The Data-Driven Economy

Applications of the M Language in Agriculture

Edmund W. Schuster and Stuart J. Allen
The Data Center Program
Laboratory for Manufacturing and Productivity
Massachusetts Institute of Technology
WHAT WE WILL DISCUSS TODAY

1. Managing **risk** in agriculture
2. An **introduction** to the M Language
3. A list of **applications** and projects programmed for the future
4. **Engineering Marketing Science** – The integration of marketing science, engineering technology, and supply chain management
www.mitdatacenter.org
An Interactive Discussion is welcome

Please feel free to ask questions or add comments


PURPOSE OF THE “HARVEST MODEL”

- Balance the growers desire to harvest all grapes before a hard frost verses capital expenditures required for maximum through-put rate.
- Historically, Welch’s used a fixed-length of harvest to plan the though-put rate.
- The fixed-length of harvest method ignored the risk of a hard freeze.
DEFINITION OF “POLICY”

• Take 100% of the crop, 85% of the time
• Implies a harvest rate (R) required to meet the policy
• By defining a “statistical” policy for receiving grapes we can make trade-offs between harvest capacity and investment in equipment
• We calculated a “loss function” and found the 85% policy to be optimal
Qualitative Comparison of Start Dates and First 28 Degree Day With Estimated Triangular Distributions

**Lawton**
- Earliest start date: 30-Aug 1991
- Latest frost date: 20-Nov 1985
- Probability distribution of start and frost dates
- Comments: Lawton has occurring; Six times days

**Grandview**
- Earliest start date: 3-Sep 1992
- Latest frost date: 28-Nov 1983
- Probability distribution of start and frost dates
- Comments: The West has occurring; Over a 50% in October

**North East**
- Earliest start date: 6-Sep 1991
- Latest frost date: 4-Dec 1986
- Probability distribution of start and frost dates
- Comments: North East in October; Largest amounts

---

Average safe days:
- Lawton: 27 safe days
- Grandview: 37 avg. safe days
- North East: 46 avg. safe days
DATA REQUIRED FOR THE HARVEST MODEL

• Harvest Size - we use the average of the LRP for Concord, for each growing area

• Historical analysis shows the harvest size to be normally distributed
• We use the “start date” and “end date” provided by National to calculate the length of season, L

• We assume the distribution of the season length to be normal (based on observations of histograms)

• L is not correlated with harvest size, H.
  – .14 correlation with significance of 53%.

• ONLY DATA AVAILABLE – POINT ESTIMATE OF TEMPERATURE

• TEMPERATURE SENSORS
M – THE BIG PICTURE

• Sensors
  “the number of deployed sensors will dwarf the number of personal computers by a thousand fold in 2010”


• A Network of Models
  – Capture 50 years of modeling
  – Something like eBay
  – The future of ERP…Packaged Software?
  – SAP and DEC, Analog Devices

• Connect to the customer, interact

• Interoperable Data
  – Something like Adobe Acrobat
SEVERAL TYPES OF WEBS

• The Web of Information
  – HTML and the World Wide Web

• The Web of Things
  – Linking physical objects together using the EPCGlobal Network and RFID

• The Web of Abstractions
  – Building a network of mathematical models
  – Link models together
  – Link data to models
  – Computer languages & protocols to create a free flow of models in a network (Internet or Intranet)
GLOBAL RFID: The Value of the EPCglobal Network and RFID for Supply Chain Management

Edmund W. Schuster, Stuart J. Allen, David L. Brock

- Publisher: Springer Verlag, Manuscript = 330 pages
- 600 citations
- 41 figures and tables
GOALS IN WRITING THE BOOK

• What does the capability of unique identification mean for supply chains and business in general?
• Insight into EPCglobal, Inc.
• Targeted for a wide audience
• Focus on implementation
• The role of data, and the future (MIT Data Center and the M Language)
• Foreword written individually by Kevin Ashton and Sanjay Sarma
TABLE OF CONTENTS

List of Photographs, Figures, and Tables

Foreword

Preface

Part I: Introduction

1 - The Emergence of a New Key Technology

2 - Hardware: RFID-Tags and Readers

3 - Infrastructure: EPCglobal Network

4 - Data: What, When and Where?
Part II: Leveraging the Supply Chain: Case Studies

