The Golden Age of Computer Architecture? Computer Architecture is just one small piece of the answer

Yale N. Patt

The University of Texas at Austin

First presented: August 16, 2019

Abstract

People talk about the next chapter in Computer Architecture as the Golden Age. Most of it, standard, mindless hype. Everywhere I turn I see the next chapter requiring more and more things other than computer architecture. Or, more correctly put, from things that need to be combined with computer architecture. The imminent end of Moore's Law suggests no more *oomph* at the bottom. But we do still need to pay attention to the bottom. And, if there are limits to what we can do at the bottom, clearly we will have to use the plenty of room at the top. Processor paradigms will need to encompass that. Many suggest the von Neumann model is dead. That suggestion is misguided. Tools, important now, will be even more important moving forward. Education needs a whole new rethinking. If anything is essential to guiding us through all this, it is the transformation hierarchy. I will try to explain.

Abstract

People talk about the next chapter in Computer Architecture as the Golden Age. Most of it, standard, mindless hype. Everywhere I turn I see the next chapter requiring more and more things other than computer architecture. Or, more correctly put, from things that need to be combined with computer architecture. The imminent end of Moore's Law suggests no more *oomph* at the bottom. But we do still need to pay attention to the bottom. And, if there are limits to what we can do at the bottom, clearly we will have to use the plenty of room at the top. Processor paradigms will need to encompass that. Many suggest the von Neumann model is dead. That suggestion is misguided. Tools, important now, will be even more important moving forward. Education needs a whole new rethinking. If anything is essential to guiding us through all this, it is the transformation hierarchy. I will try to explain.

Abstract

People talk about the next chapter in Computer Architecture as the Golden Age. Most of it, standard, mindless hype. Everywhere I turn I see the next chapter requiring more and more things other than computer architecture. Or, more correctly put, from things that need to be combined with computer architecture. The imminent end of Moore's Law suggests no more *oomph* at the bottom. But we do still need to pay attention to the bottom. And, if there are limits to what we can do at the bottom, clearly we will have to use the plenty of room at the top. Processor paradigms will need to encompass that. Many suggest the von Neumann model is dead. That suggestion is misguided. Tools, important now, will be even more important moving forward. Education needs a whole new rethinking. If anything is essential to guiding us through all this, it is the transformation hierarchy. I will try to explain.

Abstract

People talk about the next chapter in Computer Architecture as the Golden Age. Most of it, standard, mindless hype. Everywhere I turn I see the next chapter requiring more and more things other than computer architecture. Or, more correctly put, from things that need to be combined with computer architecture. The imminent end of Moore's Law suggests no more *oomph* at the bottom. But we do still need to pay attention to the bottom. And, if there are limits to what we can do at the bottom, clearly we will have to use the plenty of room at the top. Processor paradigms will need to encompass that. Many suggest the von Neumann model is dead. That suggestion is misguided. Tools, important now, will be even more important moving forward. Education needs a whole new rethinking. If anything is essential to guiding us through all this, it is the transformation hierarchy. I will try to explain.

Abstract

People talk about the next chapter in Computer Architecture as the Golden Age. Most of it, standard, mindless hype. Everywhere I turn I see the next chapter requiring more and more things other than computer architecture. Or, more correctly put, from things that need to be combined with computer architecture. The imminent end of Moore's Law suggests no more *oomph* at the bottom. But we do still need to pay attention to the bottom. And, if there are limits to what we can do at the bottom, clearly we will have to use the plenty of room at the top. Processor paradigms will need to encompass that. Many suggest the von Neumann model is dead. That suggestion is misguided. Tools, important now, will be even more important moving forward. Education needs a whole new rethinking. If anything is essential to guiding us through all this, it is the transformation hierarchy. I will try to explain.

Abstract

People talk about the next chapter in Computer Architecture as the Golden Age. Most of it, standard, mindless hype. Everywhere I turn I see the next chapter requiring more and more things other than computer architecture. Or, more correctly put, from things that need to be combined with computer architecture. The imminent end of Moore's Law suggests no more *oomph* at the bottom. But we do still need to pay attention to the bottom. And, if there are limits to what we can do at the bottom, clearly we will have to use the plenty of room at the top. Processor paradigms will need to encompass that. Many suggest the von Neumann model is dead. That suggestion is misguided. Tools, important now, will be even more important moving forward. Education needs a whole new rethinking. If anything is essential to guiding us through all this, it is the transformation hierarchy. I will try to explain.

