UMach Spezifikation

24. Juli 2012

Inhaltsverzeichnis

Tabellenverzeichnis

Abbildungsverzeichnis

1 Einführung

UMach ist eine einfache programmierbare virtuelle Maschine (VM), die einen definierten Instruktionssatz und eine definierte Architektur hat. UMach orientiert sich dabei an Prinzipien von RISC Architekturen: feste Instruktionslänge, kleine Anzahl von einfachen Befehlen, Speicherzugriff durch Load- und Store-Befehlen, usw. Die UMach Maschine ist Register-basiert. Der genaue Aufbau dieser Rechenmaschine ist im Abschnitt 2.1 ab der Seite 8 beschrieben.

Für den Anwender der virtuellen Maschine wird zuerst eine Assembler-Sprache zur Verfügung gestellt. In dieser Sprache werden Programme geschrieben und anschließend kompiliert. Die kompilierte Datei (Maschinen-Code) wird von der virtuellen Maschine ausgeführt.

Obwohl in diesem Dokument Namen von Assembler-Befehlen angegeben werden (siehe Kapitel 3, Instruktionssatz) spezifiziert dieses Dokument die UMach Maschine auf Maschinencode Ebene (Register, Bussystem, Instruktionen). Die Implementierung eines Assemblers ist frei, zusätzliche Befehle, Instruktionsformate, Aliase und sprachliche Konstrukte auf der Assembler-Ebene zu definieren. Z.B. auf Maschinencode-Ebene, sind die Befehle

```
ADD R1 R2 5
SUB R2 4
```

fehlerhaft, denn das Format der `ADD` und `SUB` Befehle verlangt die Angabe von drei Registern. Der Assembler ist frei, diese zusätzliche Formate zu definieren und zu erkennen, solange er gültigen UMach Maschinencode produziert. Gültiger Maschinencode für die obigen Befehle wäre

```
ADDI R1 R2 5 # Maschinencode 0x32 0x01 0x02 0x05
SUBI R2 R2 4 # Maschinencode 0x35 0x02 0x02 0x04
```

2 Organisation der UMach VM

2.1 Aufbau

Die virtuelle UMach Maschine hat eine einfache Konstruktion, die im wesentlichen aus den folgenden Komponenten besteht:

1. Kern (Abschnitt 2.2)
2. Speicher (2.4)
3. IO Einheit (Abschnitt 2.5)

Siehe auch die Abbildung 2.1 auf der Seite 9, die die UMach Maschine darstellt.

Wenn die Maschine startet, sucht sie nach einem Program im Speicher, holt jede Instruktion nach einander in den Kern und führt sie aus. Danach bleibt sie im Wartezustand bis sie ausdrücklich ausgeschaltet wird. Alle Eingabe- und Ausgabeoperationen werden an die IO-Einheit weitergeleitet (siehe auch die Abschnitte 2.5 und 3.11).

2.1.1 Betriebsmodi

Ein Betriebsmodus bezieht sich auf die Art, wie die UMach VM läuft. Die UMach VM kann in einem der folgenden Betriebsmodi laufen:

1. Normalmodus
2. Einzelschrittmodus

Der standard-Modus ist der Normalmodus. Der Modus wird vor dem Hochfahren der Maschine festgelegt.

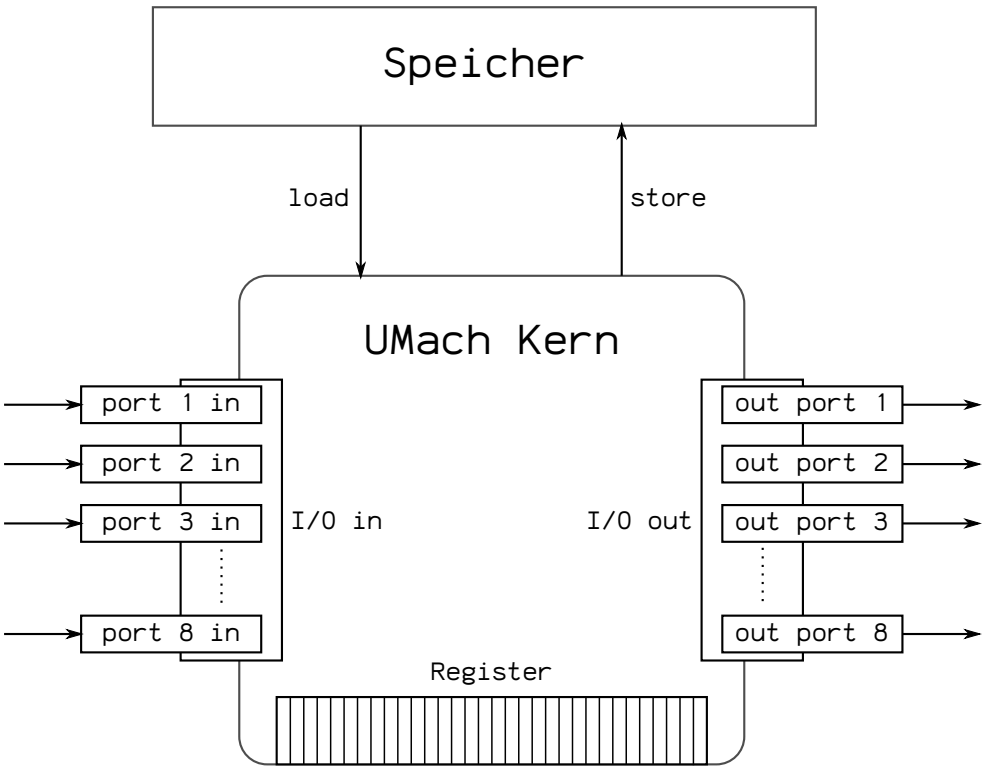

Abbildung 2.1: Aufbau der UMach Maschine

Normalmodus Die virtuelle Maschine führt ohne Unterbrechung ein Programm aus. Nach der Ausführung befindet sich die Maschine in einem Wartezustand, falls sie nicht ausdrücklich ausgeschaltet wird.

Einzelschrittmodus Die virtuelle Maschine führt immer eine einzige Instruktion aus und nach der Ausführung wartet sie auf einen externen Signal um mit der nächsten Instruktion fortzufahren. Dieser Modus soll dem Entwickler erlauben, ein Programm schrittweise zu debuggen.

2.1.2 Von Neumann Zyklus

Der Von Neumann Zyklus der UMach ist auf 4 Schritte verkürzt. Dies ist möglich, da die Instruktionsbreite immer aus 4 Wörtern besteht, und somit der „FETCH“ von Befehl und zugehörigen Operanden in einem gemeinsamen Schritt durchgeführt werden kann. Somit besteht der Zyklus aus einem beginnenden FETCH. Bei diesem wird der an der im Programmcounter `PC` liegende Adresse gespeicherte Befehl in das Instruktionsregister `IR` geladen. Danach wird der im ersten Byte liegende Befehl mit Hilfe des Befehlsdecoders

decodiert und an der ALU entsprechend eingestellt.

In einem dritten Schritt EXECUTE wird die Instruktion ausgeführt. Der theoretisch 4. Schritt, das Inkrementieren des Programmcounters PC, wird parallel zu den FETCH und DECODE Vorgängen ausgeführt. Diese Tatsache ist bei Instruktionen, welche den Inhalt des PC manipulieren, zu berücksichtigen. Siehe auch Abbildung 2.2 auf der Seite 10.

Auflistung der Schritte:

1. FETCH – Holen der Instruktion aus dem Speicher von der Adresse, welche im PC vorliegt.

2. DECODE – Decodieren des Befehles und Einstellen der ALU.

3. EXECUTE – Auführen des Befehles in der ALU.

4. Update `PC` – Inkrementieren des `PC`. Findet parallel zu FETCH und DECODE statt.

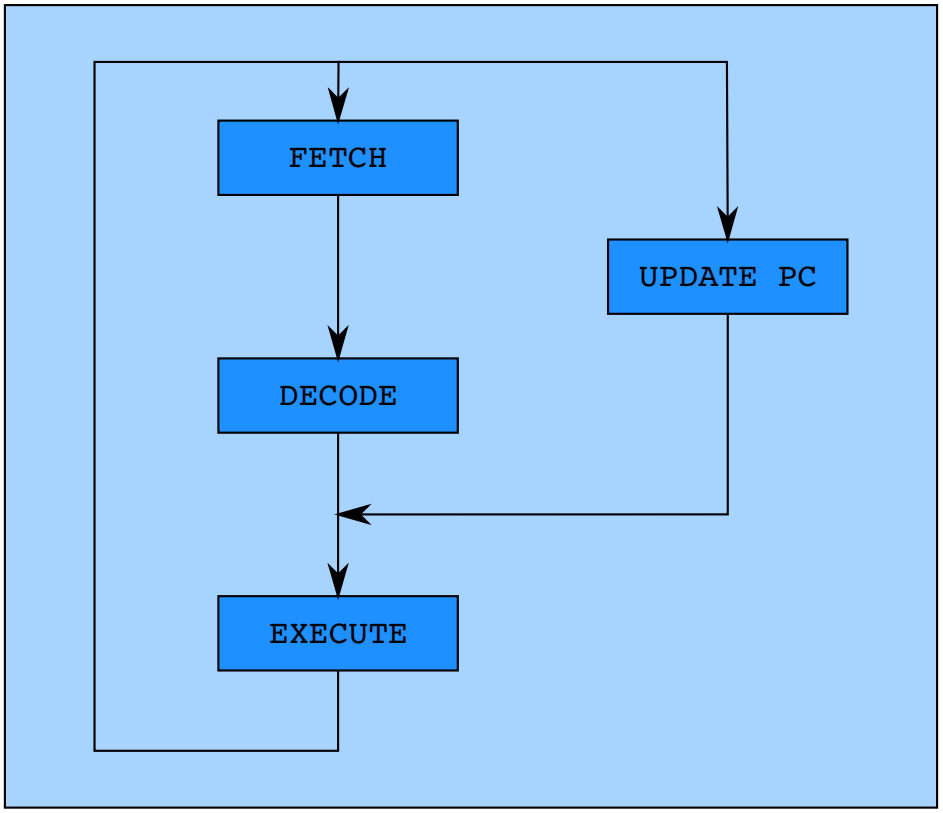

Abbildung 2.2: Von Neumann Zyklus

2.2 Kern

Der Kern der Maschine besteht aus den folgenden Komponenten:

1. Befehlsabruf Einheit (Instruction Fetch Unit), die die Programminstruktionen der Reihe nach aus dem Speicher holt und der Decodierungseinheit übergibt.

2. Decodierungseinheit, die dafür zuständig ist, eine Instruktion zu decodieren, bzw. in ihren Komponenten zu zerteilen.

3. Recheneinheit, die für die tatsächliche Ausführung der Instruktionen zuständig ist.

4. Register, kleine Speichereinheiten, die sich im Kern befinden.

2.3 Register

Die Register sind die Speichereinheiten im Prozessor. Die meisten Anweisungen an die UMach Maschine operieren auf einer Art mit den Registern.

Für alle Register gilt:

1. Jedes Register ist ein Element aus der Menge $\mathcal{R}$, die alle Register beinhaltet. Die Notation $x \in \mathcal{R}$ bedeutet, dass x ein Register ist.

