Collaborative Intelligence - the Need for Approximation

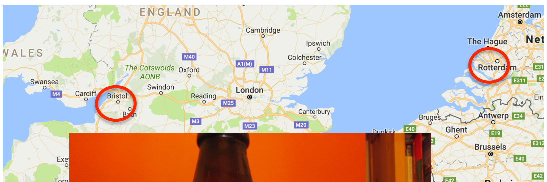
It's better to be vaguely right than precisely wrong

Background
Fuzzy in intelligent systems
Application – Situation Awareness and Security Analytics
Sequences of similar events
Provenance and approximate ontologies

Trevor Martin
Professor of Artificial Intelligence, University of Bristol
Senior Research Fellow, BT Security Futures
trevor.martin@bristol.ac.uk

University of Bristol
BT

Bristol



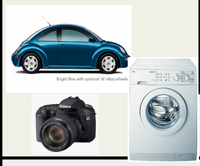
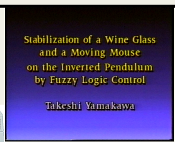
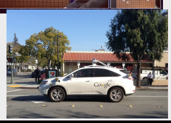
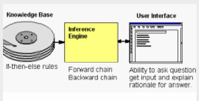
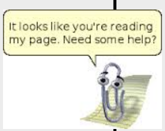
New AI vs Old AI

If the future is like the past, it can be predicted (given sufficient data).
BUT
If we can predict the future, it will not be like the past.

- new AI ≈ statistical machine learning
  - intelligent in that it copies human/natural solutions in similar circumstances
  - unarguable if we have perfect, full information
  - arguable in other circumstances (poor definitions, changing models, unseen data, …)
  - relatively easy to produce (gather data – mobiles, web, sensors - build model, run)
- previous AI: extract +reproduce human knowledge / understanding
  - intelligent in that it approximately duplicates human behaviour
  - may be able to handle exceptions, and “explain” its reasoning
  - in many cases, can exceed typical human performance
  - generally difficult to create and maintain

Collaborative AI
  - make use of complementary strengths of humans and computers
  - e.g. cyborg chess, Amazon turk, “crowdsourced” code-Kaggle, visual analytics
  - contrast to autonomous (controlling) AI

Examples

	Old AI	**New AI**
Autonomous	fuzzy control  	deep mind  google car 
Collaborative	mycin expert systems  	Cyborg-chess tesla car 
	distilled human expertise	distilled data

Approximation

Use mouse pointer to zoom image.

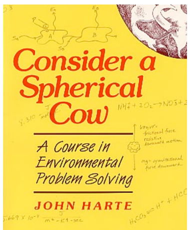
1. What is the breed of the dog in this image? Select only ONE of the following:
   - Greater Swiss Mountain
   - Japanese Chin
   - Scottish Deerhound
   - None of the above

2. Select fur Color(s). If the dog in this image has more than one fur color, you MUST provide all of them individually. If needed, click on "Add Color" to enter another color name. Example
   - select one -
   Add Color

3. Select fur texture of the dog in this image: Example
   Straight/Smooth | Wire | Curly | Corded | Hairless

4. Select the length of the fur in this image: Example
   Short | Medium | Long | None

Example from an Amazon Turk image labelling task
Is fur : "*short, straight, black / white / light brown*" ?



Be aware of approximations in data and computation (*when* they are made and *how* they are made)

$$2.4 + 2.4 = 4.8$$

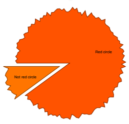
2 + 2 = 5

Rounding to nearest integers

A role for fuzzy

Human perception

Closest category



red circle

Stored data - starting point for intelligent systems

"mid-career researcher"

Computer definition:
PhD + 5 or more years experience

stored world is forced into CRISP categories (often by subjective judgment)

output – value, decision, …
CRISP (*mostly*) but more understandable if expressed using initial, approximate categories

Human description:
"*strong* possibility of a *short shower* in *central Bristol* during *rush hour* tomorrow morning"