Abstract

People talk about the next chapter in Computer Architecture as the Golden Age. Most of it, standard, mindless hype. Everywhere I turn I see the next chapter requiring more and more things other than computer architecture. Or, more correctly put, from things that need to be combined with computer architecture. The imminent end of Moore's Law suggests no more *oomph* at the bottom. But we do still need to pay attention to the bottom. And, if there are limits to what we can do at the bottom, clearly we will have to use the plenty of room at the top. Processor paradigms will need to encompass that. Many suggest the von Neumann model is dead. That suggestion is misguided. Tools, important now, will be even more important moving forward. Education needs a whole new rethinking. If anything is essential to guiding us through all this, it is the transformation hierarchy. I will try to explain.

Abstract

People talk about the next chapter in Computer Architecture as the Golden Age. Most of it, standard, mindless hype. Everywhere I turn I see the next chapter requiring more and more things other than computer architecture. Or, more correctly put, from things that need to be combined with computer architecture. The imminent end of Moore's Law suggests no more *oomph* at the bottom. But we do still need to pay attention to the bottom. And, if there are limits to what we can do at the bottom, clearly we will have to use the plenty of room at the top. Processor paradigms will need to encompass that. Many suggest the von Neumann model is dead. That suggestion is misguided. Tools, important now, will be even more important moving forward. Education needs a whole new rethinking. If anything is essential to guiding us through all this, it is the transformation hierarchy. I will try to explain.

Abstract

People talk about the next chapter in Computer Architecture as the Golden Age. Most of it, standard, mindless hype. Everywhere I turn I see the next chapter requiring more and more things other than computer architecture. Or, more correctly put, from things that need to be combined with computer architecture. The imminent end of Moore's Law suggests no more *oomph* at the bottom. But we do still need to pay attention to the bottom. And, if there are limits to what we can do at the bottom, clearly we will have to use the plenty of room at the top. Processor paradigms will need to encompass that. Many suggest the von Neumann model is dead. That suggestion is misguided. Tools, important now, will be even more important moving forward. Education needs a whole new rethinking. If anything is essential to guiding us through all this, it is the transformation hierarchy. I will try to explain.

Abstract

People talk about the next chapter in Computer Architecture as the Golden Age. Most of it, standard, mindless hype. Everywhere I turn I see the next chapter requiring more and more things other than computer architecture. Or, more correctly put, from things that need to be combined with computer architecture. The imminent end of Moore's Law suggests no more *oomph* at the bottom. But we do still need to pay attention to the bottom. And, if there are limits to what we can do at the bottom, clearly we will have to use the plenty of room at the top. Processor paradigms will need to encompass that. Many suggest the von Neumann model is dead. That suggestion is misguided. Tools, important now, will be even more important moving forward. Education needs a whole new rethinking. If anything is essential to guiding us through all this, it is the transformation hierarchy. I will try to explain.

Abstract

People talk about the next chapter in Computer Architecture as the Golden Age. Most of it, standard, mindless hype. Everywhere I turn I see the next chapter requiring more and more things other than computer architecture. Or, more correctly put, from things that need to be combined with computer architecture. The imminent end of Moore's Law suggests no more *oomph* at the bottom. But we do still need to pay attention to the bottom. And, if there are limits to what we can do at the bottom, clearly we will have to use the plenty of room at the top. Processor paradigms will need to encompass that. Many suggest the von Neumann model is dead. That suggestion is misguided. Tools, important now, will be even more important moving forward. Education needs a whole new rethinking. If anything is essential to guiding us through all this, it is the transformation hierarchy. I will try to explain.

Problem

Algorithm

Program

ISA (Instruction Set Arch)

Microarchitecture

Circuits

Electrons

What I want to do today

- ***The transformation hierarchy***
- ***Moore's Law will end, the von Neumann model won't***
- ***Room at the bottom***
- ***Room at the top***
- ***Room straddling both***
- ***What will we need to make it happen***

What I want to do today

- ***The transformation hierarchy***
- ***Moore's Law will end, the von Neumann model won't***
- ***Room at the bottom***
- ***Room at the top***
- ***Room straddling both***
- ***What will we need to make it happen***