2. Die Speicherkapazität beträgt 32 Bit.

3. Es gibt eine eindeutige natürliche Nummer, die innerhalb der Maschine das Register identifiziert. Diese Nummer heißt Registernummer und wird von einer Instruktion auf Maschinencode-Ebene verwendet, wenn sie das Register anspricht – mit anderen Worten, auf Maschinecode-Ebene wird das Register durch seine Nummer angegeben.

4. Die UMach Maschine erwartet die Angabe eines Registers als numerischer Wert (Maschinenname). Jedoch verwendet der Programmierer der Maschine auf Assembler Ebene einen eindeutigen Namen dieses Registers. Dieser Name heißt Assemblername.

Die UMach Maschine hat zwei Gruppen von Registern: die Allzweckregister und die Spezialregister.

2.3.1 Allzweckregister

Es gibt 32 Allzweckregister, die dem Programmierer für allgemeine Zwecke zur Verfügung stehen. Diese 32 Register werden beim Hochfahren der Maschine mit dem Wert Null (`0x00`) initialisiert. Außer dieser Initialisierung, verändert die Maschine den Inhalt der Allzweckregister nur auf explizite Anfrage, bzw. infolge einer Instruktion.

Die 32 Register werden auf Maschinencode-Ebene von 1 bis 32 nummeriert (`0x01` bis einschliesslich `0x20` im Hexadezimalsystem). Diese Nummer ist die Registernummer des Registers. Auf Assembler-Ebene, werden sie mit den Namen $R1, R2 \ldots$ bis $R32$ angesprochen (Assemblername). Die Zahl nach dem Buchstaben R ist im Dezimalsystem angegeben und ist fester Bestandteil des Registernamens.

Assemblername	`R1`	`R2`	`R3`	`...`	`R32`
Registernummer	`0x01`	`0x01`	`0x03`	`...`	`0x20`

Das Register mit Nummer Null, oder das Null-Register, ist ein Spezialregister, dass immer den Wert Null hat und nicht beschreibbar ist . Es wird mit dem Namen `ZERO` angesprochen (siehe auch den Abschnitt 2.3.2 auf Seite 12).

Beispiel für die Verwendung von Registernamen:

Assembler Maschinencode	`ADD` `0x50`	`R1` `0x01`	`R2` `0x02`	`R3` `0x03`
Bytes	erstes Byte	zweites Byte	drittes Byte	viertes Byte
Algebraisch	$R_1 \leftarrow R_2 + R_3$			

2.3.2 Spezialregister

Die Spezialregister werden von der UMach Maschine für spezielle Zwecke verwendet, sind aber dem Programmierer sichtbar. Der Inhalt der Spezialregister kann von der Maschine während der Ausführung eines Programms ohne Einfluss seitens Programmierers verändert werden.

Nicht alle Spezialregister können durch Instruktionen überschrieben werden (sind schreibgeschützt).

Die Registernummern der Spezialregister setzen die Nummerierung der Allzweckregister zwar fort, die Assemblernamen aber nicht: es gibt kein Register $R33$. Die Tabelle 2.1 auf

Seite 13 enthält die Liste aller Spezialregister. In der ersten Spalte steht der Assemblername, so wie er vom Programmierer verwendet wird. In der zweiten Spalte steht der Maschinenname im Dezimalsystem, so wie er im Maschinencode steht. Die dritte Spalte enthält eine kurze Beschreibung und Bemerkungen. Falls die Beschreibung nicht anders spezifiziert, ist das Register nicht schreibgeschützt.

Tabelle 2.1: Liste der Spezialregister

`PC`	33	„Instruction Pointer“, oder „Program Counter“. Enthält zu jeder Zeit die Adresse der nächsten Instruktion. Wird auf Null gesetzt, wenn die Maschine hochfährt. Wird nach dem Abfangen einer Instruktion in das Register `IR` automatisch inkrementiert. Schreibgeschützt.
`SP`	34	„Stack Pointer“. Enthält die Speicheradresse des höchsten Eintrag auf dem Stack. Wird beim Hochfahren der Maschine auf die maximal erreichbare Speicheradresse plus 1 gesetzt. Die 1, die zur maximalen Adresse addiert wird, setzt den stack pointer auf eine ungültige Adresse und sorgt dafür, dass keine Werte mittels `POP` gelesen werden, bevor Werte auf den Stack gepusht wurden. Siehe auch 2.4.2, Seite 17.
`FP`	35	„Frame Pointer“. Enthält die Startadresse des Stack Frames einer Subroutine und unterstützt die Implementierung von Funktionen.
`IR`	36	„Instruction Register“. Enthält die gerade ausgeführte Instruktion. Schreibgeschützt.
`STAT`	37	Enthält Status-Informationen. Diese Informationen sind in den einzelnen Bits dieses Register gespeichert. Welche Informationen vorhanden sind liegt an der jeweiligen Anwendung.
`ERR`	38	„Error“. Fehlerregister. Die einzelnen Bits dises Registers geben Auskunft über Fehler, die mit der Ausführung des Programms verbunden sind. Liegt kein Fehler vor, so ist der Inhalt dieses Registers gleich Null. Ist ein bestimmtes Bit gesetzt, so wird dadurch der entsprechende Fehler signalisiert. Für eine Liste der verwendeten Bits un deren Bedeutung, siehe Tabelle 2.2 auf der Seite 14.
`HI`	39	„High“. Falls eine Multiplikation durchgeführt wird, enthält dieses Register die höchstwertigen 32 Bits des Ergebnisses der Multiplikation und bildet zusammen mit dem Register `LO` das volle Ergebnis der Multiplikation. Falls eine Division durchgeführt wird, enthält dieses Register den Quotient der Division.
`LO`	40	„Low“. Falls eine Multiplikation durchgeführt wird, enthält dieses Register die niedrigstwertigen 32 Bits des Ergebnisses der Multiplikation und bildet zusammen mit dem Register `HI` das volle Ergebnis der Multiplikation. Falls eine Division durchgeführt wird, enthält dieses Register den Rest der Division.

CMPR	41	„Comparison Result“. Enthält das Ergebnis eines Vergleichs. Siehe auch Abschnitt 3.7, Seite 48.
HIR	42	„Hardware Interrupt“. Enthält die Nummer des auslösenden Ports einer Unterbrechung. Siehe auch Abschnitt 2.6.1, Seite 19.
ZERO	00	Enthält die Zahl Null. Schreibgeschützt.

Das ERR Register

Die Tabelle 2.2 listet alle Fehler auf, die in dem ERR Register signalisiert werden können. Zu jedem Fehler gehört ein Bit im Register. Die Bitstellen werden dabei entsprechend deren Stelligkeiten durchnummeriert: Bit mit Stelligkeit 2^0 hat Position 0, Bit mit Stelligkeit 2^{31} hat die Position 31.

Tabelle 2.2: Bedeutung der einzelnen Bits im ERR Register

0	Division durch Null
	Ungültige Registernummer
	Stack Fehler
	Ungültige Speicheradresse
	Ungültige Befehlsargumenten
	Ungültiger Befehl
	I/O Port existiert nicht

2.4 Der Speicher

2.4.1 Adressierungsarten

Als RISC-orientierte Maschine, greift die UMach lediglich in zwei Situationen auf den Speicher zu: zum Schreiben von Registerinhalten in den Speicher (Schreibzugriff) und zum Lesen von Speicherinhalten in einen Register (Lesezugriff). Die Adressierungsart beschreibt dabei, wie der Zugriff auf den Speicher erfolgen sollte, bzw. wie die angesprochene Speicheradresse angegeben wird. Anders ausgedrückt, beantwortet die Adressierungsart die Frage „wie kann eine Instruktion der Maschine eine Adresse angeben?“.

Die UMach Maschine kennt eine einzige Adressierungsart: die (register-) indirekte Adressierung. Dabei werden alle Speicheradressen nicht direkt angegeben, sondern es wird ein Register angegeben, der die Adresse beinhaltet. Zum Beispiel, die Instruktion

```
0x13 0x01 0x02 0x00
```

bedeutet nicht „lade das Wort an der Adresse `0x02` in das Register mit Nummer `0x01`“, sondern „lade das Wort, das an der Adresse liegt, die in Register mit Nummer `0x02` angegeben ist, in das Register mit Nummer `0x01`“. Das entpricht der Instruktion

```
LW R1 R2
```

Analog gilt für Schreibbefehle.

2.4.2 Speicherstruktur

Der Speicher der UMach Maschine hat zur Laufzeit eine bestimmte Struktur, bzw. wird in bestimmten Bereichen unterteilt. Diese Bereiche sind

1. Unterbrechungstabelle
2. Programm und Daten
3. Stack

Siehe auch die Abbildung 2.3 auf der Seite 15.

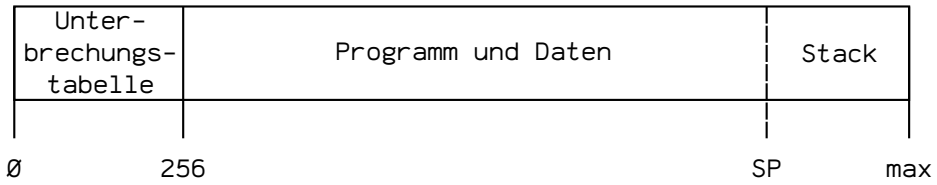

Abbildung 2.3: Speicherstruktur zur Laufzeit

Unterbrechungstabelle

Die Unterbrechungstabelle besteht aus einer Reihe von Sprungadressen zum ausführbaren Code, bzw. zu sogenannten Unterbrechungsroutinen, oder „interrupt handlers“, die in Ausnahmefällen ausgeführt werden sollen. Jedes mal, wenn eine Ausnahmesituation auftritt (meistens eine Fehlersituation), wird intern ein Unterbrechungssignal erzeugt, das mit

einer Kennnummer (Unterbrechungsnummer) versehen ist. Die Unterbrechungsnummer wird als Index in dieser Tabelle verwendet.

Hier ist die Rede von Unterbrechungen, die direkt mit der Ausführung von Instruktionen verbunden sind (Hardware Interrupts). Sie können entweder von der Maschine selbst erzeugt werden, wie z.B. im Falle einer Division durch Null, oder im Falle, dass eine unerlaubte Speicheradresse verwendet wird, oder sie können unter Programmkontrolle erzeugt werden. Siehe auch den Abschnitt 2.6, ab der Seite 19.

Die Unterbrechungstabelle wird von der Maschine nicht gefüllt, sondern es ist Sache der Software, die entsprechenden Stellen im Speicher zu füllen und entsprechende Funktionalität zur Verfügung zu stellen. Wird eine solche Adresse nicht gesetzt, d.h. ist der entsprechende Tabelleneintrag auf Null gesetzt, so reagiert die Maschine auf die Ausnahmesituation mit seiner Standardfunktion: die Maschine hält an. Für weitere Informationen bzgl. der Fehlerbehandlung siehe den Abschnitt 2.6, ab der Seite 19.

Die Unterbrechungstabelle besteht aus 64 Einträgen, wobei der Eintrag mit der Nummer Null reserviert ist. Die folgenden acht (Eintrag 1 bis 8) dienen für die Hardware-Unterbrechungen, die von den Eingabeports erzeugt werden. Nicht alle Einträge müssen belegt sein.