5 -Warehousing: Improving Customer Service

6 - Maintenance: Service Parts Inventory Management

7 - Pharmaceuticals: Preventing Counterfeits

8 - Medical Devices: Smart Healthcare Infrastructure
Part II: Leveraging the Supply Chain: Case Studies

9 - Agriculture: Animal Tracking

10 - Food: Dynamic Expiration Dates

11 - Retailing: Theft Prevention

12 - Defense: Improving Security and Efficiency
Part III: Creating Business Value

13 - The Role of Data in Enterprise Resource Planning

14 - Building a Business Case for the EPCglobal Network

15 - Enhancing Revenue Using the EPC

16 - Outlook: Navigating the Sea of Data

Notes

Glossary
Problem

40% to 60% annual data increase

From Forbes
PROBLEM

What are you going to do with all your Data?
“Semantics is a hot industry sector right now – a $2 billion per year market and projected to grow to over $50 billion by the year 2010.”

“Leading analysts have estimated that 35-65% of our System Integration costs are due to Semantic issues.”

And in every sector of the market...our biggest software challenges come down to creating and resolving meaning. In other words: semantics.

2006 Semantic Technology Conference
San Jose, Ca
The market in semantics might be 10 times larger than RFID

This is a good area to add value, create new research, and make profits

There are many internal company benefits as well

We think the MIT Data Center might be larger than the Auto-ID effort
The market for mathematical models might be even larger

We want to become the “Henry Ford of Modeling”
MIT DATA CENTER PROGRAM

• First work conducted in 2003
• Smartworld 2004 – over 300 attendees
• Administrative Unit within MIT 2005
• Laboratory for Manufacturing and Productivity
  – New Master of Engineering in Manufacturing
  – A number of high tech manufacturing experts
• First member, MorganFranklin Corporation, then LG, Raytheon, Siemens, and ReadyTouch

• Winner of the 2004 E. Grosvenor Plowman Award given by the Council of Logistics Management for best contribution to the study of logistics.
THE M LANGUAGE

- David Brock, Chief Architect
- Initial Design – Dictionary and Rules
- A way to link data together semantically
VISION

Model A

Model C

Model B

Database A

Database B

Database C
EXAMPLE - LOGISTICS
DATA
Data
Data
VISION

Mission

- Make sense of your data

Task

- Create the standards and systems for interoperable data and modeling
How do we synchronize data?

Data System A

Data System B

Data = Symbol

CONTEXT

Target: Vocabulary and Syntax
XML

Must agree on vocabulary and syntax
XML is like a form.
PROBLEM

Different forms?
Can’t we just agree on one form?
STANDARD?

Whose form?
M
A Modeling Language

The Fundamental Idea:

Separate vocabulary and grammar
DATA “BLOCK”
DICTIONARY ENTRY

WORD
  call.5

DEFINITION
  call  n.  a telephone connection.

RELATIONS
  Synonyms: phone_call.1, telephone_call.1
  Type of:  telephone.2, telephony.1
  Part of:

DATA
  Data:  ^[+][0-9]\d{2}-\d{3}-\d{4}$
  Attributes:  party.5, duration.1, telephone_number.1

TRANSLATIONS
  Data:  电话, telefoongesprek, 전화, телефонный вызов
DATA “ATOM”

Data that “self identifies”
• Web accessible
• Web editable – A “wiki” dictionary
• Web community
• Staged approval
  • Proposal – Universal accessible and editable
  • Draft – Universal accessible and limited editable
  • Pre-approval – Universal accessible and limited comments
  • Recommendation – Universal accessible
HOW DOES M COMPARE TO SEMANTIC WEB?

• The M Language works with W3C standards
  – XML
  – XSLT
• M focuses on atomic elements, Semantic Web focuses on creating an ontology
• M is designed for “many to many” communication, across industry disciplines
• Semantic Web plans to use some elements of Artificial Intelligence and Knowledge Management
Applications
CNN.com delivers up-to-the-minute news and information on the latest top stories, weather, entertainment, politics and more.