Room at the Bottom

- *The run-time system (because latency is important)*
- *Dark silicon (not everything is needed all the time)*
- *The silliness of a multiple-ISA, multi-core chip*
  - *Endianness, page size, data types*
  - *The good sense of a single-ISA, multiple-uarch cores*
- *SSMT (helper threads for improving infrastructure)*
  - *Also, recompile*
- *ILP is not dead (contrary to some misguidance)*
- *New materials → new tools*
- *Asynch for a few cycles, then synch*

What I want to do today

- ***The transformation hierarchy***
- ***Moore's Law will end, the von Neumann model won't***
- ***Room at the bottom***
- ***Room at the top***
- ***Room straddling both***
- ***What will we need to make it happen***

Room at the Top

- ***Charles Leiserson and his MIT colleagues***
  - ***Algorithms, Languages***
- ***Two interfaces***
  - ***One for those who understand microarchitecture***
  - ***One for those who don't***
  - ***Tools to bridge the gap***

What I want to do today

- ***The transformation hierarchy***
- ***Moore's Law will end, the von Neumann model won't***
- ***Room at the bottom***
- ***Room at the top***
- ***Room straddling both***
- ***What will we need to make it happen***

Most *importantly: Room Straddling Both*

- ***Processor Paradigms***
  - ***Why? They yield higher performance, better energy needs***
  - ***From yesterday: HPS, GPU***
  - ***For tomorrow: Accelerators of all sorts***
- ***What's involved***
  - ***Lots of domain specific***
  - ***We will need to keep von Neumann***
  - ***We will have to break the layers***
  - ***We will have to deal with portability***

Some Paradigms

- ***Tomasulo (what was good, what was bad)***
- ***Data Flow (what was good, what was bad)***
- ***HPS***
- ***CDC6600***
- ***HEP (what was good, what was the problem)***
- ***SMT (what was good, what was bad)***
- ***SIMD***
- ***GPU***
- ***Systolic Array***
- ***Spatial Computing***
- ***Non-Von, BVM, Connection Machine***
- ***Multi-core***
- ***User-writeable Control Store***

Examples of Paradigm Evolution

- ***HPS***
- ***SMT***
- ***GPU***
- ***VLIW***
- ***Spatial Computing***
- ***Accelerators***

HPS

- ***Tomasulo + Data Flow → HPS***
- ***Tomasulo had out-of-order, NOT precise exceptions***
  - ***Also, ONLY the floating point unit***
  - ***Also, ONLY one operation per instruction***
  - ***Also, Stall on a branch (no steady supply of operations)***
- ***Data Flow had micro-ops, but too unwieldly***
  - ***Hard to take interrupts***
  - ***Hard to debug***

HPS took the good, added: in-order retirement,

Restricted window, wide issue, aggressive br.predictor

SMT

- ***HEP + ooo + wide issue → SMT***
- ***HEP was brilliant, ahead of its time (SPIE 1977)***
  - ***But issued only one instruction each clock cycle***
- ***Actually, CDC6600 → HEP***
- ***SMT (Hirata, ISCA 1992, Nemirovsky 1994, UW 1995)***

GPU

- ***SIMD + SMT + Predicated Execution → GPU***

- ***If the software can pay attention to branches***
- ***If the software can organize memory accesses***

VLIW

- ***Horizontal microcode → VLIW***
- ***Not good for General Purpose Computing***
- ***But good for domain specific stuff***
  - ***Microcode Emulation of the ISA***
  - ***DSP chips***
- ***i.e., when the software is known in advance***

Spatial Computing

- ***Systolic Array + FPGAs → Spatial Computing***
- ***HT Kung (1979): not enough transistors, too "asic"***
- ***Today, stream data through a data flow graph***
- ***If the software can produce the data flow graph***

The future: Accelerators of all kinds

- ***Many implementation mechanisms***
  - ***ASICs***
  - ***FPGAs***
  - ***EMT instruction (with writeable control store)***
- ***Examples (Quantum computing, Machine learning)***
- ***Requires the attention of***
  - ***The person writing the algorithm***
  - ***The programmer***
  - ***The compiler writer***
  - ***The microarchitect***
- ***Straddling the layers? -- Break the layers!!!***

Accelerators will play a big role

- ***They will need CPUs to control them***
- ***Which means Von Neumann will continue to thrive!***

The Von Neumann Paradigm

- ***A straightforward model of executing programs***
  - ***Fetch, Decode, Evaluate address, Fetch Data, …***
- ***Many have suggested its demise***
  - ***Wrong!***
  - ***Our best mechanism for maintaining order***
- ***But it will be augmented with accelerators***
  - ***The pervasive computing engine of the future***