Jeder Eintrag beträgt wie ein Register 32 Bit, oder 4 Byte. Da es 64 Einträge gibt, ist die Unterbrechungstabelle $64 \cdot 4 = 256$ Bytes groß. Siehe auch die Abbildung 2.3 auf der Seite 15.

Die Tabelle 2.3 gibt alle Unterbrechungsnummer an und deren Bedeutung. Undefinierte Unterbrechungsnummer werden nicht explizit angegeben.

Tabelle 2.3: Unterbrechungsnummer

`0`	Reserviert
`1-8`	Hardware Interrupts
`16`	Division durch Null
`99`	Noch nicht fertig

Programm und Daten

Nach der Unterbrechungstabelle, die 256 Bytes groß ist (64 mal 4), folgt das eigentliche Program, das von der Maschine ausgeführt werden soll. Dieses Programm wird also ab der Adresse 256 gelesen und ausgeführt.

Insbesondere ist dieses Program dafür zuständig, die Unterbrechungstabelle zu füllen, falls (bestimmte) Unterbrechungen behandelt werden sollten.

Der Stack

Der Stack ist ein spezieller Bereich im Speicher. Dieser Bereich fängt am Ende des Speichers mit der größten Adresse an und erstreckt sich bis zur derjenigen Adresse, die im Register `SP` gespeichert ist. Die Stack-Größe ist damit dynamisch, denn das Register `SP` wird sowohl durch die Instruktionen `PUSH` und `POP`, als auch direkt vom Programmierer geändert.

Das Wachsen des Stacks bedeutet, dass das Register `SP` immer kleinere Werte annimmt. Das Schrumpfen des Stacks bedeutet, dass `SP` immer größere Werte annimmt. Wird versucht, den Inhalt von `SP` kleiner Null oder größer als die maximale Speicheradresse zu setzen, so wird dies von der Maschine verweigert und als Fehler im Register `ERR` signalisiert.

Beim Hochfahren der Maschine, wird das Register `SP` auf die maximal-erreichbare Speicheradresse plus Eins gesetzt. Damit können keine Werte gelesen werden, bevor Werte geschrieben wurden.

2.5 I/O Einheit

Die I/O-Architektur der UMach Maschine ist am sogenannten „Port-Mapped I/O“ angelehnt[1]. Obwolhl diese Architektur deutliche Einschränkungen hat, wird sie wegen ihrer Einfachheit benutzt.

Die I/O-Einheit der UMach Maschine besteht aus einer Reihe von jeweils 8 Eingangs- und Ausgangschnittstellen, auch Ports genannt. An diesen Ports können verschiedene physikalische Geräte angeschlossen werden, die die entsprechenden Daten generieren bzw. verarbeiten können. Siehe dazu auch die Abbildung 2.1 auf der Seite 9. Die Eingabe und Ausgabe erfolgt dadurch, dass die Maschine die Port-Signale (Inhalte) in Register kopiert oder umgekehrt.

Die I/O Ports sind in zwei Kategorien unterteilt: 8 Eingabeports und 8 Ausgabeports. Von der Bauart und Struktur her, gibt es innerhalb der jeweiligen Kategorie keine Unterschiede zwischen Ports. Sie werden lediglich anhand deren Nummern identifiziert.

[1] Als Gegenstück von „Memory Mapped I/O“.

Die Nummerierung der Ports fängt bei 1 an. Der Port mit Nummer 0 (Null) ist eine ungültige Angabe.

Die Eingabe- und Ausgabefunktionen können durch bestimmten Instruktionen gesteuert werden (siehe auch den Abschnitt 3.11, auf der Seite 54).

2.5.1 Eingabeports

Die acht möglichen Eingabeports sind mit den Nummern 1 bis 8 gekennzeichnet. Dem Lesen von einem Eingabeport geht immer ein Hardware-Unterbrechung vorraus. Dabei wird die Portnummer, von welcher die Unterbrechung ausgeht, in das Spezialregister HIR abgelegt. Nun kann vom Port wiederholt mit dem Befehl IN immer 32 Bit oder ein Wort von 4 Byte in ein angegebenes Register gelesen werden. Sind alle Daten gelesen, wird das Register HIR auf Null gesetzt.

2.5.2 Ausgabeports

Die Ausgabe eines Registerinhaltes erfolgt dadurch, dass die UMach Maschine diesen Inhalt an einen ihrer Ausgabeports schickt.

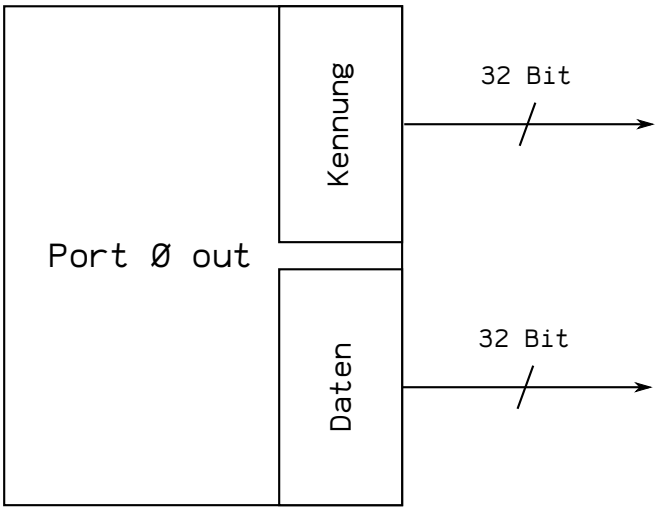

Abbildung 2.4: I/O Ausgabeport

Ein Ausgabeport besteht aus zwei Teilen, bzw. zwei Ausgängen, die jeweils 32 Bit-Signale an das angeschlossene Gerät weitergeben (siehe auch die Abbildung 2.4):

1. Kennung, die das Verwendungszweck der Daten gekennzeichnet.

2. Daten, die den eigentlichen ausgegebenen Inhalt darstellen.

Die Interpretation und Verwendung der Kennung und der Daten liegt in der Verantwortlichkeit des Programmierers, der sie verwendet, bzw. des Gerätes, das sie empfängt. Es werden seitens dieser Spezifikation keine Einschränkungen oder Richtlinien bzgl. dieser Verwendung gegeben. Insbesondere, könnte ein Gerät die Kennung ignorieren.

2.6 Unterbrechungen

Es gibt drei verschiedene Unterbrechungsursachen:

1. Hardware-Unterbrechungen, die von den Input-Ports verursacht werden können,

2. geziehlte Software-Unterbrechungen die vom der Software angefordert werden können, und

3. Unterbrechungen, die aufgrund von Fehlern geschehen und eine rudimentäre Fehlerbehandlung bei schwerwiegenden Fehlern wie die Nulldivision ermöglichen sollen.

Bei jeder Unterbrechung wird der PC auf den Stack gepuscht und bei Rückkehr automatisch gepopt. Daher sind bei einer Unterbrechung zwei Dinge zu beachten:

Der Stackpointer muss auf der Ausgangsposition stehen. Verwendete Register müssen gesichert und vor der Rückkeht wieder hergestellt werden.

2.6.1 Hardware-Unterbrechungen

Wird eine Hardware-Unterbrechung ausgelöst, so wird ohne Zutun des Programmierers der aktuelle Programmablauf unterbrochen und der PC auf den Stack gepuscht. Weiterhin wird die Nummer des für die Hardware-Unterbrechung verantwortlichen Ports im Spezialregister HIR hinterlegt. Dies dient dazu um den passenden Eintrag in der Unterbrechungstabelle aufzufinden. Dieser wird in den PC geladen und der dort Adressierte Programmcode zur Behandlung der Unterbrechung wird ausgeführt.

Ein Hardwareinterrupt entspricht somit einem INT Portnummer, wobei die CPU zusätzlich in einen Modus versetzt welcher jede weitere Hardwareunterbrechung ignoriert. Dieser Modus wird automatisch mit dem IRET Befehl zurückgesetzt, welcher dazu Dient aus der Unterbrechungsbehandlung zurückzukehren.

Das Ende einer Hardwareunterbrechung kann entweder vom Verursacher als für beendet erklärt werden, indem er das HIR nach dem Ausführen einer IN Instruktion auf Null setzt. Oder aber durch die Unterbrechungsbehandlung selbst, indem einfach IRET aufgerufen wird. Dieser Befehl setzt das HIR auch auf Null zurück, falls dies noch nicht der Fall sein sollte. Auserdem wird dem Unterbrecher das Ende der Unterbrechungsbehandlung signalisiert.

Dies ist besonderst von Bedeutung, wenn noch Daten am Port anstehen sollten. Diese werden im Falle eines vorzeitigigen Beendens der Unterbrechungsbehandlung verworfen.

2.6.2 Software-Unterbrechungen

2.6.3 Fehlerbedingte Unterbrechungen

Im Normalfall wird eine Instruktion ohne Weiteres ausgeführt. In Ausnahmefällen – wenn die Instruktion nicht ausgeführt werden kann, meistens wegen ungültigen Parametern – wird eine Ausnahmesituation signalisiert und der Programmfluss unterbrochen.

Mit jeder Ausnahmesituation wird eine Unterbrechungsnummer assoziiert. Diese Nummer wird als Index in der Unterbrechungstabelle verwendet, um die Speicheradresse einer Unterbrechungsroutine zu finden (siehe den Abschnitt 2.4.2, besonders die Tabelle 2.3 auf der Seite 16). Ist diese Adresse ungleich Null, so wird das `PC` Register entsprechend gesetzt und die Unterbrechungsroutine ausgeführt. Wird keine solche Routine gefunden, so wird die gesamte Programmausführung unterbrochen und die Maschine hält an.

Beispiel Zum Beispiel, die `DIV` Instruktion (Division) erzeugt die Unterbrechung mit Nummer 16, falls ihr zweites Argument gleich Null ist (siehe auch die Tabelle 2.3 auf der Seite 16). Wenn die Unterbrechung erzeugt wird, schaut die Maschine in der Unterbrechungstabelle an der Position 16 nach. Ist der Eintrag ungleich Null, so wird der Eintrag als Adresse einer Routine behandelt und dorthin gesprungen: das Register `PC` wird zuerst auf den Stack gespeichert und dann gleich dem Tabelleneintrag gesetzt. Ist der Eintrag dagegen Null, so hält die Maschine an.

3 Instruktionssatz

In diesem Kapitel werden alle Instruktionen der UMach VM vorgestellt.

3.1 Instruktionsformate

Eine Instruktion besteht aus einer Folge von 4 Bytes. Das Instruktionsformat beschreibt die Struktur einer Instruktion auf Byte-Ebene. Das Format gibt an, ob ein Byte als eine Registerangabe oder als reine numerische Angabe zu interpretieren ist.

Instruktionsbreite Jede UMach-Instruktion hat eine feste Bitlänge von 32 Bit (4 mal 8 Bit). Instruktionen, die für ihren Informationsgehalt weniger als 32 Bit brauchen, wie z.B. `NOP`, werden mit Nullbits gefüllt. Alle Daten und Informationen, die mit einer Instruktion übergeben werden, müssen in diesen 32 Bit untergebracht werden.

