Computer Science 2.0

Algorithms = Logic + Control
Programs = Algorithms + Data-Structures

Programs = Logic + Data

Vijay Saraswat
May 12, 2021

All views and opinions expressed are mine, and not those of my employer. They are for informational purposes only. All examples noted are from public documents.
Agenda

- Computer Science 1.0, and 2.0
  - Deep Learning as a transformational set of ideas
- Documents in Finance
- (Briefly) Quantitative Finance (QF)
- The Big Challenges
Computer Science 1.0

Algorithms = Logic + Control
Programs = Algorithms + Data-Structures

Human-generated Code

Programming Models

Assembly
Fortran
Pascal, C
C++, Java
Scala
Rust

Lisp
Prolog
ML
Esterel
X10
Go
Python

Declarative v Prescriptive
Correctness – Type systems, Verification
Transformational v Reactive
Sketching

“If you know how to do it, its not AI.”
Transformational vs Reactive

**Agent**

```
int W=8, logW = 3;
    bit[W] I = W;
    while (ret > 0) {
        i--;
        if (in[i]) break;
    }
    return i;
}
```

**World**

```
loop {
    m = 0 ;
    trap (WAIT ) {
        loop {
            present N then m += 5;
            present D then m += 10;
            if (m >= 15) then exit WAIT;
            pause
        }
        emit GUM; pause
    }
}
```

- **One and done**
- **Maintain continuous interaction**
Sketching (paraphrasing [3]): set of techniques for generating code based on artifacts that establish syntactic or semantic properties of the code.

1. Sketching (paraphrasing [3]): set of techniques for generating code based on artifacts that establish syntactic or semantic properties of the code.

```
int W=8, logW = 3;
    bit[W] i = W;
    while (ret > 0) {
        i--;
        if (in[i])
            break;
    }
    return i;
}
```

```
int W=8, logW = 3;
    bit[W] ret = 0;
    loop (logW) {
        if (in & ??) {
            in >>= ??;
            ret |= ??
        }
        if (in & 0xf0) { in >>= 4; ret |= 4; }
        if (in & 0xc0) { in >>= 2; ret |= 2; }
        if (in & 0x02) { in >>= 1; ret |= 1; }
    }
    return ret;
}
```

Computer Science 2.0

Programs = Logic + Data

Data is noisy, uncertain, high-dimensional [1]

We move to compute with continuous functions.

Can we synthesize the code for the agent, simply given (a lot of) examples of input/output pairs?

What if we don’t know how to write the code?

Deep Representations

• (Parametrized) Function Approximation
• 2 layers of neurons (linear weights plus non-linear activation) are adequate to represent any function
• But: Deep layers provide an exponential advantage
• Distributed representations
• Over-parametrized -- \( \sim 1.6t! \)

Let the data speak!

Stochastic Gradient Descent

- [1] Find the minimum of a function by moving in the direction of the steepest descent.
- Not guaranteed to work for non-convex problems.
- But: [2] Large, multi-layer networks have many equivalent local minima
- [3] For over-parametrized n/w SGD can find global minima in polynomial time.

cf constrained optimization

The Financial Industry Runs on Documents...

Market Making
- Pricing
- Trading

Mergers & Acquisitions
- Marking
- Monitoring

Risk Management
- Risk Assessment
- Renegotiation

Investment Management
- NAV extraction
- Investing

Operational Review
- Rolling reviews
- Onboarding
- Reconciliation

Compliance & Legal
- Due diligence
- Regulatory
- Negotiation

ISDA
Schedule and CSA Term Extraction

COB
Org, Alias, Address Extraction

Due Diligence
Cap, Auth Evidence Extraction

Fin Statements
Identify, extract values, normalize terms

NAV
Extract from email, attachments, spreadsheets, tables

There are scores of different document types across multiple finance functions. Our goal is for code to understand and use text just as it uses structured data today.
Working with Text

- Set up a path between any pair of tokens in a layer, to learn the strength of that pair.
- Build up these patterns recursively, layer on layer.
- Self-attention supports permutation-equivariance
  - Key to compositional generalization? (cf Lake and Baroni, ICML 2018).

Breakthrough in NLU

Identifying invariants and building them into the network is key to “programming.”

Vaswani et al. “Attention is all you need”, 2017
A (brief!!) review of SoTA in NL

- Field focused primarily on understanding mass media
  - Web, Wikipedia, news articles
  - (Though also a bit in Enterprise AI -- Health-care, RFC response...)