But the Best Accelerators Destroy Portability

- *Companies find that unacceptable*
- *i.e., the economics of allowing non-portability*

Without portability:

- ***The economics isn't there;***
- ***Industry won't buy into it.***
- ***I have two responses:***
  - ***Economics Be Damned!***
  - ***More soberly: Tools to the rescue***

I believe there are things we need to tackle in spite of the apparent economics

- ***Things that should not have a price tag***
  - ***Like curing cancer***
  - ***Predicting tsunamis***
  - ***Driverless cars that don't crash into walls***
  - ***Or lots of other things we will "trust" computers to do.***
- ***Things that need much greater performance***
- ***We already have existence proofs***
  - ***US DOE's Supercomputers for controlling nuclear plants***
  - ***David Elliott Shaw's ANTON for understanding science***
  - ***Google's recent Tensor Processing Unit and follow-ons***
- ***What else if we reject a quick payoff***

What I want to do today

- ***The transformation hierarchy***
- ***Moore's Law will end, the von Neumann model won't***
- ***Room at the bottom***
- ***Room at the top***
- ***Room straddling both***
- ***What will we need to make it happen***

What we will need to make it happen

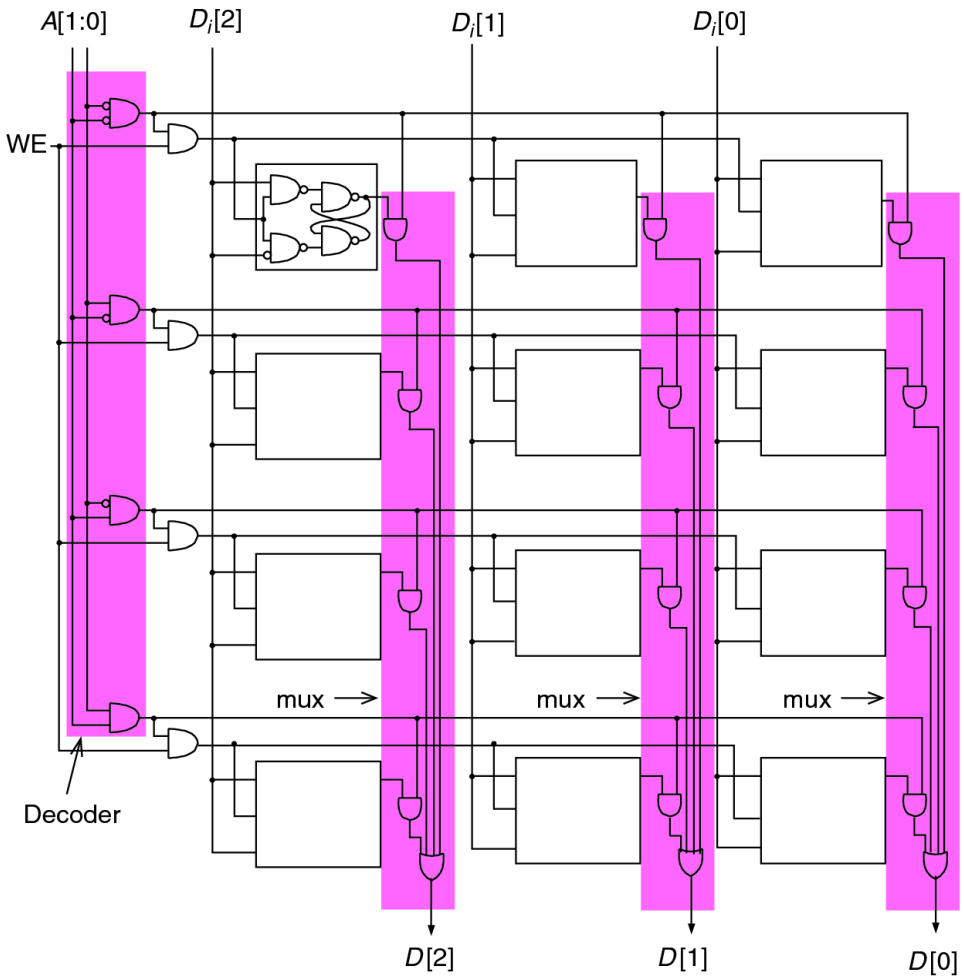
- ***New tools**, as we have said again and again!*
- ***Education**: the motivated bottom-up approach.*

Education: What do we do?