Byte Order Die Byte Order (Endianness) der gelesenen Bytes ist big-endian. Die zuerst gelesenen 8 Bits sind die 8 höchstwertigen (Wertigkeiten 2^{31} bis 2^{24}) und die zuletzt gelesenen Bits sind die niedrigstwertigen (Wertigkeiten 2^7 bis 2^0). Bits werden in Stücken von n Bits gelesen, wobei $n = k \cdot 8$ mit $k \in \{1, 4\}$ (byteweise oder wortweise).

Allgemeines Format Jede Instruktion besteht aus zwei Teilen: der erste Teil ist 8 Bit lang und entspricht dem tatsächlichen Befehl, bzw. der Operation, die von der UMach virtuellen Maschine ausgeführt werden soll. Dieser 8-Bit-Befehl belegt also die 8 höchstwertigen Bits einer 32-Bit-Instruktion. Die übrigen 24 Bits, wenn sie verwendet werden, werden für Operanden oder Daten benutzt. Beispiel einer Instruktionszerlegung:

Instruktion (32 Bit)	`00000001`	`00000010`	`00000011`	`00000100`
Hexa	`01`	`02`	`03`	`04`
Byte Order	erstes Byte	zweites Byte	drittes Byte	viertes Byte
Interpretation	Befehl (8 Bit)	Operanden, Daten oder Füllbits		

Die Instruktionsformate unterscheiden sich lediglich darin, wie sie die 24 Bits nach dem 8-Bit Befehl verwenden. Das wird auch in der 3-buchstabigen Benennung deren Formate wiedergeben.

In den folgenden Abschnitten werden die UMach-Instruktionsformate vorgestellt. Jede Angegebene Tabelle gibt in der ersten Zeile die Reihenfolge der Bytes an. Die nächste Zeile gibt die spezielle Belegung der einzelnen Bytes an.

3.1.1 000

erstes Byte	zweites Byte	drittes Byte	viertes Byte
Befehl	nicht verwendet		

Eine Instruktion, die das Format 000 hat, besteht lediglich aus einem Befehl ohne Argumenten. Die letzen drei Bytes werden von der Maschine nicht ausgewertet und sind somit Füllbytes. Es wird empfohlen, die letzten 3 Bytes mit Nullen zu füllen.

3.1.2 NNN

erstes Byte	zweites Byte	drittes Byte	viertes Byte
Befehl	numerische Angabe N		

Die Instruktion im Format NNN besteht aus einem Befehl im ersten Byte und aus einer numerischen Angabe N (einer Zahl), die die letzten 3 Bytes belegt. Die Interpretation der numerischen Angabe wird dem jeweiligen Befehl überlassen.

3.1.3 R00

erstes Byte	zweites Byte	drittes Byte	viertes Byte
Befehl	R_1	nicht verwendet	

Die Instruktion im Format R00 besteht aus einem Befehl im ersten Byte gefolgt von der Angabe eines Registers im zweiten Byte. Die letzten zwei Bytes werden nicht verwendet, bzw. werden ignoriert.

3.1.4 RNN

erstes Byte	zweites Byte	drittes Byte	viertes Byte
Befehl	R_1	numerische Angabe N	

Eine Instruktion im Format RNN besteht aus einem Befehl, gefolgt von einer Register Nummer R_1, gefolgt von einer festen Zahl N, die die letzten 2 Bytes der Instruktion belegt. Die genaue Interpretation der Zahl N wird dem jeweiligen Befehl überlassen. Zum Beispiel, die Instruktion

erstes Byte	zweites Byte	drittes Byte	viertes Byte
`0x20`	`0x01`	`0x02`	`0x03`

wird folgenderweise von der UMach Maschine interpretiert: die Operation mit Nummer `0x20` soll ausgeführt werden, wobei die Argumenten dieser Operation sind das Register mit Nummer `0x01` und die numerische Angabe `0x0203`.

3.1.5 RR0

erstes Byte	zweites Byte	drittes Byte	viertes Byte
Befehl	R_1	R_2	nicht verwendet

Eine Instruktion im Format RR0 besteht aus einem Befehl im ersten Byte, gefolgt von der Angabe zweier Register in den folgenden 2 Bytes. Das dritte Byte wird nicht verwendet, bzw. wird ignoriert.

3.1.6 RRN

erstes Byte	zweites Byte	drittes Byte	viertes Byte
Befehl	R_1	R_2	numerische Angabe N

Eine Instruktion im Format RRN besteht aus einem Befehl, gefolgt von der Angabe zweier Registers R_1 und R_2, jeweils in einem Byte, gefolgt von einer numerischen Angabe N (festen Zahl) im letzten Byte. Zum Beispiel, die Instruktion

erstes Byte	zweites Byte	drittes Byte	viertes Byte
`0x52`	`0x01`	`0x02`	`0x03`

soll wie folgt interpretiert werden: die Operation mit Nummer `0x52` soll ausgeführt werden, wobei die Argumenten dieser Operation sind Register mit Nummer `0x01`, Register mit Nummer `0x02` und die Zahl `0x03`.

3.1.7 RRR

erstes Byte	zweites Byte	drittes Byte	viertes Byte
Befehl	R_1	R_2	R_3

Eine Instruktion im Format RRR besteht aus der Angabe eines Befehls im ersten Byte, gefolgt von der Angabe dreier Register R_1, R_2 und R_3 in den jeweiligen folgenden drei Bytes. Die Register werden als Zahlen angegeben und deren Bedeutung hängt vom jeweiligen Befehl ab.

3.1.8 Zusammenfassung

Im folgenden werden die Instruktionsformate tabellarisch zusammengefasst.

Format	erstes Byte	zweites Byte	drittes Byte	viertes Byte
000	Befehl	nicht verwendet		
NNN	Befehl	numerische Angabe N		
R00	Befehl	R_1	nicht verwendet	
RNN	Befehl	R_1	numerische Angabe N	
RR0	Befehl	R_1	R_2	nicht verwendet
RRN	Befehl	R_1	R_2	numerische Angabe N
RRR	Befehl	R_1	R_2	R_3

3.2 Verteilung des Befehlsraums

Zur besseren Übersicht der verschiedenen UMach-Instructionen, unterteilen wir den Instruktionssatz der UMach virtuellen Maschine in den folgenden Kategorien:

Maschinencodes	Kategorie
00 - 0F	Kontrollbefehle
10 - 2F	Lade-/Speicherbefehle
30 - 4F	Arithmetische Befehle
50 - 6F	Logische Befehle
70 - 7F	Vergleichsbefehle
80 - 8F	Sprungbefehle
90 - 9F	Unterprogrambefehle
A0 - AF	Systembefehle
B0 - BF	IO Befehle

Tabelle 3.1: Verteilung des Befehlsraums nach Befehlskategorien. Die Zahlen sind im Hexadezimalsystem angegeben.

1. Kontrollinstruktionen, die die Maschine in ihrer gesamten Funktionalität betreffen, wie z.B. den Betriebsmodus umschalten.

2. Lade- und Speicherbefehle, die Register mit Werten aus dem Speicher, anderen Registern oder direkten numerischen Angaben laden und die Registerinhalte in den Speicher schreiben.

3. Arithmetische Instruktionen, die einfache arithmetische Operationen zwischen Registern veranlassen.

4. Logische Instruktionen, die logische Verknüpfungen zwischen Registerinhalten oder Operationen auf Bit-Ebene in Registern anweisen.

5. Vergleichsinstruktionen, die einen Vergleich zwischen Registerinhalten angeben.

6. Sprunginstruktionen, die bedingt oder unbedingt sein können. Sie weisen die UMach Maschine an, die Programmausführung an einer anderen Stelle fortzufahren.

7. Unterprogramm-Steuerung, bzw. Instruktionen, die die Ausführung von Unterprogrammen (Subroutinen) steuern.

8. Systeminstruktionen, die die Unterstützung eines Betriebssystem ermöglichen.

9. IO Instruktionen

Die oben angegebenen Instruktionskategorien unterteilen den Befehlsraum in 9 Bereiche. Es gibt 256 mögliche Befehle, gemäß $2^8 = 256$. Die Verteilung der Kategorien auf die verschiedenen Maschinencode-Intervallen wird in der Tabelle 3.1 auf Seite 25 angegeben.

Tabelle 3.2: Befehlentabelle

	0	1	2	3	4	5	6	7
0	NOP				EOP			
1	SET	CP	LB	LW	SB	SW		
	PUSH	POP						
2								
3	ADD	ADDU	ADDI	SUB	SUBU	SUBI		
	MUL	MULU	MULI	DIV	DIVU	DIVI		
4	ABS	NEG	INC	DEC				
	MOD	MODI						
5	AND	ANDI	OR	ORI	XOR	XORI	NOT	NOTI
	NAND	NANDI	NOR	NORI				
6	SHL	SHLI	SHR	SHRI	SHRA	SHRAI		
	ROTL	ROTLI	ROTR	ROTRI				
7	CMP	CMPU	CMPI					
8	BE	BNE	BL	BLE	BG	BGE		
	JMP							
9	GO	CALL	RET					
A	INT	IRET						
B	IN							
	OUT							
C								
D								
E								
F								
	8	9	A	B	C	D	E	F

Die Tabelle 3.2 auf der Seite 26 enthält eine Übersicht aller Befehle und deren Maschinencodes. Diese Tabelle wird folgenderweise gelesen: in der am weitesten linken Spalte wird die erste hexadezimale Ziffer eines Befehls angegeben (ein Befehl ist zweistellig im Hexadezimalsystem). Jede solche Ziffer hat rechts Zwei Zeilen, die von links nach rechts gelesen werden: eine Zeile für die Ziffern von 0 bis 8, die anderen für die übrigen Ziffern 9 bis F (im Hexadezimalsystem). Die Assemblernamen (Mnemonics) der einzelnen Befehle sind an der entsprechenden Stelle angegeben.

Definitionsstruktur Im den folgenden Abschnitten werden die einzelnen Instruktionen beschrieben. Zu jeder Instruktion wird der Assemblername, die Parameter, der Maschinencode und das Instruktionsformat, das die Typen der Parameter definiert, formal angegeben. Zudem werden Anwendungsbeispiele angegeben. Die Instruktionsformate können im Abschnitt 3.1 ab der Seite 21 nachgeschlagen werden.

Zur Notation Mit $\mathcal{R}$ wird die Menge aller Register gekennzeichnet[1]. Die Notation $X \in \mathcal{R}$ bedeutet, dass X ein Element aus dieser Menge ist, mit anderen Worten, dass X ein Register ist. Analog bedeutet die Schreibweise $X, Y \in \mathcal{R}$, dass X und Y beide Register sind.

Gilt $X, Y \in \mathcal{R}$ und ist $\sim$ eine durch einen Befehl definierte Relation zwischen X und Y, so bezieht sich die Schreibweise $X \sim Y$ nicht auf die Maschinennamen von X und Y, sondern auf deren Inhalte. Zum Beispiel, haben die Register $R1$ und $R2$ die Maschinencodes `0x01` und `0x02` und sind sie mit den Werten 4 bzw. 5 belegt, so bedeutet $R1 + R2$ das gleiche wie $4 + 5 = 9$ und nicht `0x01 + 0x02 = 0x03`.