- Driven, in part, by corporate interests of / funding from consumer companies, reflecting their strategic priorities

- Techniques: Significant recent advances based on learnt, high-parameter, non-linear function approximation: language models, recurrent/recursive techniques, Seq2Seq, attention (“is all there is”)

- Tasks: information retrieval, (POS) tagging, parsing, text simplification, paraphrase, summary, question answering / reasoning ...
Overview of challenges (and tasks)

### New tasks
- Information Extraction
- Sentence simplification
- Semantic parsing

### Old tasks
- Document Consolidation
- Template extraction from doc corpus
- Formula extraction
- Spreadsheet (w formulas) extraction (from pdf tables)
- Professional question answering – meaningful vs literal reasoning

| 1. | Background theory – taken as given by professionals |
| 2. | Complex text – very low Flesch scores, multi-paragraph sentences (!). |
| 3. | Careful, precise, usually well-crafted language |
| 5. | Text with explicit and implicit (inline) definitions |
| 6. | Constructed text (amendments) – complex references, text-deletion / replacement, construction. |
| 7. | “Deontic irrealis” mood (cf Leora Morgenstern) – hypothetical event between generic parties; complex temporal constructs, normative definitions |
| 8. | Disjunctive, Conditional definitions, with multiple levels of nested exceptions |
| 9. | English verbalizations of mathematical expressions |
| 10. | Sophisticated (comparative) chains of argumentation |
| 11. | (Financial) Table understanding, comparison |

| 1. | Derivates agreements / credit support annexes |
| 2. | Loan / corporate credit agreements |
| 3. | Analyst / management reports |

### Different corpora

| 1. Derivates agreements / credit support annexes |
| 2. Loan / corporate credit agreements |
| 3. Analyst / management reports |

### New tasks
- Document Consolidation
- Template extraction from doc corpus
- Formula extraction
- Spreadsheet (w formulas) extraction (from pdf tables)
- Professional question answering – meaningful vs literal reasoning
Achieving professional facility in deep domains

- Value extraction
- Formula extraction
- Template identification
- Ontology construction
- Text consolidation
- Report analysis
- Answering questions

Complex (new) tasks over [long] documents with exploitable structure

- **Text is not text!**
ISDA 2002 MASTER AGREEMENT

Exhibit 10.1

ISDA
International Swaps and Derivatives Association, Inc.

2002 MASTER AGREEMENT
dated as of March 22, 2011

Bank of America, N.A. and LKQ Corporation

have entered and/or anticipate entering into one or more transactions (each a “Transaction”) that are or will be governed by this 2002 Master Agreement, which includes the schedule (the “Schedule”), and the documents and other confirming evidence (each a “Confirmation”) exchanged between the parties or otherwise effective for the purpose of confirming or evidencing those Transactions. This 2002 Master Agreement and the Schedule are together referred to as this “Master Agreement”.

Accordingly, the parties agree as follows:

1. Interpretation
   (a) Definitions. The terms defined in Section 14 and elsewhere in this Master Agreement will have the meanings therein specified for the purpose of this Master Agreement.
   (b) Inconsistency. In the event of any inconsistency between the provisions of the Schedule and the other provisions of this Master Agreement, the Schedule will prevail. In the event of any inconsistency between the provisions of any Confirmation and this Master Agreement, such Confirmation will prevail for the purpose of the relevant Transaction.
   (c) Single Agreement. All Transactions are entered into in reliance on the fact that this Master Agreement and all Confirmations form a single agreement between the parties (collectively referred to as this “Agreement”), and the parties would not otherwise enter into any Transactions.

2. Obligations
   (a) General Conditions.

Copyright © 2002 by International Swaps and Derivatives Association, Inc.
Section 1. Definitions. Capitalized terms used in this Agreement, including in its preamble and recitals, shall have the following meanings:

(a) “Affiliation Agreement” means the Affiliation Agreement, dated as of the date hereof, between Total Gas & Power USA, SAS and the Company.

(b) “Aggregate L/C Amount” means, as of any time, the sum, calculated on a Dollar-Equivalent Basis, of (i) the aggregate amount then-available to be drawn under all L/Cs issued under any Guaranteed Facility, (ii) the then-remaining amount of L/Cs available to be issued under any Guaranteed Facility (based on the maximum aggregate amount of L/Cs that could from time to time in the future be issued under any such Guaranteed Facility), and (iii) the aggregate amount of draws (including accrued but unpaid interest thereon) on any L/Cs issued under any Guaranteed Facility that have not yet been reimbursed by the Company to either (x) the applicable Bank or (y) the Guarantor (following a payment by the Guarantor to the Bank pursuant to a Guaranty).

(c) “Agreement” has the meaning given in the Preamble.

