MEMORY MAP

mit Wandervorschlägen Teil 5

Auf unserer Reise durch den Speicherdschungel unserer beiden Commodore-Computer treffen wir diesmal auf die Speicherstellen 47 bis 56. Sie teilen uns sehr interessante und wichtige Informationen über Variablen-Felder mit.

Bevor wir unsere Wanderung beginnen, möchte ich Ihnen, lieber Leser, kurz in eigener Sache etwas erklären. Erst nachdem ich diesen fünften Teil des Kurses bereits geschrieben, aber noch nicht abgeliefert hatte, bekam ich die Ausgabe 1/85 des 64'er in die Hände ...und darin fand ich den ausgezeichneten Aufsatz von Boris Schneider über das Thema der »Garbage Collection« (Seite 122 ff). Vieles, was er beschreibt, beschreibe ich im 5. Teil dieser Serie auch. Und einige meiner Erklärungen von heute hätte ich mir auch sparen können, wenn ich den Aufsatz von Boris Schneider gekannt hätte.

Und genau darum geht es mir in meiner Erklärung: Die Autoren im 64'er wissen voneinander nichts, und es ist reiner Zufall, wenn dasselbe Thema von zwei Seiten gleichzeitig aufgegriffen wird. Nun, was soll man da machen? Die Redaktion jedenfalls hat mich ermutigt, dies ruhig geschehen zu lassen, da wichtige Themen, von verschiedenen Seiten betrachtet, sicher an Verständlichkeit gewinnen. Jetzt, wo Sie mir hoffentlich glauben, daß Duplizitäten nicht abgesprochen sind, will ich natürlich auf die hauptsächliche Duplizität zwischen Herrn Schneider und mir hinweisen. Wir stellen beide eine Methode vor, wie man die im Arbeitsspeicher stehenden Variablen und Zeichenketten sichtbar machen kann. Die Grundidee ist bei uns beiden gleich, nicht aber die Ausführung. Da diese Duplizität so schön demonstriert, daß dieselbe Idee in der Computerei oft auf verschiedene Weise gelöst werden kann, lasse ich die Erklärung meiner Methode in ihrer Ausführlichkeit bestehen.

In der letzten Folge haben wir die Bedeutung der Speicherzellen 45-46 für den Speicherbereich der Ganzzahl-, Gleitkomma- und Stringvariablen behandelt.

Zur »Sichtbarmachung« habe ich Ihnen ein Kochrezept gezeigt, mit dem wir die Variablen, so wie sie im Rechner stehen, auf dem Bildschirm anschauen können. Allerdings galt diese Methode nur für den C 64. Heute liefere ich Ihnen das entsprechende Kochrezept für den VC 20 nach (siehe Texteinschub 1).

Alle interessierten VC 20-Besitzer sollten sich jetzt noch einmal den letzten Teil des Kurses vornehmen und sich die normalen Variablen ansehen.

Heute kommen wir zu weiteren Zeigern im Speicherbereich 0 bis 1024, welche ebenfalls den Variablenspeicher beeinflussen.

Adresse 47-48 ($2F-$30)

Zeiger auf die Anfangsadresse des Speicherbereichs für Felder (Arrays)

Dieser Zeiger, in der Low/High-Byte-Darstellung, gibt dem Basic-Übersetzer (Interpreter) an, ab welcher Speicherzelle die Felder (Arrays) eines Basic-Programms gespeichert sind. Da die Felder direkt nach den normalen Variablen gespeichert werden, zeigt dieser Zeiger natürlich gleichzeitig auf das Ende des Speichers für normale Variablen.

Durch POKEn einer Adresse in die Speicherzellen 47-48 kann der Speicherbereich am Anfang eines Programms beinahe beliebig verschoben werden, beinahe deswegen, weil die Verschiebung im Zusammenhang mit den anderen Bereichen (siehe Bild 1) einen Sinn haben muß. Im übrigen gilt für diesen Zeiger dasselbe, was schon für den Zeiger in 45-46 gesagt worden ist. Die Darstellung der Feld-Variablen selbst kann mit den genannten Methoden angesehen werden, ihre Erklärung finden Sie im Texteinschub 2 und 3.

Wie aus den Erklärungen hervorgeht, wird bei Feldern mit Zeichenketten (Strings) in dem von Zeiger 47-48 bezeichneten Speicherbereich nur die Definition beziehungsweise die Dimensionierung gespeichert. Die eigentlichen Zeichenketten stehen wie bei den normalen Variablen im 4. Block, vom Speicherende rückwärts angeordnet.