Andere verwendeten Schreibweisen:

$\mathbb{N}$	Menge aller natürlichen Zahlen: $0, 1, 2, \ldots$
$\mathbb{N}_{\setminus 0}$	$\mathbb{N}$ ohne die Null: $1, 2, \ldots$
$\mathbb{Z}$	Menge aller ganzen Zahlen: $\ldots, -2, -1, 0, 1, 2, \ldots$
$\mathbb{Z}_{\setminus 0}$	$\mathbb{Z}$ ohne die Null: $\ldots, -2, -1, 1, 2, \ldots$
$N \in \mathbb{N}$	N ist Element von $\mathbb{N}$, oder liegt im Bereich von $\mathbb{N}$
$X \leftarrow Y$	X wird auf Y gesetzt
$mem(X)$	Speicherinhalt an der Adresse X (1 Byte)
$mem_n(X)$	n-Bytes-Block im Speicher ab Adresse X
`mem[n]`	äquivalent zu $mem(n)$

[1] Nicht verwechseln mit den Symbolen $\mathbb{R}$ und $\mathbb{R}$, die die Menge aller reellen Zahlen bedeuten.

3.3 Kontrollinstruktionen

3.3.1 NOP

Assemblername	Parameter	Maschinencode	Format
NOP	keine	0x00	000

Diese Instruktion („No Operation“) bewirkt nichts. Der Sinn dieser Instruktion ist, den Maschinencode mit Nullen füllen zu können, ohne dabei die gesamte Ausführung zu beeinflussen, außer, Zeitlupen zu schaffen.

3.3.2 EOP

Assemblername	Parameter	Maschinencode	Format
EOP	keine	0x04	000

„End Of Program“. Die Maschine ausschalten.

3.4 Lade- und Speicherbefehle

3.4.1 SET

Assemblername	Parameter	Maschinencode	Format
SET	$X \in \mathcal{R}$, $N \in \mathbb{Z}$	0x10	RNN

Setzt den Inhalt des Registers X auf den ganzzahligen Wert N. Da N mit 16 Bit und im Zweierkomplement dargestellt wird, kann N Werte von -2^{15} bis $2^{15} - 1$ aufnehmen, bzw. von -32768 bis $+32767$. Werte außerhalb dieses Intervalls werden auf Assembler-Ebene entsprechend gekürzt (es wird modulo berechnet, bzw. nur die ersten 16 Bits aufgenommen).

Beispiele:

```
label:
  SET R1   8     # R1 ← 8
  SET R2  -3     # R2 ← −3
  SET R3 65536 # R3 ← 0, da 65536 = 2^16 ≡ 0 mod 2^16
  SET R4 70000 # R3 ← 4464 = 70000 mod 2^16
  SET R7 label # Adresse 'label' ins R7
```

3.4.2 CP

Assemblername	Parameter	Maschinencode	Format
CP	$X, Y \in \mathcal{R}$	0x11	RR0

Kopiert den Inhalt des Registers Y in das Register X. Register Y wird dabei nicht geändert. Entspricht

$$X \leftarrow Y$$

Beispiel:

```
SET   R1 5   # R1 ← 5
CP    R2 R1  # R2 ← 5
```

3.4.3 LB

Assemblername	Parameter	Maschinencode	Format
LB	$X, Y \in \mathcal{R}$	0x12	RR0

Lade ein Byte aus dem Speicher mit Adresse Y in das niedrigstwertige Byte des Registers X. Die anderen Bytes von X werden von diesem Befehl nicht geändert. Insbesondere, werden sie nicht auf Null gesetzt.

Äquivalenter C Code:

```
x = (x & 0xFFFFFF00) | (mem(y) & 0x00FF);
```

Beispiel Angenommen, der Speicher an den Adressen 100 und 101 hat den Wert 5, bzw. 6. Dann können diese zwei nacheinander folgenden Bytes auf die folgende Weise in $R2$ gelesen werden. $R1$ wird dabei als Zeiger im Speicher verwendet.

```
SET  R1     100   # Speicheradresse R1 = 100
SET  R2       0
LB   R2  R1       # R2 = 5 (mem(100))
SHLI R2  R2   8   # shift left 8 Bit, R2 = 1280
INC  R1           # R1++, R1 = 101
LB   R2  R1       # R2 = 1286 (R2 + mem(101))
```

3.4.4 LW

Assemblername	Parameter	Maschinencode	Format
LW	$X, Y \in \mathcal{R}$	0x13	RR0

„Load Word“. Lade ein Wort (4 Byte) aus dem Speicher mit Adresse Y in das Register X. Alle Bytes von X werden dabei überschrieben. Die Bytes aus dem Speicher werden nacheinander gelesen. Es werden also die Bytes mit Adressen $Y + 0$, $Y + 1$, $Y + 2$ und $Y + 3$ zu einem 4-Byte Wort zusammengesetzt und so in X ablegt.

Beispiel Abgenommen, die Adressen von 100 bis 103 sind mit den Werten 0, 1, 2 und 3 belegt und bilden somit den Wert 66051.

```
SET R1   100
LW  R2 R1      # R2 ← mem4(R1) = 66051
```

3.4.5 SB

Assemblername	Parameter	Maschinencode	Format
SB	$X, Y \in \mathcal{R}$	0x14	RR0

„Store Byte“. Speichert den Inhalt des niedrigstwertigen Byte von X an der Speicherstelle Y. X und Y sind dabei Register.

Entspricht dem algebraischen Ausdruck

$$X \rightarrow mem_1(Y)$$

$mem_1(x)$ bedeutet dabei 1 Byte an der Adresse x.

```
SET R1 128      # R1 = Speicheradresse 128
SET R2 513      # R2 = 0x0201
SB  R2 R1       # Speicher mit Adresse 128 wird auf 1 gesetzt
```

3.4.6 SW

Assemblername	Parameter	Maschinencode	Format
SW	$X, Y \in \mathcal{R}$	0x15	RR0

„Store Word“. Speichert den Inhalt aller Bytes in X an die Speicheradressen Y bis $Y+3$.

$$X \rightarrow mem_4(Y)$$

Beispiel Es wird das Register $R2$ mit dem Wert 0x01020304 geladen und an die Adresse 128 gespeichert. Dabei werden die Byte-Werten in „big-endian“ Reihenfolge gespeichert: das höchstwertige Byte aus $R2$ (0x01) wird an der Adresse 128 gespeichert, das niedrigstwertige Byte (0x04) an die Adresse 131.

```
SET R1 128          # R1 = Speicheradresse 128
SET R2 0x01020304   # Wert zum Speichern
SH  R2 R1           # mem[128] = 0x01
                    # mem[129] = 0x02
                    # mem[130] = 0x03
                    # mem[131] = 0x04
```

3.4.7 PUSH

Assemblername	Parameter	Maschinencode	Format
PUSH	$X \in \mathcal{R}$	0x18	R00

„Push Word“. Erniedrigt das Register SP um 4 und kopiert das ganze Register X auf den Stack, wobei der „Stack“ ist der Speicherbereich mit Anfangsadresse in SP. Die

Byte-Reihenfolge der Lese- und Schreiboperationen ist „Big-Endian“ und wird in der nachfolgenden Tabelle dargestellt:

X Wertigkeiten	$2^{31} \leftrightarrow 2^{24}$	$2^{23} \leftrightarrow 2^{16}$	$2^{15} \leftrightarrow 2^{8}$	$2^{7} \leftrightarrow 2^{0}$
	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$
Stack-Bereich	`mem[SP + 0]`	`mem[SP + 1]`	`mem[SP + 2]`	`mem[SP + 3]`

Entspricht

$$SP \leftarrow SP - 4$$
$$X \rightarrow mem_4(SP)$$

Beispiel Der folgende Code speichert das 4-Byte Wort `0x01020304` auf den Stack. Die Stack-Struktur wird in Kommentaren gezeigt.

```
SET  R1 0x01020304 # Wert zum pushen
PUSH R1            # mem[SP + 0] = 0x01
                   # mem[SP + 1] = 0x02
                   # mem[SP + 2] = 0x03
                   # mem[SP + 3] = 0x04
```

3.4.8 `POP`

Assemblername	Parameter	Maschinencode	Format
`POP`	$X \in \mathcal{R}$	`0x19`	R00

„Pop Word“. Speichert 4 Bytes ab der Adresse `SP` in das Register X und erhöht `SP` um 4. Die Byte-Reihenfolge der Lese- und Schreiboperationen ist „Big-Endian“ und wird in der nachfolgenden Tabelle dargestellt.

X Wertigkeiten	$2^{31} \leftrightarrow 2^{24}$	$2^{23} \leftrightarrow 2^{16}$	$2^{15} \leftrightarrow 2^{8}$	$2^{7} \leftrightarrow 2^{0}$
	$\uparrow$	$\uparrow$	$\uparrow$	$\uparrow$
Stack-Bereich	`mem[SP + 0]`	`mem[SP + 1]`	`mem[SP + 2]`	`mem[SP + 3]`

Diese Instruktion kann algebraisch so ausgedrückt werden:

$$X \leftarrow mem_4(SP)$$
$$SP \leftarrow SP + 4$$

Beispiel Angenommen, die ersten 4 Bytes auf dem Stack (d.h. ab der Adresse, die im Register `SP` angegeben ist) sind `0xAA 0xBB 0xCC 0xDD`.

```
POPH R1       # R1 = 0xAABBCCDD
```

3.5 Arithmetische Instruktionen

3.5.1 `ADD`

Assemblername	Parameter	Maschinencode	Format
`ADD`	$X, Y, Z \in \mathcal{R}$	`0x30`	RRR

Vorzeichen behaftete Addition der Registerinhalte Y und Z. Das Ergebnis der Addition wird in das Register X gespeichert. Entspricht dem algebraischen Ausdruck

$$X \leftarrow Y + Z$$

Beispiel:

```
SET   R1 1      # R1 ← 1
SET   R2 2      # R2 ← 2
ADD   R3 R1 R2  # R3 ← R1 + R2 = 1 + 2 = 3
#     X  Y  Z
SET   R2 -2     # R2 ← −2
ADD   R3 R3 R2  # R3 ← R3 + R2 = 3 + (−2) = 1
ADD   R3 R4  5  # Fehler! 5 kein Register
```

Vorzeichenlose Addition wird durch den Befehl `ADDU` ausgeführt.

3.5.2 `ADDU`

Assemblername	Parameter	Maschinencode	Format
`ADDU`	$X, Y, Z \in \mathcal{R}$	`0x31`	RRR

„Add Unsigned“. Vorzeichenlose Addition der Register Y und Z. Das Ergebnis wird in das Register X gespeichert. Enthält Y oder Z ein Vorzeichen (höchstwertiges Bit auf 1 gesetzt), so wird es nicht als solches interpretiert, sondern als Wertigkeit, die zum Betrag des Wertes hinzuaddiert wird ($+2^{31}$).

```
SET   R1 1      # R1 ← 1
SET   R2 -2     # R2 ← −2
ADDU  R3 R1 R2  # R3 ← (1 + 2 + 2^31) = 2147483651
```

3.5.3 ADDI

Assemblername	Parameter	Maschinencode	Format
ADDI	$X, Y \in \mathcal{R}$, $N \in \mathbb{Z}$	0x32	RRN

„Add Immediate“. Hinzuaddieren eines festen vorzeichenbehafteten ganzzahligen Wert N zum Inhalt des Registers Y und speichern des Ergebnisses in das Register X. Entspricht dem algebraischen Ausdruck

$$X \leftarrow Y + N$$

N wird als vorzeichenbehaftete 8-Bit Zahl in Zweierkomplement-Darstellung interpretiert und kann entsprechend Werte von -128 bis 127 aufnehmen.