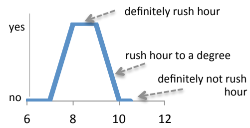
Computer definition:
Event: uniform liquid precipitation composed exclusively of drops with diameters < 0.5mm
Duration: 5 - 20 minutes
Location: within 2 miles of 51°27'N 2°36'W
Time : 0800 - 0900
Probability : 0.8

Fuzzy is needed because humans mostly work with approximate categories

Fuzzy thinking ... bridging the gap between human and machine

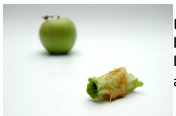
we use *fuzzy terms* to communicate efficiently

- mathematical representation of approximate definitions
- a fuzzy set allows elements to have partial membership
  typically $\mu : U \rightarrow [0,1]$



more generally, membership *lattice*

- indicates the degree to which an object has some property
- relative membership is important – B is *good*,
  C is *better*, A is *about the same* as B
- property *X* is fuzzy if an object can be *very X, slightly X, etc*



How many bites before an apple becomes an apple core?

– there is also a hierarchical aspect

why is this
a dog in a field,
not a mammal in the countryside
or a black labrador surrounded by grass?

- fuzzy categories = maximum information
  with the least cognitive effort

Interface to intelligent systems : model fuzzy uncertainty in *categories* and *hierarchies*

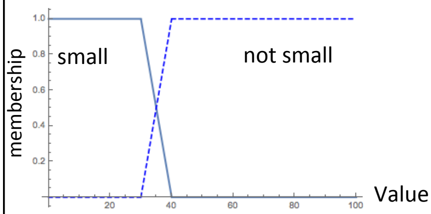
see also

Rocacher /Bosc "Relative integers"
Dubois/Prade : "Gradual numbers / elements"
Sanchez – RL representation

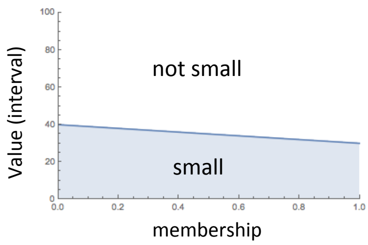
Fuzzy and the excluded middle

Zadeh's original motivation : avoid over-precision

- the recipe for *"fuzzy-X"*
  - choose a method/algorithm/representation
  - replace crisp sets (or singletons) by fuzzy sets
  - e.g. fuzzy control, fuzzy rules, fuzzy databases, fuzzy arithmetic, ...
  - problem – computation uses sets/intervals instead of single values



Viewpoint 1 – fuzzy predicates are intrinsically "gradual". Partial membership in *small* means partial membership in not-*small*. Losing the excluded middle is the price we pay for truth-functional operators

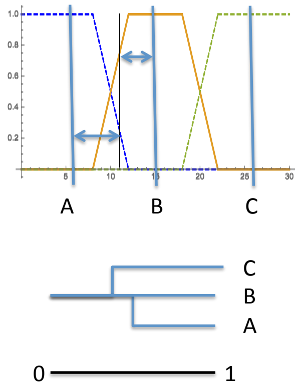


Viewpoint 2 – *X-μ* fuzzy predicate has a crisp extension (but the boundary is not precisely specified). An object cannot be *small* and not-*small*

Indiscernible centres

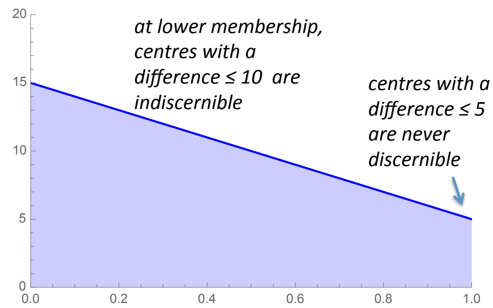
- a fuzzy partition is equivalent to a set of centres
  - traditional membership can be expressed in terms of the distance from a *point* to each *centre*