Adresse 49-50 ($31-$32)

Zeiger auf die Endadresse (+1) des Speicherbereichs für Felder (Arrays)

Der Inhalt dieser Speicherzellen zeigt auf die Adresse, wo der Speicherbereich für Felder aufhört. Wie aus Bild 1 hervorgeht, werden die Zeichenketten vom Ende des verfügbaren RAM-Speichers rückwärts abgespeichert. Man kann also auch sagen, daß der Zeiger in 49-50 die letzte mögliche Adresse für Zeichenketten angibt. Wenn in einem Programm neue Variablen definiert werden, rutscht diese Adresse weiter nach oben und nähert sich dem Ende der Zeichenketten, die durch den Zeiger in 51-52 angegeben wird.

Wenn sich die Speicherbereiche der Felder und Zeichenketten berühren, bleibt der Computer stehen und führt die »Garbage Collection« (Müllabfuhr) durch — ein Prozeß, in dem nicht mehr gebrauchte Zeichenketten entfernt und der Zeichenketten-Speicher reduziert wird. Ist danach immer noch kein Platz, wird OUT OF MEMORY gegeben.

Der Befehl FRE löst immer eine solche Garbage Collection aus und gibt dann die Differenz zwischen den Adressen in den Zeigern 49-50 und 51-52 als verbleibenden noch verfügbaren Speicherbereich aus.

Adresse 51-52 ($33-$34)

Zeiger auf die untere Grenze des Speicherbereichs für den Text der Zeichenketten-Variablen

Der Inhalt dieser Speicherzellen zeigt in Low/High-Byte-Darstellung auf das jeweilige untere Ende (siehe Bild 1) des Textspeichers von Zeichenketten, er bezeichnet aber zugleich auch das obere Ende des frei verfügbaren RAM-Bereichs. Das entsteht dadurch, daß der Text der Zeichenketten vom Ende des RAM-Bereichs nach unten abgespeichert wird. In Bild 1 ist das durch den Pfeil dargestellt.

Beim Einschalten des Computers und nach einem RESET wird dieser Zeiger auf das oberste Ende des RAM-Bereichs gesetzt. Beim C 64 ist das 40960 ($A000). Beim VC 20 hängt es von den eingesetzten Speichererweiterungen ab, ohne Erweiterung ist die Adresse 7680 ($1E00).

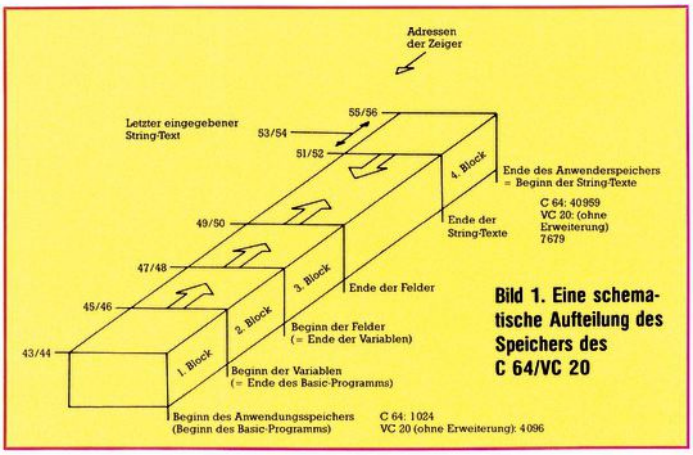

Bild 1. Eine schematische Aufteilung des Speichers des C 64/VC 20

Texteinschub 1

Darstellung der normalen Variablen beim VC 20

In der letzten Folge habe ich Ihnen gezeigt, wie die drei Typen der normalen Variablen auf dem Bildschirm sichtbar gemacht werden können, auch wenn kein Disassembler oder Monitor-Programm zur Verfügung steht. Die Erklärung galt aber nur für den C 64.

Heute ist die Methode für den VC 20 an der Reihe. Die Grundidee ist natürlich dieselbe wie beim C 64:

Wir legen den Bildschirmspeicher in den Speicherbereich, der für die Variablen reserviert ist, und schon sehen wir alle Variablen in der vom Computer verwendeten internen Darstellung.