- ***We start in the freshman year***
- ***Start with what they "know"***
  - ***The transistor as light switch***
  - ***Not quantum mechanics***
- ***Choose a computer model that is simple***
  - ***As the genius said: simple, but still rich***
- ***Continually build on what they know***
- ***Continually raising the level of abstraction***
- ***Memorizing as little as absolutely necessary***
- ***Trying very hard to not introduce magic***

We start with light switches, and quickly move to transistors (p-type and n-type)

- *Then gates (inverter, NAND, NOR, …)*
- *Then combinational structures (MUX, Decoder, ALU)*
- *Then latches, flip flops*
- *Then memory*
- *Then finite state machines*
- *Then a computer (LC-3), specially invented*
- *Then programming, debugging in machine language*
- *Then assembly language*
- *Then data structures*
- *Then physical I/O, and interrupts*
- *Then programming in C, C++*


A[1:0]
Di[2]
Di[1]
Di[0]
WE
mux
mux
mux
Decoder
D[2]
D[1]
D[0]

The ISA

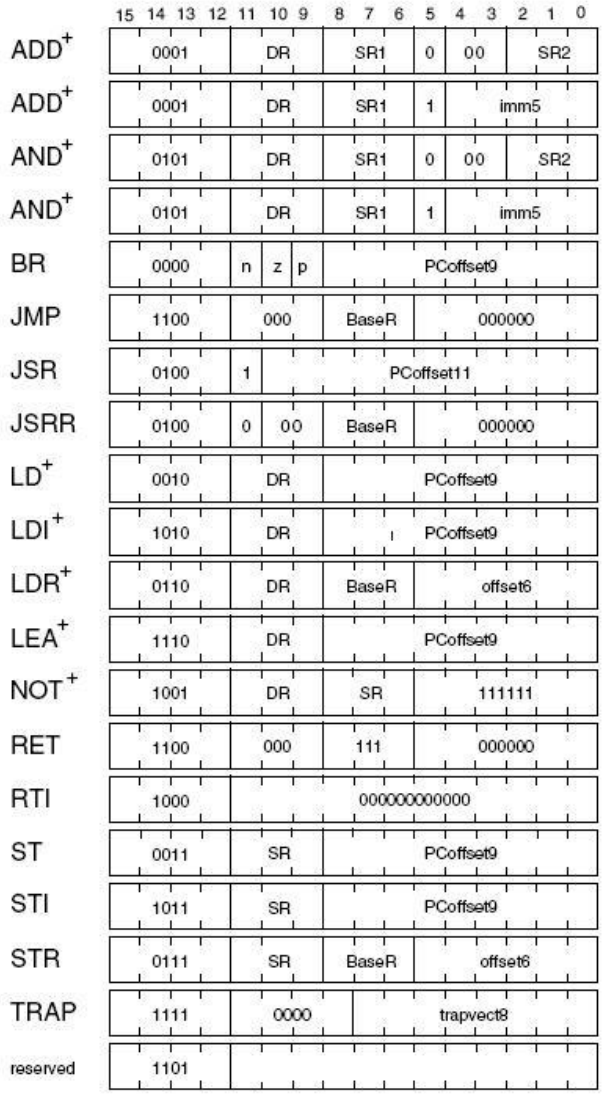


Figure A.2 Format of the entire LC-3 instruction set. Note: + indicates instructions that modify condition codes