(d) “Annual Operating Plan” means, for any fiscal year, the projected income statement, cash flow statement and balance sheet of the Company broken out by quarter for such fiscal year and approved by the Company’s Board of Directors following its review of supporting material such as operational metrics (including regional MW, ASPs and COGS) and credit support requirements.

(e) “Assignment Fee” means a fee, equal to $20 million as of January 1, 2014, and reduced by $2 million per calendar quarter until such Assignment Fee is reduced to zero. As an example, the Assignment Fee that would be payable in connection with an assignment that occurred on October 15, 2014, would be $14 million.

Agreements may have dozens of pages of (mutually recursive) definitions.

Definition may define (a fluent using) a nested formula.

References can be at abstract levels: has the meaning given in the Preamble.

What does a language model mean for text with nested scopes?
Deconstruct document as a graph over (titled) stanzas.

Stanza: Smallest directly addressable contiguous passage

Inline or out-of-line

Title: heading

Links:
- Reading order
- Section reference
- Named reference (defined terms)
“Maturity Date” shall mean the Stated Maturity Date, provided that
(a) in the event of the exercise by Borrower of the First Extension Option pursuant to Section 2.7, the Maturity Date shall be the First Extended Maturity Date,
(b) in the event of the exercise by Borrower of the Second Extension Option pursuant to Section 2.7, the Maturity Date shall be the Second Extended Maturity Date,
(c) in the event of the exercise by Borrower of the Third Extension Option pursuant to Section 2.7, the Maturity Date shall be the Third Extended Maturity Date, and
(d) in the event of the exercise by Borrower of the Fourth Extension Option pursuant to Section 2.7, the Maturity Date shall be the Fourth Extended Maturity Date,
or such earlier date on which the final payment of principal of the Note becomes due and payable as herein and therein provided, whether at the Stated Maturity Date, by declaration of acceleration, extension or otherwise.

What are all the possible values for the Maturity Date?

Complex nested exception conditions
(c) Any overdue principal of, or interest on, any Euro-Dollar Loan shall bear interest, payable on demand, for each day until paid at a rate per annum equal to the higher of:

(i) the sum of 2% plus the Euro-Dollar Margin for such day plus the quotient obtained (rounded upward, if necessary, to the next higher 1/100 of it) by dividing (x) the average rate per annum (rounded upward, if necessary, to the next higher 1/100 of 1%) of the respective rates per annum at which one day (or, if such amount due remains unpaid more than three Business Days, then for such other period of time not longer than three months as the Administrative Agent may select) deposits in Dollars in an amount approximately equal to such overdue payment due to the Administrative Agent is offered to the Administrative Agent in the London interbank market for the applicable period determined as provided above by (y) one minus the Reserve Percentage (or, if the circumstances described in clause (a) or (b) of Section 8.1 shall exist, at a rate per annum equal to the sum of 2% plus the rate applicable to Base Rate Loans for such day) and

(ii) the sum of 2% plus the Euro-Dollar Margin for such day plus the London Interbank Offered Rate applicable to such Loan at the date such payment was due.

\[
\begin{align*}
i &= 0.02 + v(\text{Euro-Dollar-Margin}) + \operatorname{cei}(x/y, 0.01) \\
x &= v(\text{avg rate per annum for one day deposits in Dollars offered to AA in the London interbank market}) \\
y &= (1 - (0.02 + v(\text{Base Rate Loan}))) \text{ if } v(\text{Section 8.1 (a) or (b) holds}) \text{ else } (1 - v(\text{Reserve Percentage})) \\
i &= 0.02 + v(\text{Euro-Dollar Margin}) + v(\text{London Interbank Offered Rate}) \\
\text{max} (i, ii)
\end{align*}
\]

- Terms need to be normalized, e.g. \(v(\text{London Interbank Offered Rate}) = \text{LIBOR}\)
- Some terms are document specific
- Note: significant portions of text are not relevant to formula extraction.
Construct the basic rate table for all loan facilities. (row = type of facility, currency, underlying interest rate,

Coordinate multiple structured extractions
- Number of “rows” unknown – dependent on parsing text in the doc
- Number of (relevant) passages unknown
- Location of passages unknown
- May need to follow links
- Need to handle disjunction and conditionals
Template Identification

- Given a corpus of documents of the same genre (e.g. loan agreement)
- Represent them as a collection of templates + variations
  - Deletions, Additions, Replacements
- Doc(i) = (Template(j) + Variations)
- Perhaps multiple tiers of templates
  - Identify “background” or “understood” assumptions
- (Dual of “Text infilling”, [Xing et al 2018])

Life-cycle of a loan document:

Company A:
- \( \rightarrow S: \) (biz people) ~ 2 pages
- \( \rightarrow D: \) (legal) ~ 200 pages

