



Äußerst fruchtbar

Selbstreproduzierende Programme

Hartwig Thomas

Ehrlos sei, wer Schlechtes dabei denkt. In diesem Beitrag geht es um eine recht bekannte Übungsaufgabe, deren Lösung selten in der Literatur auftaucht. Für eine laue Stunde, in der man gerade einmal nicht mit real existierenden Softwareproblemen kämpft (und auch sonst nichts Besseres zu tun hat), ist ein kleines Denkspiel auf jeden Fall bestens geeignet.

Gesucht ist ein Programm, das sich selber im Quellcode ausgibt. Das Programm soll dazu als ASCII-File vorliegen, dessen Kompilat (Turbo-Pascal >= 4.0) genau den Dateiinhalt der Quelle auf den Bildschirm schreibt. Die Aufgabe gilt als gelöst, wenn zwei identische Dateien entstehen, sofern man die Bildschirmausgabe in eine zweite Datei umleitet. Selbstverständlich gilt jede

'billige Lösung' als böses Foul (zum Beispiel Quelldatei als vorhanden annehmen, sie einlesen und wieder ausgeben).

Bevor Sie zur Lösung weiterblättern oder die Überlegungen dazu lesen, dürfen Sie ruhig einmal ein paar Minuten Bedenkzeit einlegen und einen eigenen Weg suchen. Um Ihnen die Entscheidung leichter zu machen, habe ich

die Lösung geringfügig verschlüsselt.

Denkpause

Nun gut. Die gestellte Aufgabe verwirrt anfangs hauptsächlich durch ihre Unbestimmtheit. Sie scheint nicht wesentlich von der gewählten Programmiersprache abzuhängen. Gesucht ist also erst einmal eine abstrakte Programmiersprache, in der die Aufgabe garantiert lösbar ist. Eine solche Sprache muß mindestens einen Befehl kennen, der ein Argument auf dem Bildschirm ausgibt. Wenn alle Ausgabebefehle Ihre Argumente nur höchstens einmal ausgeben, ist die Aufgabe offensichtlich nicht lösbar.

Die minimale Programmiersprache, die diesem Anspruch genügt, besitzt einen einzigen Befehl, der das nachfolgende Zeichen *zweimal* auf den Bildschirm schreibt. Der Quellcode für diesen Befehl sei A; eine weitere Vorgabe, daß die Programmiersprache keine Garnierungen für Beginn und Ende verlangt. In dieser Sprache lautet die Lösung:

AA

Offenbar schreibt dieses Programm seinen eigenen Quelltext auf den Bildschirm, da der Befehl A das darauffolgende A zweimal ausgibt.

Konkretisierung

In Pascal und ähnlichen Sprachen muß man einerseits Bezeichner als Zeichenketten und andererseits am Anfang und Ende eine Minimalgarnierung in Kauf nehmen. Um dieser Tatsache Rechnung zu tragen, modifiziere ich die obige Programmiersprache etwas. Sie hat jetzt folgende Syntax:

'A' betrachtet die zwei folgenden Symbole als Argumente;

'B' hat keine Argumente und die Semantik:

'A' startet das Programm und gibt jedes der beiden Argumente zweimal auf dem Bildschirm aus;

'B' beendet das Programm ordnungsgemäß.

Die Problemlösung in dieser erweiterten Sprache lautet

AABB.

Nach dieser Grundidee entwickelt man dann das Pascal-Programm, das die Aufgabe löst. Die Probe führt eine Batch-Datei durch: Sie kompiliert das Original, führt es mit Redirection aus, kompiliert das Ergebnis, startet es mit erneuter Ausgabeumleitung und vergleicht die beiden ASCII-Files (siehe Kasten). Der Zusatzschritt mit doppelter Kompilation ist übrigens nur nötig, wenn der verwendete Editor unerwünschte Returns, Linefeeds oder Ctrl-Zs einfügt oder anhängt.

Diskussion

Das Lösungsprogramm lehnt sich recht nah an das Schema AABB an. Damit keine Schwierigkeiten mit Zeilenenden innerhalb der Argumente auftreten, fehlt der in Standard-Pascal nötige Header 'program(input,output);', wie es in Turbo-Pascal ja erlaubt ist. Daher paßt auch das ganze A auf eine Zeile (die für den Compiler höchstens 129 Zeichen lang sein darf), was eine zusätzliche Aufblähung durch Mehrfachzeilenbehandlung einspart.

Trotz ihrer Plausibilität und einem Anschein der Natürlichkeit besteht die oben angeführte Herleitung den letzten Test auf Perfektion nicht. Die Notwendigkeit der hergeleiteten Struktur der Lösung ist nicht gegeben. Die Frage, ob ein selbstreproduzierendes Programm am doppelten Auftauchen seiner Teile zu erkennen ist, bleibt offen.

Mittel ohne Zweck

Die kleine Programmieraufgabe ist sicherlich nur Spielerei. Die Eigenschaft der Selbstreproduktion hat unter dem Namen 'Computerviren' erst einmal einen schlechten Ruf erhalten. Im anbrechenden Zeitalter der Computernetze schüren zum Teil dieselben EDV-Manager die Angst vor dem Virus, die vor nicht allzu langer Zeit auf ihrem Time-Sharing-Mainframe alle Benutzern ihre Macht unterwarfen und ihnen ihre Aufsicht zumuteten. Der PC ohne Festplatte ist die angepriesene Lösung, und Daten-