Alle Angaben gelten für den VC 20 ohne Speichererweiterung, also ziehen Sie bitte alle Speichermodule heraus. Der Speicherbereich für Programme und deren Variablen beginnt jetzt ab Adresse 4096, das ist Block 1 im Bild 1. Der Bildschirmspeicher beginnt ab 7680. Wir verlegen jetzt den Bildschirmspeicher in den Block 1, so daß er ebenfalls ab Adresse 4096 beginnt. Danach müssen wir noch eine Farbe — am besten Schwarz — in den Farbspeicher POKEn, der in dieser neuen Konfiguration von 37888 bis 38399 liegt. Warum das so ist, erklärt Christoph Sauer in seinem Aufsatz »Der gläserne VC 20« Teil 4 im 64'er 1/85 Seite 131.

Das High-Byte der Adresse, in welcher der Bildschirmspeicher beginnt, steht in der Speicherzelle 648. Sie können das jederzeit mit PRINT PEEK(648) nachprüfen. Umgekehrt können wir eine Zahl hineinPOKEn, wodurch der Bildschirmspeicher verschoben wird. In unserem Fall erhalten wir das High-Byte für 4096 durch 4096/256 = 16.

Machen Sie jetzt bitte folgende Schritte:

1) direkt eingeben: POKE 648,16 (RETURN),
2) RUN/STOP und RESTORE drücken, bis der Cursor wieder da ist,
3) direkt eingeben:
FOR J=37888 TO 38399: POKE J,0: NEXT J (RETURN),
4) mit der DELETE-Taste (nicht mit CLR !) den ganzen Text des Bildschirms löschen,
5) mit dem Cursor etwa acht Zeilen nach unten gehen,
6) mit der Commodore- und SHIFT-Taste zusammen auf die Groß- und Kleinschrift umstellen,

Schritt 1 und 3 habe ich oben schon erklärt. Schritt 4 ist nicht absolut notwendig, aber ein leerer Bildschirm ist für uns besser. Die CLR-Taste würde Schritt 3 zunichte machen. Schritt 5 erlaubt uns, weiter unten auf dem Bildschirm Variablen einzugeben, ohne den oberen Teil vollzuschreiben. Schritt 6 schließlich erleichtert das Erkennen der Variablen-Darstellung.

Geben Sie jetzt bitte direkt ein:

```
VARIABLE = 3
```

und drücken Sie die RETURN-Taste. Was jetzt passiert und alle folgenden Erklärungen sind identisch mit dem Texteinschub für den C 64 in der letzten Folge. Ich verweise deshalb darauf, ohne sie zu wiederholen.

Der Befehl CLR setzt den Zeiger auf die Adresse, welche durch den Zeiger in den Speicherzellen 55-56 als das Ende des Basic-Speichers angegeben wird. Wozu das dient, erkläre ich Ihnen bei der Beschreibung dieses Zeigers weiter unten.

Adresse 53-54 ($35-$36)

Zeiger auf die Adresse der zuletzt eingegebenen Zeichenkette

In diesen Speicherplätzen steht die Adresse (im 4. Block, siehe Bild 1) der Zeichenkette, die als letzte von Routinen (Programme, Direkteingabe) zur String-Manipulation abgespeichert worden ist. Mit dem folgenden kleinen Programm können Sie das genau sehen:

```
10 PRINT PEEK(53)+256*PEEK(54),
20 PRINT PEEK(51)+256*PEEK(52)
30 INPUT A$
40 GOTO 10
```

Zeile 10 druckt uns zuerst (links) den Zeiger auf die zuletzt eingegebene Zeichenkette aus, Zeile 20 rechts daneben den Zeiger auf die untere Speichergrenze der Zeichenketten. Zeile 30 fordert zur Eingabe einer Zeichenkette auf.

Wenn Sie bei frisch eingeschaltetem Computer das Programm starten, sehen Sie eine 0 (= vorher noch kein String eingeben) und daneben die Adresse dez. 40960 (C 64) beziehungsweise dez. 7680 (VC 20 ohne Erweiterung). Wenn Sie auf das Fragezeichen des INPUT hin zum Bei-

Texteinschub 2

Darstellung der Felder-(Array)-Variablen

Die Felder-Variablen kommen in drei Arten vor:
— als ganze Zahlen,
— als Gleitkomma-Zahlen,
— als Zeichenketten,

Sie sind in dem Texteinschub 3 »Felder in Basic« kurz beschrieben.