-----

Company B:
- \( \rightarrow D: \) (legal) ~ 200 pages
- \( \rightarrow S': \) (legal) ~ 2 pages

Idea: Learn \( S \rightarrow D \) and \( D \rightarrow S \) translator simultaneously, cf recent FB work on unsupervised translators
Text Consolidation

• An agreement may be long-lived, and may have multiple amendments
  • ~ 20 year loan agreement (~200 pages)
  • Sixteen amendments (each multiple pages)

• Construct current applicable agreement – “Consolidated Agreement”

Identify, interpret and apply changes to original document
Maintain chain of changes (provenance)
Analyze Analyst Reports

Strong Q1 FY14 & Multiple Catalysts Ahead

- **Q1 Results:** Jazz Pharmaceuticals reported strong Q1 results on Thursday, with revenue of $247 million vs. $196 million in Q1 of FY 13, representing 26% YoY growth. The GAAP Net Loss was $93 million vs. GAAP Net Income of $43 million in Q1 of FY 13, but $127 million of that was due to an upfront license fee and milestone payment for JZP-110. Adjusted Net Income was $101 million vs. $84 million in Q1 of FY 13. These figures were slightly ahead of Goldman Stanley estimates of $240 million in revenue and $95 million in Adjusted Net Income.

- **FY 14 Guidance:** The Company reaffirmed its FY 14 guidance across the board, with revenue of $1.10 billion – $1.16 billion, Xyrem sales of $755 – $755 million, Erwinaze at $185 – $200 million, Defitelio at $42 – $52 million, and Adjusted Net Income of $496 – $520 million. While we believe the market has already priced in these expectations, we continue to see Jazz as an undervalued, high-growth story going forward, and we believe that its longer-term revenue, EPS, and EBITDA are likely to exceed consensus estimates in FY 15, FY 16, and beyond.

- **Catalysts:** 1) **Possible price increases for Xyrem** – Given the historical price increases and the price ranges for comparable orphan drugs, we believe the company is likely to announce another round of price increases at the end of FY 14 or early FY 15, and that there is significant room to grow pricing beyond the current levels. 2) **Launch of new marketing campaigns for Xyrem** – Jazz management is in the process of launching awareness campaigns for narcolepsy patients in key geographies, and has already reported 11,400 Xyrem patients in Q1, above our FY 14 estimate of ~11,300. 3) **Settlement of Roxane lawsuit** – We believe this will be decided in Jazz’s favor, resulting in a delayed entrance for Xyrem generics. Current market expectations point to generics in FY 19 or FY 20, but we believe FY 21 is more likely (with peak sales of ~$3.0 billion in FY 20).

- Our **$170.00** target price is based on an FY 14 EV / EBITDA multiple of 20.7x and an FY 15 EV / EBITDA multiple of 15.3x, vs. median peer company multiples of 21.8x and 15.3x, respectively. Given Jazz’s higher revenue growth, margins, and EBITDA growth, we believe this is still quite conservative. A DCF analysis with our long-term FCF projections, a discount rate of 8.07%, and a Terminal FCF growth rate of 0.3% also produces an implied share price of $168.71.

---

What is the analyst’s rating?
Why?

Has s/he changed any other metric? Why?

What is the valuation framework the analyst is using? Does s/he vary from the conventional ways of using it?

What are the key upside risks/downside risks?

- One research tack: Obtain semantic forms, reason with them
- Semantic form in a tailored logical vocabulary (fluent over qual/quant domains)
Are there restrictions on selling participations in the loan?

No

Why?
- “... Each Noteholder may ... sell participations ...”
- “without the consent of the Borrower”
- Passage is conditional on (A) ... (D).
  These are “boilerplate” (not meaningful).

Are there restrictions on post-payment in case of event of default?

...
Deep Hedging

Reward measures cash flow from held instruments and costs of transaction, with risk-adjustments.

- Separate market model needed, e.g. Black Scholes.
- Potentially, model-free: train in (augmented) market data.

State: market observations ... current/past prices, news etc.
Action: Decide what to trade, what to hold

Challenges

1. Develop integrated programming / reasoning / validation / risk-analysis framework for integrated CS 2.0 programs
   1. Sketching (resolved through constrained optimization) as a unifying framework
2. Develop framework for exploiting background knowledge
   1. Sketched function is translation invariant $\Rightarrow$ use convnets
   2. Permutation invariant $\Rightarrow$ Use Deep Sets [1]
   3. ([2] General theory emerging)
3. Develop differentiable (learnable) data-structures, e.g. stacks, queues, deques [3], tries, arbitrary graphs, ..., control-structures [4]

Conclusion