Beispiel:

```
SET   R1 1      # R1 ← 1
ADDI  R2 R1 2   # R2 ← R1 + 2 = 1 + 2 = 3
#     X  Y  N
ADDI  R2 R2 -3  # R2 ← R2 + (−3) = 3 + (−2) = 1
ADDI  R2 R3 R4  # Fehler! R4 kein n ∈ Z
```

3.5.4 SUB

Assemblername	Parameter	Maschinencode	Format
SUB	$X, Y, Z \in \mathcal{R}$	0x33	RRR

Subtrahiert die Registerinhalte von Y und Z und speichert das Ergebnis in das Register X. Entspricht dem Ausdruck

$$X \leftarrow (Y - Z)$$

Wobei X, Y und Z Register sind.

3.5.5 SUBU

Assemblername	Parameter	Maschinencode	Format
SUBU	$X, Y, Z \in \mathcal{R}$	0x34	RRR

„Subtract Unsigned“. Analog zur Instruktion SUB mit dem Unterschied, dass alle Werte und Operationen vorzeichenlos sind.

3.5.6 SUBI

Assemblername	Parameter	Maschinencode	Format
SUBI	$X, Y \in \mathcal{R},\ N \in \mathbb{Z}$	0x35	RRN

„Subtract Immediate“. Funktioniert wie SUB aber N ist eine direkt angegebene Zahl (kein Register).

Beispiel Folgendes Beispiel demonstriert die Verwendung von SUBI und zeigt zugleich einen Fehler.

```
SET  R1 50      # R1 ← 50
SUBI R2 R1 30   # R2 ← (R1 − 30) = 20
SUBI R2 R1 R1   # Fehler! da R1 ∉ Z
```

3.5.7 MUL

Assemblername	Parameter	Maschinencode	Format
MUL	$X, Y \in \mathcal{R}$	0x38	RR0

„Multiply“. Multipliziert die Inhalte der Register X und Y und speichert das Ergebnis in die Spezialregister HI und LO. Diese zwei Spezialregister werden als eine 64-Bit Einheit betrachtet, wobei jedes eine Hälfte des 64-Bit Ergebnisses enthält. Dabei werden die höchstwertigen 32 Bit des Ergebnisses in das Register HI und die 32 niedrigstwertigen Bits des Ergebnisses in das Register LO gespeichert. Siehe auch die Tabelle 2.1 auf der Seite 13.

Falls das Ergebnis der Multiplikation gänzlich in den 32 Bit des Registers `LO` passt, wird das Register `HI` trotzdem auf Null gesetzt.

Äquivalenter algebraischer Ausdruck:

$$(HI, LO) \leftarrow X \cdot Y$$

Beispiel Der folgende Code demonstriert die Verwendung der `MUL` Instruktion.

```
SET   R1 4     # R1 ← 4
SET   R2 5     # R2 ← 5
MUL   R1 R2    # HI ← 0
               # LO ← 20

SET   R1 0xAAAAAAAA
MUL   R1 R1    # R1^2
               # HI = 0x71C71C70
               # LO = 0xE38E38E4
COPY  R2 LO    # R1^2 mod 2^32
```

3.5.8 `MULU`

Assemblername	Parameter	Maschinencode	Format
`MULU`	$X, Y \in \mathcal{R}$	`0x39`	RR0

„Multiply Unsigned“. Funktioniert wie die Instruktion `MUL` mit dem Unterschied, dass die Multiplikationoperanden X und Y vorzeichenlos behandelt werden.

3.5.9 `MULI`

Assemblername	Parameter	Maschinencode	Format
`MULI`	$X \in \mathcal{R}$, $N \in \mathbb{Z}$	`0x3A`	RNN

„Multiply Immediate“. Multipliziert den Inhalt des Registers X mit der ganzen Zahl N und speichert das 64-Bit Ergebnis in die Register `HI` und `LO`, die als ein einziges 64-Register betrachtet werden: `HI` enthält die ersten 32 Bits (die höchstwertigen) und `LO` die letzten 32 Bits (die niedrigstwertigen). Siehe auch die Instruktion `MUL`.

3.5.10 DIV

Assemblername	Parameter	Maschinencode	Format
DIV	$X, Y \in \mathcal{R}$	0x3B	RR0

„Divide“, ganzzahlige Division. Dividiert den Inhalt des Registers X durch den Inhalt des Registers Y und speichert den Quotient in das Register HI und den Rest in das Register LO. Nach der Ausführung gilt

$$X = Y \cdot HI + LO$$

Algebraisch ausgedrückt:

$$HI \leftarrow \lfloor X/Y \rfloor$$
$$LO \leftarrow X \bmod Y$$

$\lfloor x \rfloor$ bedeutet in diesem Kontext, dass x auf die betragsmässig nächstkleinste ganze Zahl gerundet wird, oder die Nachkommastellen von x werden abgeschnitten.

Beispiel Der folgende Code demonstriert die Verwendung von DIV.

```
SET R1 10    # R1 ← 10
SET R2  3    # R2 ← 3
DIV R1 R2    # HI ← 3
             # LO ← 1
```

3.5.11 DIVU

Assemblername	Parameter	Maschinencode	Format
DIVU	$X, Y \in \mathcal{R}$	0x3C	RR0

„Divide Unsigned“. Funktioniert wie DIV mit dem Unterschied, dass ganzzahlige vorzeichenlose Division durchgeführt werden. Die Ergebnis-Register HI und LO enthalten entsprechend vorzeichenlose Werte.

3.5.12 DIVI

Assemblername	Parameter	Maschinencode	Format
DIVI	$X \in \mathcal{R}$, $N \in \mathbb{Z}_{\setminus 0}$	0x3D	RNN

„Divide Immediate“. Dividiert den Inhalt des Registers X durch die feste ganze Zahl N und speichert den Quotient in das Register `HI` und den Rest in das Register `LO`. N nimmt Werte aus dem Intervall $[-2^{15}, 2^{15} - 1] \setminus 0$. Nach der Durchführung der Division gilt:

$$X = HI \cdot N + LO$$

Beispiel Der folgende Code demonstriert die Verwendung von `DIVI`.

```
SET   R1  10    # R1 ← 10
DIVI  R1   3    # HI ← 3
                # LO ← 1
```

3.5.13 ABS

Assemblername	Parameter	Maschinencode	Format
`ABS`	$X, Y \in \mathcal{R}$	`0x40`	RR0

„Absolute“. Speichert den absoluten Wert des Registers Y in das Register X. Algebraisch ausgedrückt:

$$X \leftarrow \begin{cases} Y & \text{falls } Y \geq 0 \\ (-1) \cdot Y & \text{falls } Y < 0 \end{cases}$$

3.5.14 NEG

Assemblername	Parameter	Maschinencode	Format
`NEG`	$X, Y \in \mathcal{R}$	`0x41`	RR0

„Negate“. Wechselt das arithmetische Vorzeichen des Registers Y und speichert das Ergebnis in das Register X. Entspricht der Zweierkomplement Bildung. Algebraische Schreibweise:

$$X \leftarrow \Big((-1) \cdot Y\Big)$$

Um eine bitweise Inversion zu erreichen (Einerkomplement), siehe die Instruktion `NOT`.

3.5.15 INC

Assemblername	Parameter	Maschinencode	Format
INC	$X \in \mathcal{R}$	0x42	R00

„Increment“. Inkrementiert den Inhalt des Registers X.

$$X \leftarrow (X + 1)$$

3.5.16 DEC

Assemblername	Parameter	Maschinencode	Format
DEC	$X \in \mathcal{R}$	0x43	R00

„Decrement“. Dekrementiert den Inhalt des Registers X.

$$X \leftarrow (X - 1)$$

3.5.17 MOD

Assemblername	Parameter	Maschinencode	Format
MOD	$X, Y, Z \in \mathcal{R}$	0x48	RRR

Modulo Operation. Berechnet den Rest der Division Y/Z und speichert den Rest in das Register X. Dabei ist das Ergebnis immer positiv. Äquivalent zu den Instruktionen:

```
DIVU Y Z
COPY X LO
```

Algebraische Schreibweise:

$$X \leftarrow Y \bmod Z$$

oder

$$X \leftarrow \left(Y - \left\lfloor \frac{Y}{Z} \right\rfloor \cdot Z\right)$$

3.5.18 MODI

Assemblername	Parameter	Maschinencode	Format
MODI	$X, Y \in \mathcal{R}$, $N \in \mathbb{N}$	0x49	RRN

„Modulo Immediate“. Analog zur Instruktion MOD, berechnet MODI den Rest der ganzzahligen Division Y/N und speichert ihn in das Register X. Der Unterschied liegt darin, dass N eine fest angegebene natürliche Zahl ist.

$$X \leftarrow Y \bmod N$$

3.6 Logische Instruktionen

3.6.1 AND

Assemblername	Parameter	Maschinencode	Format
AND	$X, Y, Z \in \mathcal{R}$	0x50	RRR

Berechnet bitweise $Y \wedge Z$ (und-Verknüpfung) und speichert das Ergebnis in das Register X.

$$X \leftarrow (Y \wedge Z)$$

3.6.2 ANDI

Assemblername	Parameter	Maschinencode	Format
ANDI	$X, Y \in \mathcal{R}$, $N \in \mathbb{N}$	0x51	RRN

„And Immediate“. Berechnet bitweise $Y \wedge N$ (und-Verknüpfung) und speichert das Ergebnis in das Register X. N ist dabei eine feste Zahl aus dem Bereich $[0, 255]$. Wird eine größere Zahl angegeben, so wird sie modulo 256 berechnet – mit anderen Worten, nur das letzte Byte zählt. Andere Schreibweise:

$$X \leftarrow (Y \wedge (N \bmod 256))$$

Bemerkung Die natürliche Zahl N wird auf die Länge von 32 Bit mit Nullen verlängert (Links-Verlängerung). Deshalb setzt diese Instruktion alle Bits mit Wertigkeiten von 2^8 bis 2^{31} im Register X auf Null.