Wir wollen sie uns hier mit den Methoden anschauen, welche ich das letzte Mal für den C 64 und heute für den VC 20 in den Texteinschüben »Darstellung der normalen Variablen« beschrieben habe.

Beim C 64 ist allerdings ein Zusatz dabei. Sie müssen, am besten gleich am Anfang, noch eingeben:

```
POKE 44,4:NEW
```

Ein eventuell auftretender SYNTAX ERROR soll uns nicht weiter stören.

Wenn Sie also das jeweilige Kochrezept ausgeführt und damit den Bildschirm- und den Variablenspeicher auf dieselbe Adresse gelegt haben, können wir anfangen.

Gleitkomma-Feld

Geben Sie direkt ein:

```
DIM AB(1,2,3)
```

Wir dimensionieren also ein Feld mit dem Namen AB, es hat drei Dimensionen, die erste Dimension hat zwei (0,1) Werte, die zweite hat drei und die dritte hat vier Werte. Sobald Sie die RETURN-Taste drücken, erscheint das Feld auf dem Bildschirm. Wir sehen:

```
... (plus 120 Klammeraffen @ ...)
```

Die ersten zwei Stellen sind der Name des Feldes in der Darstellung für Gleitkomma-Variable, wie in der letzten Folge beschrieben wurde. Die dritte und vierte Stelle geben im Bildschirmcode als Low- und High-Byte die Länge des Feldes an (das inverse c = 131, das @ = 0, bitte nachzählen). Die fünfte Stelle zeigt die Anzahl der Dimensionen (c = 3) an. Ab der sechsten Stelle stehen die Anzahl der Elemente der Dimension (diesmal als High- und Low-Byte) und zwar beginnend mit der letzten Dimension. In unserem Falle ist das also in Stelle 6 und 7 ein @ und d (0 — 3 = 4 = d), Stelle 8 und 9 sind dasselbe für die zweite Dimension und schließlich Stelle 10 und 11 für die erste Dimension (0 — 1 = 2 = b). Danach folgen entsprechend der Anzahl der dimensionierten Elemente (2x3x4 = 24) fünf Bytes pro Element (24x5 = 120), die vorerst auf 0 = @ stehen, die aber mit den Werten der Elemente aufgefüllt werden.

Dieses Auffüllen wollen wir nachvollziehen. Geben Sie bitte direkt ein:

```
AB(0,0,0)=5
```

Wir weisen damit dem allerersten Element des Feldes den Wert 5 zu.

In der oberen Darstellung des Feldes AB ändern sich dadurch Byte 12 und 13. Das neu erschienene inverse C und die Leerstelle mit den drei nachfolgenden @ ist die Gleitkommadarstellung (Mantisse und Exponent) der Zahl 5. Auf diese Darstellung werde ich später im Verlauf dieses Kurses bei der Besprechung der Speicherzelle 97 noch genauer eingehen.

Wenn wir jetzt (durch Überschreiben der vorigen Anweisung) zusätzlich noch eingeben:

```
AB(1,0,0)=6
```

erreichen wir eine entsprechende Änderung der Bytes 17 und 18, also des zweiten Elements des Feldes.

In Bild 2 sind die Stellen eines Gleitkomma-Feldes grafisch dargestellt.

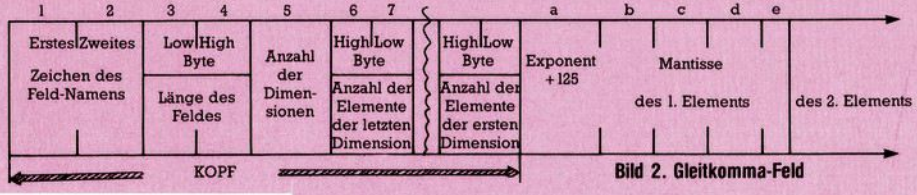

Bild 2. Gleitkomma-Feld

Ganzzahliges Feld

Im Vergleich zu dem Gleitkomma-Feld dimensionieren wir als nächstes ein ganzzahliges Feld:

```
DIM AB%(1,2,3)
```

Jetzt erscheint auf dem Bildschirm gleich anschließend an das erste Feld eine neue Darstellung:

... (plus 48 Klammeraffen @) ...