The Data Path

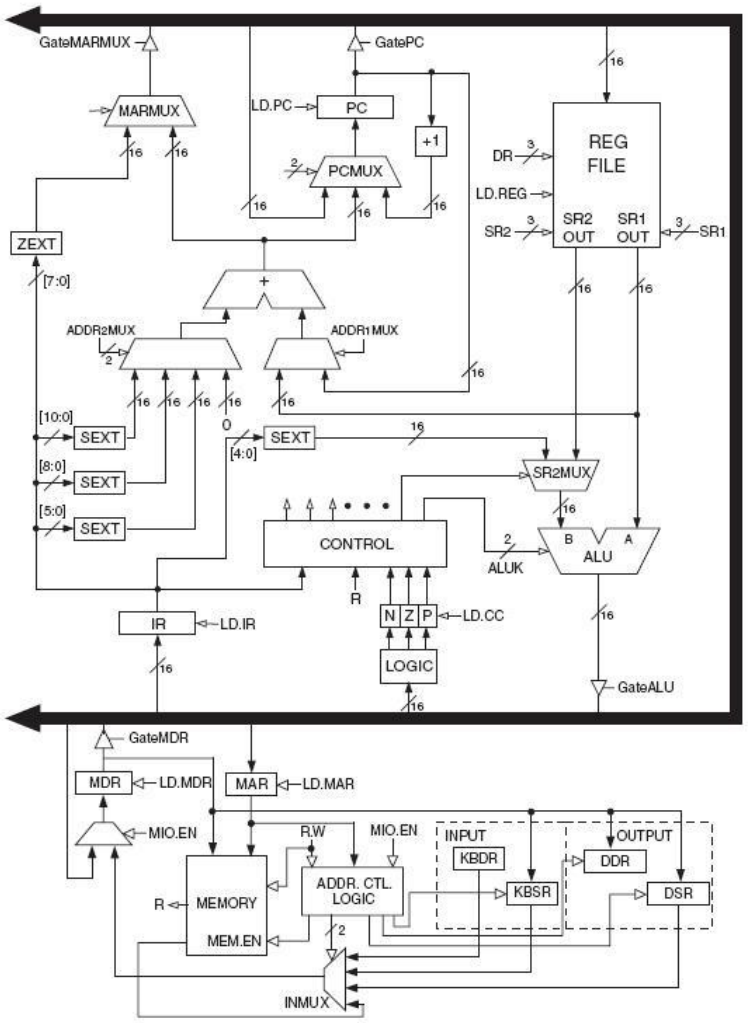


Figure C.3 The LC-3 data path

The State Machine

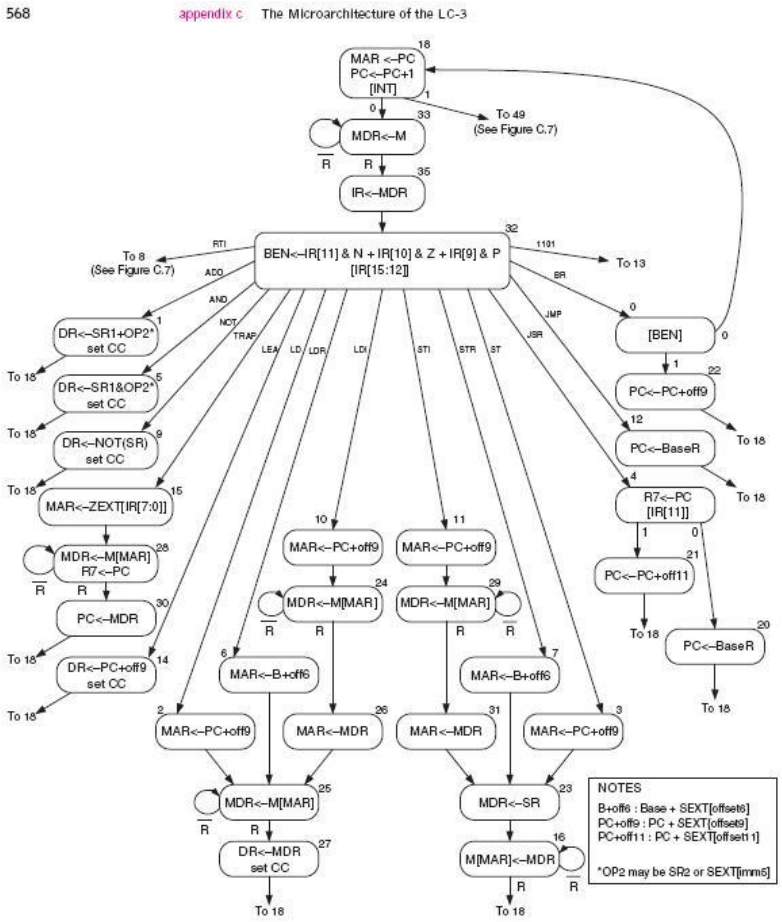


Figure C.2 A state machine for the LC-3

The result:

The graduate is ready

To straddle the layers

So,

- ***Moore's Law will end***
- ***But there is good reason to be hopeful***
- ***There is room at the bottom and the top***
- ***If we break the layers, more room straddling both***
- ***We will need to keep the von Neumann model***
- ***We will need new tools***
- ***And we will need to do something about Education***

Thank you!