Beispiel für die Verwendung von `ANDI`:

```
SET  R1    0x0102
ANDI R2 R1   0x01  # R2 = 0
```

3.6.3 `OR`

Assemblername	Parameter	Maschinencode	Format
`OR`	$X, Y, Z \in \mathcal{R}$	`0x52`	RRR

Berechnet bitweise $Y \vee Z$ (oder-Verknüpfung) und speichert das Ergebnis in das Register X.

$$X \leftarrow (Y \vee Z)$$

3.6.4 `ORI`

Assemblername	Parameter	Maschinencode	Format
`ORI`	$X, Y \in \mathcal{R}$, $N \in \mathbb{N}$	`0x53`	RRN

„Or Immediate“. Berechnet wie `OR` ein bitweises „oder“zwischen Y und N und speichert das Ergebnis in das Register X. Dabei wird die natürliche Zahl $N \in [0, 255]$ auf die Länge von 32 Bit mit Nullen verlängert (Links-Verlängerung). Wird für N einen Wert größer als 255 angegeben, so wird er modulo 256 berechnet.

$$X \leftarrow (Y \vee (N \bmod 256))$$

3.6.5 `XOR`

Assemblername	Parameter	Maschinencode	Format
`XOR`	$X, Y, Z \in \mathcal{R}$	`0x54`	RRR

Berechnet bitweise $Y \oplus Z$ (xor-Verknüpfung) und speichert das Ergebnis in das Register X.

$$X \leftarrow (Y \oplus Z)$$

3.6.6 XORI

Assemblername	Parameter	Maschinencode	Format
XORI	$X, Y \in \mathcal{R}$, $N \in \mathbb{N}$	0x55	RRN

„Xor Immediate“. Analog zur Instruktion XOR mit dem Unterschied, dass N eine direkt angegebene Zahl aus dem Intervall $[0, 255]$ ist.

$$X \leftarrow (Y \oplus (N \bmod 256))$$

3.6.7 NOT

Assemblername	Parameter	Maschinencode	Format
NOT	$X, Y \in \mathcal{R}$	0x56	RR0

Invertiert alle Bits aus dem Register Y und speichert das Ergebnis in das Register X. Entspricht der Einerkomplement-Bildung.

$$X \leftarrow \overline{Y}$$

Beispiel Der folgende Code invertiert alle Bits aus dem Register $R2$ und speichert das Ergebnis in das Register $R1$.

```
NOT R1 R2 # R1 ← ¬R2
```

3.6.8 NOTI

Assemblername	Parameter	Maschinencode	Format
NOTI	$X \in \mathcal{R}$, $N \in \mathbb{N}$	0x57	RNN

„Not Immediate“. Wie die Instruktion `NOT` aber N ist eine natürliche Zahl aus dem Intervall $[0, 2^{15} - 1]$. Diese konstante Zahl nach links auf die Länge von 32 Bit mit Nullen verlängert.

Beispiel Der folgende Code invertiert alle Bits der Zahl 5 und speichert das Ergebnis in das Register $R1$.

```
NOT R1 5 # R1 ← 5̄
```

3.6.9 NAND

Assemblername	Parameter	Maschinencode	Format
`NAND`	$X, Y, Z \in \mathcal{R}$	`0x58`	RRR

Berechnet $Y \overline{\wedge} Z$ (nand-Verknüpfung) und speichert das Ergebnis in das Register X. Gemäß dem Format RRR sind X, Y und Z Register. Algebraische Schreibweise:

$$X \leftarrow (Y \overline{\wedge} Z)$$

oder

$$X \leftarrow \overline{(Y \wedge Z)}$$

3.6.10 NANDI

Assemblername	Parameter	Maschinencode	Format
`NANDI`	$X, Y \in \mathcal{R}$, $N \in \mathbb{N}$	`0x59`	RRN

„Nand Immediate“. Berechnet eine nand-Verknüpfung zwischen dem Inhalt des Registers Y und der konstanten Zahl N und speichert das Ergebnis in das Register X. $N \in [0, 255]$.

$$X \leftarrow \left(Y \overline{\wedge} (N \bmod 256)\right)$$

3.6.11 NOR

Assemblername	Parameter	Maschinencode	Format
`NOR`	$X, Y, Z \in \mathcal{R}$	`0x5A`	RRR

Verknüpft bitweise die Inhalte der Register Y und Z mit der nor-Operation und speichert das Ergebnis in das Register X. „nor“ ist die Negation von „or“.

$$X \leftarrow (Y \overline{\vee} Z)$$

```
NOR R1 R2 R3 # R1 ← ¬(R2 ∨ R3)
```

3.6.12 NORI

Assemblername	Parameter	Maschinencode	Format
NORI	$X, Y \in \mathcal{R}$, $N \in \mathbb{N}$	0x5B	RRN

„Nor Immediate“. Analog zur Instruktion NOR mit dem Unterschied, dass N eine konstante Zahl aus dem Bereich $[0, 255]$ ist. Diese Zahl wird modulo 256 berechnet und auf der linken Seite auf die Länge von 32 Bit mit Nullen aufgefüllt.

3.6.13 SHL

Assemblername	Parameter	Maschinencode	Format
SHL	$X, Y, Z \in \mathcal{R}$	0x60	RRR

„Shift Left“. Verschiebt die Bits aus dem Register Y Z Stellen nach links und speichert das Ergebnis in das Register X. Auf der rechten Seite wird X mit Nullen aufgefüllt.

Die Stellenangabe in Z muss positiv sein.

Andere Schreibweise:

$$X \leftarrow Y << Z$$

oder

$$X \leftarrow (Y \cdot 2^Z) \bmod 2^{32}$$

3.6.14 SHLI

Assemblername	Parameter	Maschinencode	Format
SHLI	$X, Y \in \mathcal{R}$, $N \in \mathbb{N}$	0x61	RRN

„Shift Left Immediate“. Verschiebt die Bits im Register Y N Stellen nach links und speichert das Ergebnis in das Register X. N ist eine positive Zahl im Bereich $[0, 255]$. Anstelle der nach links verschobenen Bits werden Nullen aufgefüllt.

N muss positiv sein.

3.6.15 SHR

Assemblername	Parameter	Maschinencode	Format
SHR	$X, Y, Z \in \mathcal{R}$	0x62	RRR

„Shift Right“. Verschiebt die Bits aus dem Register Y Z Stellen nach rechts und speichert das Ergebnis in das Register X. Anstelle der verschobenen Bits werden im Register X auf der linken Seite Nullen aufgefüllt. Gemäß dem Instruktionsformat RRR sind X, Y und Z Register. Alle Register-Inhalte werden vorzeichenlos interpretiert.

Bemerkung Wird in dem Register Y eine negative Zahl gespeichert, so löscht diese Verschiebung das Vorzeichen, da auf der linken Seite Nullen aufgefüllt werden. Um das Vorzeichen zu behalten, sollte man die Instruktion SHRA verwenden.

Beispiel Der folgende Code verschiebt die Bits im $R1$ eine Stelle nach rechts. Es wird praktisch $R3 \leftarrow (5 \div 2)$ berechnet.

```
SET   R1 5       # R1 ← 5
SET   R2 1       # R2 ← 1
SHR   R3 R1 R2   # R3 ← 2
```

3.6.16 SHRI

Assemblername	Parameter	Maschinencode	Format
SHRI	$X, Y \in \mathcal{R}$, $N \in \mathbb{N}$	0x63	RRN

„Shift Right Immediate“. Das Bitmuster im Register Y wird N Stellen nach rechts verschoben und das Ergebnis in das Register X gespeichert. Dabei ist N eine positive natürliche Zahl aus dem Intervall $[0, 255]$. Auf der linken Seite werden die versetzten Bits mit Nullen ersetzt. Siehe auch SHRAI.

3.6.17 SHRA

Assemblername	Parameter	Maschinencode	Format
SHRA	$X, Y, Z \in \mathcal{R}$	0x64	RRR

„Shift Right Arithmetical“. Funktioniert wie die Instruktion SHR mit dem Unterschied, dass auf der linken Seite nicht mit Nullen, sondern mit dem ersten (höchstwertigen) Bit aufgefüllt wird. Dies führt dazu, dass das Vorzeichenbit in Y erhalten bleibt.

3.6.18 SHRAI

Assemblername	Parameter	Maschinencode	Format
SHRAI	$X, Y \in \mathcal{R}$, $N \in \mathbb{N}$	0x65	RRN

„Shift Right Arithmetical Immediate“. Wie SHRA, N ist aber eine natürliche Zahl aus dem Intervall $[0, 255]$.

3.6.19 ROTL

Assemblername	Parameter	Maschinencode	Format
ROTL	$X, Y, Z \in \mathcal{R}$	0x68	RRR

„Rotate Left“. Die Bits aus dem Register Y werden so viele Stellen nach links „rotiert“ wie der Inhalt des Registers Z und das Ergebnis in das Register X gespeichert.

$$X \leftarrow Y \circlearrowleft Z$$

Die Rotation bedeutet, dass die Bits nach links verschoben werden und diejenigen Bits, die über die linke Grenze hinausfallen auf der rechten Seite wieder eingefügt werden. Dies bedeutet, dass mit jedem verschobenen Bit, ein Bit ganz links (Wertigkeit 2^{31}) fällt aus und wird wieder ganz rechts mit Wertigkeit 2^0 eingefügt. Nach 32 Rotationen eine Stelle nach links ist das ursprüngliche Bitmuster wieder hergestellt.

Beispiel Der folgende Code zeigt ein Beispiel für die Verwendung dieser Instruktion.

```
SET  R1 2
SET  R2 1
ROTL R3 R1 R2       # links-rotation von R1 eine Stelle
                    # R3 = 4
SET  R1 0x80000000 # nur das linke Bit gesetzt
ROTL R4 R1 R2       # R4 = 0x01
```

3.6.20 ROTLI

Assemblername	Parameter	Maschinencode	Format
ROTLI	$X, Y \in \mathcal{R}$, $N \in \mathbb{N}$	0x69	RRN

„Rotate Left Immediate“. Wie ROTL aber N ist eine konstante Zahl aus dem Intervall $[0, 255]$.

Beispiel Verwendung:

```
ROTL R4 R1  6 # rotiere die Bits in R1 6 stellen
ROTL R4 R1 -6 # Fehler! da −6 ∉ ℕ
```

3.6.21 ROTR

Assemblername	Parameter	Maschinencode	Format
ROTR	$X, Y, Z \in \mathcal{R}$	0x6A	RRR

„Rotate Right“. Funktioniert genauso wie die Instruktion ROTL mit dem Unterschied, dass die Rotationen (Verschiebungen) nach rechts geführt werden.

$$X \leftarrow Y \circlearrowright Z$$

3.6.22 ROTRI

Assemblername	Parameter	Maschinencode	Format
ROTRI	$X, Y \in \mathcal{R}$, $N \in \mathbb{N}$	0x6B	RRN

„Rotate Right Immediate“. Funktioniert genauso wie ROTLI, nur die Rotationen sind nach rechts geführt.

3.7 Vergleichsinstruktionen

Die Vergleichsinstruktionen vergleichen den Inhalt eines Registers mit dem Inhalt eines anderen Registers oder mit einer angegebenen ganzen Zahl. Das Ergebnis der Vergleichsinstruktion wird in das Register CMP gespeichert. Dieses Ergebnis ist -1, 0 oder $+1$ und wird folgenderweise berechnet: werden zwei Werte x und y vergliechen, so wird das Register CMP wie folgt gesetzt:

$$CMPR \leftarrow \begin{cases} -1 & \text{falls } x < y \\ \pm 0 & \text{falls } x = y \\ +1 & \text{falls } x > y \end{cases}$$

3.7.1 CMP

Assemblername	Parameter	Maschinencode	Format
CMP	$X, Y \in \mathcal{R}$	0x70	RR0

„Compare“. Vergleicht die Inhalte der Register X und Y.