Die ersten elf Byte haben dieselbe Bedeutung wie beim Gleitkomma-Feld, aber nur deswegen, weil wir dieselben drei Dimensionen mit identischer Elementenzahl dimensioniert haben. Bei mehr Dimensionen wäre dieser Kopf natürlich länger. Die inverse Darstellung des Feldnamens signalisiert ein ganzzahliges Feld. Die dritte Stelle zeigt das »;« — im Bildschirmcode ist das die 59. In der Tat ist das Feld nur 59 Byte lang, also wesentlich weniger als das Gleitkomma-Feld. Die 2x3x4 = 24 Elemente benötigen in der Ganzzahl-Darstellung nur je zwei Byte (24x2 = 48 + 11 = 59). Womit bewiesen ist, daß eine Ganzzahl-Darstellung mit dem Zeichen % erheblich Speicherplatz spart — allerdings nur bei Feldern!

Jetzt wollen wir noch den Inhalt des Feldes füllen, so wie vorher mit:

AB%(0,0,0) = 5

... und prompt ändert sich Byte Nummer 13 in ein e (e = 5).

Eine Eingabe für das zweite Element:

AB%(1,0,0) = 6

verändert das 15. Byte in ein f.

In Bild 3 ist der Inhalt eines Ganzzahl-Feldes grafisch dargestellt.

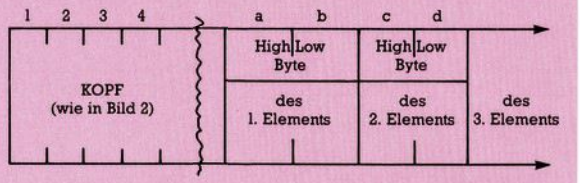

Bild 3. Ganzzahliges Feld

Felder mit Zeichenketten

Die Dimensionierung eines Feldes mit Zeichenketten sieht so aus:

DIM AB$(1,2,3)

Auf dem Bildschirm erscheint jetzt ein Feld in folgender Darstellung:

... (plus 72 Klammeraffen @) ...

Auch hier zeigen die ersten elf Stellen dieselbe Information wie bei den anderen Feldern. Zur Kennzeichnung des Zeichenketten-Feldes ist das zweite Zeichen des Feldnamens invers dargestellt. Zeichen 3 und 4 geben wieder die Länge des Feldes an. Das S hat den Bildschirmcode 83. (Vorsicht! Da wir im Groß-/Kleinbuchstaben-Modus sind, müssen wir die jeweils rechte Seite der Spalten in der Code-Tabelle nehmen). Die Länge 83 minus 11 Kopfstellen ergibt 72 Bytes, geteilt durch 24 (2x3x4 = 24 Elemente) erhalten wir 3 Byte zur Darstellung eines Elements.

Das erste Byte gibt die Länge der Zeichenkette an, das zweite und dritte Byte (Low/High-Byte) die Adresse, ab der die Zeichenkette im vierten Block (siehe Bild 1) gespeichert ist.

Die Methode ist also dieselbe wie bei den »normalen« Zeichenketten-Variablen. Das wollen wir uns auch noch ansehen. Geben Sie direkt ein:

AB$(0,0,0) = "AAAAAA"

In der Darstellung des Feldes ändern sich dadurch die Stellen 12, 13 und 14, und wir sehen

— beim C 64:

— beim VC 20:

Im Bildschirm steht dafür:

— C 64: 6 250 159 das heißt 6 Zeichen,
ab Adresse 250 + 159x256 = 40959

— VC 20: 6 250 29 das heißt 6 Zeichen
ab Adresse 250 + 29x256 = 7674

Jetzt weisen wir dem letzten Element auch noch eine Zeichenkette zu:

AB$(1,2,3) = "BB"

Die letzten drei Stellen des Feldes ändern sich ebenfalls, wobei die erste mit dem b eine Zeichenkettenlänge von 2 angibt, dementsprechend muß die Anfangsadresse um 2 niedriger sein als die vorher definierte Kette: Das Low-Byte 250 — 2 = 248, in der Codetabelle finden wir dafür das, was auch im Feld steht. Das High-Byte bleibt unverändert.

Bild 4 zeigt die grafische Darstellung des Zeichenketten-Feldes.