Hurricane Rita battered South Florida and the Keys with heavy rain and strong winds Tuesday after strengthening to a Category 2 storm. Gov. Jeb Bush warned residents to stay vigilant as the storm -- with maximum sustained winds of 100 mph -- passed through the Straits of Florida without so far making official landfall. Radar indicated Rita spawned tornadoes near Hollywood, Florida, and a water spout or tornado near Islamorada, in the upper Keys.
M DATA FEEDS

NOAA NDBC

Raw Data Feed

YYYY MM DD hh mm WD WSPD GST WVHT DPD APD MWD BARO ATMP WTMP DEWP VIS PTDY TIDE
2005 07 11 17 50 MM MM MM 1.2 5 MM MM 1011.8 16.2 13.8 13.6 MM -0.7 MM

‘M’ Data Feed

<timestamp.1>
2005-07-11T17:50
</timestamp.1>
<wave.5_height.2>
1.2
</unit.5>foot.11</unit.5>
</wave.5_height.2>
M BROWSER
‘M’ MODEL EXPLORER

Model Explorer

THE DATA CENTER
Making sense of your data
Engineering Marketing Science
Mass Advertising is taking a big hit

“Advertising is scary”

Prof. Duncan Simester
MIT Sloan School of Management
The integration of marketing science, engineering technology, and supply chain management.

Supply chains that sense and respond to the physical world.

This requires an Intelligent Infrastructure for management, control, automation and interaction.

The M Language is an open system that will form the base.
OBJECTIVE

Develop new ways of influencing customer decision-making at the point of sale

Interactive Marketing

Use the M Language as the data aggregator between venders and retailers
Research Question – measuring and modeling the effect of increased amount of data at the point-of-sale.

Commercial Question – employ interactive marketing to increase sales at retail outlets.
M LANGUAGE APPLICATION

- Serve as a data aggregator
- Effective solution for the “many to many” problem
- Open source system
- **Key Point** – no standard exists today
In-Store Informational Kiosk

Self-service, interactive, networked terminals in the aisles for:

- Product information
- Comparisons
- Targeted marketing
- Promotions

www.readyTouch.com
Brand Owner Benefits

- Consistent messaging
- Direct access to retail shoppers
- Access to customer shopping metrics
- Highly creative marketing options
- Deliver timely promotions

www.readyTouch.com
Retailer Benefits

- **Improve bottom line**
  - More sales
  - Higher margin sales
- **Improve customer service**
  - More satisfied customer
- **Hi-impact cross-promotions**
  - Increase basket size
- **Optimize staffing**
  - Kiosk is product “expert”

Adapted from ReadyTouch, Inc  
www.readyTouch.com
A Supply Chain for Product Information

New Content

New Content

Content changes

Approved Content

Publish by Location

Publish by Location

Staging

Content changes

www.readyTouch.com
Expansion options

- Put the information on a cell phone
- RFID enabled devices – auto detect
- Loyalty card tie-in
- Web history tie-in
- Blogs and customer reviews
- Reporting tools to show marketing trends
Changes in the Supply Chain

• **Orientation 1 – Power to Retailer**
  - Consumer has the ability to compare across brands at point of sale

• **Orientation 2 – Power to Manufacturer**
  - Control information flow about a brand

• **Orientation 3 – Marketing Research Tool**
  - Gain store level data on customer behavior
Marketing Spatial Diffusion

IMPLEMENTATION

• You can use the M Language now!
• We are refining the Dictionary, Browser, and Rules
• Distributed dictionary approach
• Controlled “wiki” process
• Industry leaders of the MIT Data Center Program drive use and control future direction
1. Building an intelligent network that links **models to data** (Alternative to ERP)

2. Translating data at the **edge** of computing systems (Data integration in supply chains)

3. Internet Search tool that uses the **definition** of the word (common understanding within MFG. organizations)

4. Various forms of **visualization** of data through a tangible user interface
M Language – Application Goals (continued)

5. A Standard for **Spatial** Data (impact of weather on logistical systems)

6. **Data Aggregation** (tech. manuals, maint. records)

7. **Human** Language Translation (SCM documents)

8. A standard for **sensors** (capital equipment)

9. A standard for **location** (general supply chain)

10. Improve data **quality** (general supply chain)
www.mitdatacenter.org