3.7.2 CMPU

Assemblername	Parameter	Maschinencode	Format
CMPU	$X, Y \in \mathcal{R}$	0x71	RR0

„Compare Unsigned“. Vergleicht analog zu CMP – aber vorzeichenlos – X mit Y (die Inhalte der Register X und Y werden als vorzeichenlose Werte ausgewertet).

3.7.3 CMPI

Assemblername	Parameter	Maschinencode	Format
CMPI	$X \in \mathcal{R}$, $N \in \mathbb{Z}$	0x72	RNN

„Compare Immediate“. Vergleicht X mit angegebenen festen Wert N. Dabei nimmt N Werte aus dem Intervall $[-2^{15}, 2^{15} - 1]$, entsprechend dem Datentyp „Half“.

3.8 Übersprungsbefehle

Alle Sprungbefehle, außer dem GO Befehl, veranlassen einen relativen Übersprung im Programmcode – relativ im Sinne, dass die Parameter der Übersprungbefehle einen ganzzahligen Versatz zur aktuellen Programmadresse angeben. Dabei bedeutet der Versatz, wieviele Instruktionen müssen bis zur Zielinstruktion übersprungen werden. Z.B. die Instruktion

```
JMP 2
```

bedeutet „Überspringe 2 Instruktionen und fahre mit der 3. fort“. Die Instruktion

```
BE 5
```

bedeutet „Falls das Register CMP den Wert 0 hat, überspringe 5 Instruktionen und fahre mit der 6. fort“.

Der Versatz wird nicht in Bytes angegeben, sondern in Instruktionen – wobei eine Instruktion der UMach Maschine 4 Bytes beträgt.

Die Übersprungbefehle bauen auf die Vergleichsbefehle auf: jeder Übersprungbefehl (außer JMP und GO) untersuchen das Spezialregister CMP und verzweigen die Programmausführung anhand seines Wertes.

3.8.1 BE

Assemblername	Parameter	Maschinencode	Format
BE	$N \in \mathbb{Z}$	0x80	NNN

„Branch Equal“. Wenn das Register CMP den Wert 0 hat, wird über N Instruktionen vorwärts oder rückwärts gesprungen. Ein negatives N bedeutet einen Sprung rückwärts, ein positives N bewirkt einen Sprung vorwärts. Der Sprung wird dadurch erreicht, dass das Register PC gemäß der folgenden Formel modifiziert wird:

$$PC \leftarrow PC + 4 \cdot N$$

Die Multiplikation mit 4 wird deshalbt ausgeführt, weil ein Befehl immer aus 4 Bytes besteht, sodass der Adressoffset zwischen zwei Befehlen immer 4 ist. Somit ist N die Anzahl der zu überspringenden Befehle bis zur nächsten Instruktion. Der dazu benötigte Offset N wird vom Assembler automatisch berechnet.

Bemerkung Die UMach Maschine erhöht den Programmcounter (Register PC) nach der Ausführung jeder Instruktion (siehe 2.1.2, Seite 9). Die Modifizierung des PC-Registers durch den BE Befehl wirkt sich nicht störend auf die automatische Inkrementierung des Programmcounters.

Beispiel Der folgende Code lädt zwei Bytes in die Register $R2$ und $R3$ und addiert diese arithmetisch, falls sie ungleiche Werte haben. Sind die Werte gleich, wird stattdessen der Inhalt von $R2$ mit 2 multipliziert. Ein mögliches Überlaufen wird nicht berücksichtigt.

```
SET   R1 100     # R1 ← 100
LBUI  R2 R1  0   # Lade Byte von Adresse 100 nach R2
LBUI  R3 R1  1   # Lade Byte von Adresse 100+1 nach R3
CMPU  R2 R3      # Vergleiche Inhalt von R2 und R3
                 # Ergebnis geht ins CMP
#BE    equal     # Asm Schreibweise
BE    3          # Wenn CMP gleich 0 ist, überspringe
                 # die folgenden 3 Instruktionen
                 # (gehe zum label equal)
                 # N ist in diesem Fall 3
```

```
  ADDU  R2 R2 R3  # Addiere Inhalt von R2 und R3
  SBUI  R2 R1 0   # Speichere R2 nach Adresse 100
  #JMP   finish   # Asm Schreibweise
  JMP   2         # Überspringe die folgenden 2 Befehle,
                  # N von JMP ist 2.
equal:
  MULIU R2 2      # Multipliziere Inhalt von R2 mit 2
  SBUI  LO R1 0   # Speichere Inhalt von LO nach Adresse in R1
finish:
```

3.8.2 BNE

Assemblername	Parameter	Maschinencode	Format
BNE	$N \in \mathbb{Z}$	0x81	NNN

„Branch Not Equal “. Entspricht dem Verhalten von BE mit dem Unterschied, dass der angegebene Übersprung ausgeführt wird, wenn CMP nicht 0 ist.

3.8.3 BL

Assemblername	Parameter	Maschinencode	Format
BL	$N \in \mathbb{Z}$	0x82	NNN

„Branch Less“. Überspringt N Instruktionen, wenn der Inhalt von CMP kleiner 0 ist.

3.8.4 BLE

Assemblername	Parameter	Maschinencode	Format
BLE	$N \in \mathbb{Z}$	0x83	NNN

„Branch Less Equal“. Überspringt N Instruktionen, wenn der Inhalt von CMP kleiner oder gleich 0 ist.

3.8.5 BG

Assemblername	Parameter	Maschinencode	Format
BG	$N \in \mathbb{Z}$	0x84	NNN

„Branch Greater“. Überspringt N Instruktionen, wenn der Inhalt von CMP größer als 0 ist.

3.8.6 BGE

Assemblername	Parameter	Maschinencode	Format
BGE	$N \in \mathbb{Z}$	0x85	NNN

„Branch Greater Equal“. Überspringt N Instruktionen, wenn der Inhalt von CMP größer oder gleich 0 ist.

3.8.7 JMP

Assemblername	Parameter	Maschinencode	Format
JMP	$N \in \mathbb{N}$	0x88	NNN

„Jump“. Überspringt N Instruktionen.

3.9 Unterprogramminstruktionen

3.9.1 GO

Assemblername	Parameter	Maschinencode	Format
GO	$X \in \mathcal{R}$	0x90	R00

Setzt *PC* auf die angegebene absolute Adresse. Hierbei ist zu beachten, dass nicht in die Mitte eines Befehles gesprungen wird. Dies zu gewährleisten liegt in der Verantwortung des Programmierers.

3.9.2 `CALL`

Assemblername	Parameter	Maschinencode	Format
`CALL`	$N \in \mathbb{N}$	`0x91`	NNN

3.9.3 `RET`

Assemblername	Parameter	Maschinencode	Format
`RET`	keine	`0x92`	000

3.10 Systeminstruktionen

3.10.1 `INT`

Assemblername	Parameter	Maschinencode	Format
`INT`	$N \in \mathbb{N}$	`0xA0`	NNN

„Interrupt“. Ruft das Betriebssystem auf, eine Aktion zu unternehmen. Die Aktion wird als numerischer Code angegeben und ist systemspezifisch. Vergleichbar mit einem „syscall“.

3.10.2 `IRET`

Assemblername	Parameter	Maschinencode	Format
`IRET`	$N \in \mathbb{N}$	`0xA1`	NNN

„Interrupt Return“. Rücksprung aus einer Hardware-Unterbrechung. Setzt HIR zusätzlich auf Null und Setzt den Status so, dass Hardware-Unterbrechungen wieder erlaubt sind.

3.11 IO Instruktionen

3.11.1 IN

Assemblername	Parameter	Maschinencode	Format
IN	$R, P \in \mathcal{R}$	0xB0	RR0

Liest den Inhalt des Eingabeports mit Nummer P in das Register R. Diese Operation blockiert solange die Programausführung, bis Daten eingelesen wurden.

3.11.2 OUT

Assemblername	Parameter	Maschinencode	Format
OUT	$R, P, K \in \mathcal{R}$	0xB8	RRR

Sendet den Inhalt des Registers R zum Ausgangsport mit Nummer P. Dabei wird die Kennung der Daten aus R auf den Inhalt des Registers K gesetzt. Siehe auch den Abschnitt 2.5.2, auf der Seite 18.

$$R \rightarrow P$$

R, P und K sind alle Register.

Beispiel Im folgenden Beispiel wird den Inhalt des Registers $R1$ an verschiedenen Ports ausgegeben. Die Register R2 und ZERO werden dabei für die Kennung und Portnummer verwendet.

```
OUT R1 ZERO ZERO # R1 auf Port 0 mit Kennung 0 ausgeben
SET R2 5        # port 5

SET R1 10       # Inhalt in R1 ist newline
OUT R1 R2   ZERO # newline zum Port 5, Kennung 0
OUT R1 R2   R2   # nochmal, mit Kennung 5 auf Port 5
```

4 Debugging

Die UMach Maschine stellt eine Debugging-Schnittstelle zur Verfügung. Um diese Schnittstelle zu aktivieren, muss man die Maschine im Einzelschrittmodus starten.

Glossar

A | B | I | R | S

A

Adressdekoder
: Das Herz des Bussystems. Wählt anhand einer eingegebenen Adresse einen Bus-Port, an den ein Befehl weitergeleitet wird. Funktioniert wie eine Art Demultiplexer.

Adressierungsart
: Die Art, wie eine Instruktion die Umach Maschine dazu veranlasst, einen Speicherbereich zu adressieren. Siehe auch Abschnitt 2.4.1.

Assemblername
: Der Name eines Registers oder eines Befehls, so wie er in einem textuellen Programm (ASCII) verwendet wird. $R1$, $R2$, ADD sind Assemblernamen von Registern und Befehlen.

B

Befehl
: Die ersten 8 Bits in einer Instruktion. Operation code.

Befehlsraum
: Die Anzahl der möglichen Befehle, abhängig von der Befehlsbreite. Beträgt die Befehlsbreite 8 Bit, so ist der Befehlsraum $2^8 = 256$.

Betriebsmodus
: Die Art, wie die UMach Maschine die einzelnen Instruktionen abarbeitet. Siehe auch 2.1.1.

Bussystem
: Die Mainboard.

Byte
Eine Reihe oder Gruppe von 8 Bit.

I

Instruktion
Eine Anweisung an die UMach VM etwas zu tun. Eine Instruktion besteht aus einem Befehl (Operation Code) und eventuellen Argumenten.

Instruktionsformat
Beschreibt die Struktur einer Instruktion auf Byte-Ebene und zwar es gibt an, ob ein Byte als eine Registerangabe oder als reine numerische Angabe zu interpretieren ist. Siehe 3.1.

Instruktionssatz
Die Menge aller Instruktionen, die von der UMach Maschine ausgeführt werden können.

R

Register
Eine sich im Prozessor befindende Speichereinheit. Das Register ist dem Programmierer sichbar und kann mit Werten geladen werden. Siehe Abschnitt 2.3, Seite 11.

Registernummer
Eine eindeutige Zahl, die ein Register identifiziert. Eine Instruktion verwendet diese Zahl, wenn sie ein Register angibt.

S

Speicher
Baukomponente, die am Bussystem angeschlossen ist und die eine feste Anzahl von Bytes für die Laufzeit der Maschine aufnehmen kann.

Index