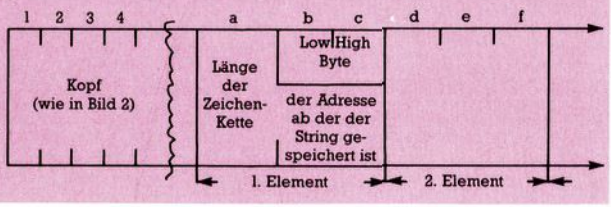

Bild 4. Zeichenketten-Feld

Als letztes zeige ich Ihnen noch die im vierten Block abgespeicherten Zeichenketten. Wir drucken einfach den CHR$-Wert der in den betreffenden Speicherzellen stehenden Codezahlen aus mit:

— VC 20:

```
FOR I=248 TO 255:PRINT CHR$(PEEK(29*256+I));:NEXT
```

— C 64:

```
FOR I=248 TO 255:PRINTCHR$(PEEK(159*256+I));:NEXT
```

... und wir erhalten die beiden Zeichenketten in umgekehrter Reihenfolge, also vom Speicherende her eingespeichert. Interessant ist, daß sich vor die Felder — wenn Sie sie noch auf dem Bildschirm hatten — die neu definierte Gleitkomma-Variable I@ geschoben hat. Auch das ist eine Demonstration des Speicherverfahrens der Variablen, genau so wie ich es Ihnen in der letzten Folge erklärt habe.

Texteinschub 3

Felder in Basic

Zur Wiederholung: Es gibt zwei Arten von Variablen, normale Variable und Felder. Jede der beiden Arten ihrerseits kann aus Gleitkomma-Zahlen, ganzen Zahlen oder Zeichenketten bestehen.

Eine normale Variable kann immer nur einen Wert haben, ein Feld enthält gleichzeitig viele Werte, alle unter demselben Variablen-Namen.

Wir können uns ein Feld mit dem Namen KARLSTRASSE als eine Liste vorstellen, in der jedes Element zwar auch den Namen Karlstraße hat, sich aber von den anderen Elementen durch eine eigene Hausnummer unterscheidet. Jede Variable in einer Hausnummer hat einen bestimmten Wert.

Während eine normale Variable einfach mit A = 3 einen Wert zugewiesen bekommt, muß ein Feld erst definiert werden, nämlich wie viele Elemente es enthält. Wir machen das mit dem Befehl

DIM KARLSTRASSE (12)

Dieses Feld hat 13 Elemente (von 0 bis 12). Jedem Element kann nun ein Variablenwert zugewiesen werden durch

KARLSTRASSE (0) = 25
KARLSTRASSE (1) = 56

Das Feld KARLSTRASSE hat in der Klammer nur eine Zahl, man sagt, es hat nur eine Dimension.

Ein zweidimensionales Feld entspricht einem Schachbrett, mit Zahlen in der einen und Buchstaben in der anderen Dimension. Wir definieren es mit:

DIM AX (7,7)

AX ist der Name, jede Dimension hat acht Elemente, insgesamt kann das Feld 64 Werte enthalten.

Ein dreidimensionales Feld entspricht einem Quader, oder bei gleicher Elementenzahl pro Dimension (Seite) einem Würfel. Dieses wird dimensioniert mit

DIM BY (125,6,2)

Die Anzahl der Dimensionen wird nur begrenzt durch den verfügbaren Speicherplatz. Wieviel Bytes pro Feld gebraucht werden, entnehmen Sie bitte der Erklärung bei der Darstellung der Feld-Variablen (Texteinschub Nummer 2).

Ein Feld, das wie bisher gezeigt dimensioniert wird, enthält Gleitkomma-Zahlen.

Ein Feld mit ganzen Zahlen wird durch das Zeichen % nach dem Namen gekennzeichnet, also:

DIM CZ%(..,..)

Ein Feld mit Zeichenketten dagegen hat nach dem Namen das übliche Zeichen $, also:

DIM DT$(..,..,..)

»Wozu brauche ich Felder, wenn ich auch normale Variable verwenden kann?«, werden Sie vielleicht noch fragen.

Felder haben den großen Vorteil, daß immer dann, wenn viele Variable in einem Programm vorkommen, die alle einen gewissen Zusammenhang haben, viel Speicherplatz gespart werden kann.

Eine normale Variable braucht sieben Byte, eine Feld-Variable nur fünf oder bei ganzen Zahlen sogar nur zwei Byte. Zugegeben, vorher steht noch ein längerer Kopf, aber halt nur einmal. Und das zahlt sich bei vielen Variablen sehr rasch aus.

Und schließlich muß ich noch darauf hinweisen, daß die »Hausnummern« oder Indizes der Elemente innerhalb eines Programms durch mathematische Operationen verändert und manipuliert werden können. Aber das ist natürlich höhere Programmierkunst und geht über diese kurze Einführung hinaus.

