolgt. Angenommen die Verhaltensweise 1 führe zunächst zum Erfolg, was durch die Schraffur der Ebene E_1 für $t_0 \leq t < t_1$ angedeutet ist. Dann wird der Umfang der Erfahrung bezüglich dieser Verhaltensweise wie gezeichnet schrittweise zunehmen. Bei jeder Startauslösung wird die erfolgreiche Verhaltensweise durch Probieren gefunden, solange die Treppenkurve noch unterhalb der Schwelle a liegt (Lernphase), dagegen zeigt der LA die erfolgreiche Verhaltensweise sofort, wenn die Treppenkurve die Schwelle a überschritten hat (Kannphase) und keine Änderung in den Bedingungen der Außenwelt eingetreten ist. An allen durch einen Pfeil gekennzeichneten Stellen baut der im Modell gleichabständig eintreffende Vergessenstakt einen gerade bestehenden Erfahrungsschatz um einen Schritt ab.

Zum Zeitpunkt t_1 sollen sich die Bedingungen der Außenwelt sprunghaft so verändern, daß nunmehr die Verhaltensweise 2 des LA (Schraffur der Ebene E_2) zum Erfolg führt. Es muß

deshalb ein Umlernprozeß einsetzen. Oberhalb der Schwelle a stellt sich der LA bei jeder Startauslösung sofort auf die gelernte und bisher erfolgreiche Verhaltensweise 1 ein und registriert „Mißerfolg“. Dies hat zur Folge, daß die vorhandene überholte Erfahrung um einen Schritt abgebaut wird. Durch anschließendes Probieren ermittelt der LA sodann die jetzt erfolgreiche Verhaltensweise 2. Wird nach mehreren Mißerfolgen die Schwelle a der alten Erfahrung unterschritten, so wird die neue Verhaltensweise 2 durch sofort einsetzendes Probieren gefunden und die alte Erfahrung dabei schrittweise bis auf Null abgebaut. Jetzt erfolgt auf der Zeitachse der Übergang in die Ebene 2. Der Neulernprozeß bezüglich der Verhaltensweise 2 kann mit einem schrittweisen Erfahrungsaufbau in der Ebene E_2 einsetzen.

Zusammenfassung

Es wurde gezeigt, daß der Vorgang des „Lernens durch Erfolg“ einschließlich des Vergessens einer erfolgreichen Verhaltensweise oder des Umlernens grundsätzlich in Form eines Flußdiagramms darstellbar und mit einem Automaten realisierbar ist. Ein einfaches Modell eines Automaten ist gebaut worden und wird beschrieben.

Der Schwierigkeitsgrad eines zu automatisierenden Lernvorgangs (Anzahl der Variablen, von welchen das erfolgreiche Verhalten abhängt) wird lediglich durch den technisch und wirtschaftlich sinnvollen Aufwand für die Realisierung begrenzt.

Literatur

- [1] ZEMANEK, H.: Logische Beschreibung von Lernvorgängen, Lernende Automaten. München: R. Oldenbourg 1961.
- [2] STEINBUCH, K.: Automat und Mensch, 2. Aufl. Berlin, Göttingen, Heidelberg: Springer 1963
- [3] LAMPERT, E.: Realisierung eines Lernenden Automaten. Studienarbeit am Institut für Nachrichtenvermittlung und Datenverarbeitung der Technischen Hochschule Stuttgart, 1963.

Dis-Ing. MANFRED HUBER,
Institut für Nachrichtenvermittlung und Datenverarbeitung der Technischen Hochschule Stuttgart.