Memory Map mit Wandervorschlägen Teil 5

spiel ein A eintippen, erhalten Sie links den vorigen Wert von rechts und rechts jetzt eine um 1 kleinere Zahl. Eine weitere Eingabe von zum Beispiel XXXXX schiebt die alte rechte Zahl nach links und die neue wird um die Anzahl der Zeichen, also 5, verringert.

Adresse 55-56 ($37-$38)

Zeiger auf das Ende des für Basic-Programme verfügbaren Speichers

Dieser Zeiger, in der Low/High-Byte-Darstellung, gibt dem Basic-Übersetzer an, welches die höchste von Basic verwendbare Speicheradresse ist. Wie aus Bild 1 ersichtlich ist, ist diese Adresse zugleich der Anfang der als Variable abgespeicherten Zeichenkette (Strings).

Normalerweise ist diese Adresse fest vorgegeben. Die folgende Tabelle gibt darüber Auskunft:

Ende des Programmspeichers

	Adresse	Zeiger in 55 56
C 64	40960	0160
VC 20 (Grundv.)	7680	030
VC 20 (+3K)	7680	030
VC 20 (+8K)	16384	064
VC 20 (+16K)	24576	096
VC 20 (+24K)	32768	0128

Beim Einschalten des Computers überprüft das Betriebssystem den gesamten RAM-Speicher, bis es zur ersten ROM-Speicherzelle kommt, setzt den Zeiger in 55-56 auf diese Adresse und druckt den bekannten Kopf mit der verfügbaren Speicherangabe auf den Bildschirm.

Normalerweise wird dieser Zeiger nicht geändert.

Es gibt aber zwei Gelegenheiten, bei denen eine Änderung dieses Zeigers sinnvoll beziehungsweise notwendig ist.

Anwendung 1:

Es kommt oft vor, daß der gesamte Speicher nicht ausschließlich für Basic benötigt wird, sondern daß ein freier Speicherbereich geschaffen wird, um zum Beispiel Maschinenprogramme, selbst definierte Zeichen oder hochaufgelöste Grafik unterzubringen, die aber nicht vom Basic-Programm überschrieben werden können.

Bei der Besprechung der Zeiger in 43-44 haben wir das auch schon gemacht, allerdings durch »Hochschieben« des Speicheranfangs. Mit dem Zeiger in 55-56 erreichen wir denselben Effekt, diesmal durch »Hinunterdrücken« des Speicherendes. Gegenüber den vier Schritten beim Hochschieben ist das Hinunterdrücken einfacher. Mit dem Befehl:

POKE 56,PEEK(56)—1:CLR

schieben wir das Speicherende um 256 Byte nach unten, egal für welchen Computer und welche Speichererweiterung. Mit —2 verschiebt sich das Ende um 512, mit —4 um 1 024 Byte (also 1 KByte) nach unten. Wenn Sie eine feinere Verschiebung als Vielfache von 256 benötigen, kommen Sie mit dem High-Byte in 56 allein nicht aus, sondern Sie müssen auch einen entsprechenden Wert in 55 hineinPOKEn.

Der Befehl CLR ist absolut notwendig, denn er setzt den Zeiger der Zellen 51-52 (siehe dort), das heißt das untere Ende des Speicherbereichs für Zeichenketten auf dieselbe Adresse wie Zeiger 55-56. Dadurch wird erzwungen, daß die Zeichenkette sozusagen als Ausgangslage unterhalb des heruntergedrückten Speicherendes abgelegt werden.

Anwendung 2:

Über den User-Port (Steckerleiste an der Rückseite, neben dem Datasetten-Anschluß) können VC 20 und C 64 mit anderen Geräten verbunden werden. Der Datentransfer über diese Verbindung — sie heißt RS232-Schnittstelle — muß allerdings programmiert werden. Diese RS232-Schnittstelle hat die Gerätenummer 2 (so wie der Drucker Nummer 4 und das Diskettengerät die Nummer 8 hat).

Wenn nun ein Gerät Nummer 2 mit einem OPEN-Befehl angewählt wird, wird automatisch der Zeiger in 55-56 und der Zeiger in 643 um 512 Bytes heruntergedrückt, um je einen Eingangs- und Ausgangspufferspeicher zu erzeugen. Da der Inhalt dieser Pufferspeicher alle Variable in diesen 512 Bytes überschreiben würde, wird auch der CLR-Befehl automatisch gegeben.

Es gilt daher als Vorschrift, daß bei RS232-Verbindungen zuerst der Datenkanal durch OPEN eröffnet werden muß, bevor Variable, Felder und Zeichenketten definiert werden.

Dieser Zeiger beschließt die Gruppe der Speicherzeiger. Das nächste Mal machen wir ab Adresse 57 weiter.

(Dr. Helmuth Hauck/gk)

Finden mit System — Eine neuartige Suchmethode (Teil 3)

Nachdem wir uns in den letzten beiden Folgen mit grundlegenden programmiertechnischen Fragen befaßt haben, sollen jetzt konkrete Anwendungen im Mittelpunkt stehen. Diesmal ist es das Suchen von Zeichenketten mit einer erst vor kurzen entwickelten Methode.

Wie versprochen, stelle ich Ihnen diesmal ein Suchprogramm vor, mit dem eine schnelle »intelligente« Suche von Strings möglich wird. Doch bevor wir uns mit der Praxis, sprich dem Programm, beschäftigen können, müssen wir uns ein wenig mit der Theorie des Suchens befassen.

Wie sucht man eigentlich?

Nun, uns soll hier nicht interessieren, wie der eine oder andere mit seiner häuslichen (Un-) Ordnung fertig wird. Uns geht es hier um das Suchen und möglichst auch Finden von einer Zeichenkette in einer anderen. Einigen wir uns jetzt schon auf zwei Begriffe, damit wir später keine Schwierigkeiten bekommen. Der Suchstring ist der String, nach dem wir suchen. Dies kann eine beliebige Zeichenkette, also ein Wort, eine Zahl, Grafikzeichen oder ein Gemisch aus alledem sein. Der Durchsuchstring ist derjenige, in dem wir suchen.

Das landläufige Suchverfahren sieht folgendermaßen aus: Der Anfang des Suchstring wird an den Anfang des Durchsuchstrings angelegt. Dann werden die beiden ersten Buchstaben verglichen. Sind diese gleich, werden die beiden zweiten Buchstaben verglichen, und so weiter. Dieses Spielchen wiederholt sich so lange, bis sich die beiden Buchstaben nicht gleichen. Dann wird der Suchstring um eine Position am Durchsuchstring verschoben, und somit der erste Buchstabe des Suchstrings mit dem zweiten des Durchsuchstrings und so weiter, verglichen. Wird beim Suchen festgestellt, daß alle Buchstaben übereinstimmen, so hat man den Suchstring im Durchsuchstring gefunden. Stellt man aber fest, daß das Ende des Suchstrings über den Durchsuchstring hinausragt, so muß man die Suche abbrechen, da der Suchstring nicht im Durchsuchstring vorhanden sein kann.

Dieses Suchverfahren war jahrelang das einzige verwendete, da es relativ einfach zu programmieren und relativ schnell war.

Es geht auch anders

In der November-Ausgabe der Zeitschrift Spektrum der Wissenschaft wurde nun ein völlig neuer Suchalgorithmus vorgestellt. Dieser ist erst vor kurzem von zwei amerikanischen Forschern entwickelt worden, und kann mit noch vertretbarem Aufwand in Maschinensprache formuliert werden. Noch einmal das Prinzip des alten Algorithmus: Man verschiebt den Suchstring am Durchsuchstring immer nur um eine einzige Position, braucht also mindestens so viele Vergleiche, wie Zeichen im Suchstring minus Zeichen im Durchsuchstring (bei Überhang wird ja abgebrochen). Findet man einen Weg, den Suchstring immer gleich um mehrere Positionen zu verschieben, ohne daß dabei ein Auftauchen des Suchstrings im Durchsuchstring übersehen werden kann, läßt sich der Suchvorgang rapide beschleunigen. Einfach um einen größeren konstanten Faktor zu verschieben, etwa um 2, klappt natürlich nicht, weil dann nur alle ungeraden Positionen abgefragt werden. Beginnt der Suchstring aber auf einer geraden Position, wird er nicht gefunden. Dieser Verschiebefaktor muß also variabel sein. Um ihn zu ermitteln, stellen wir mal fest, wie weit in einigen Fällen maximal verschoben werden darf, um den Suchstring nicht zu verpassen.

Im folgenden werden wir immer nach dem Wort »HALLO«

Fortsetzung auf Seite 151