$$\mu_A(x) = f_1(x - A)$$
$$\mu_B(x) = f_2(x - B)$$

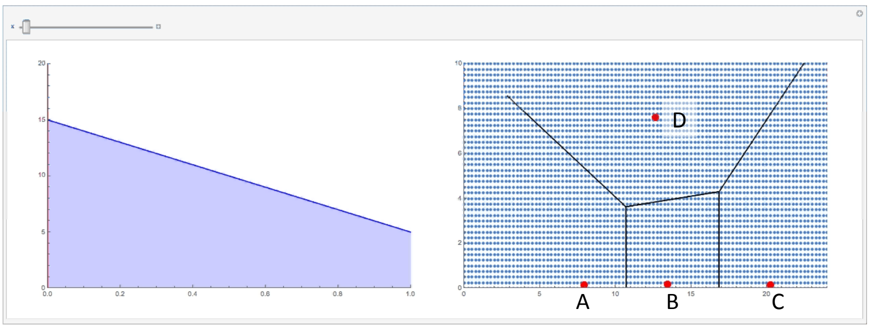


Alternatively – ignore membership, assign a point to its nearest centre but use an *X-μ* definition of *indiscernibility* to determine whether two centres are "the same"

Benefit – we have a crisp partition (but it varies according to the degree of discernibility)

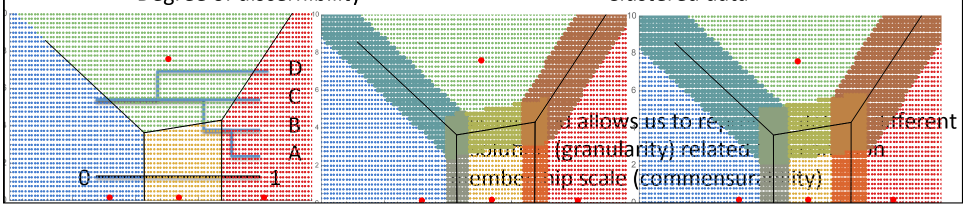


2-d example

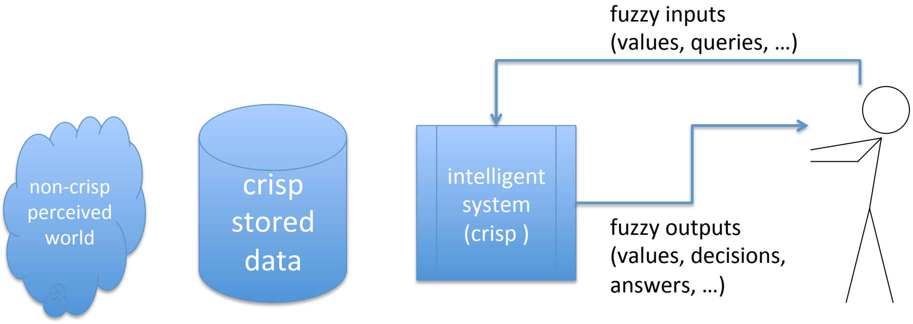


Degree of discernibility

Clustered data



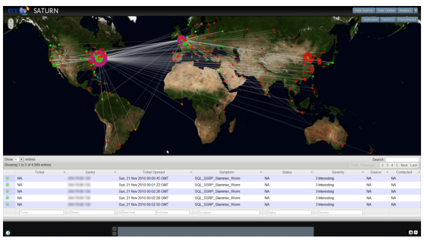
A better way to "fuzzify" crisp intelligent systems?



Problems with "standard" fuzzy
- fuzzy requires sets / intervals - arithmetic "doesn't work properly"
- fuzzy loses the excluded middle – no partitions of data
- often need to run program multiple times instead of once
- *X-μ* discernibility approach avoids problems by always working with crisp data – "fuzziness" arises from different levels of discernibility (resolution)

Collaborative Security Analytics

- dealing with large tables of (linked) events
  - human insight via visualisation (SATURN - BT)

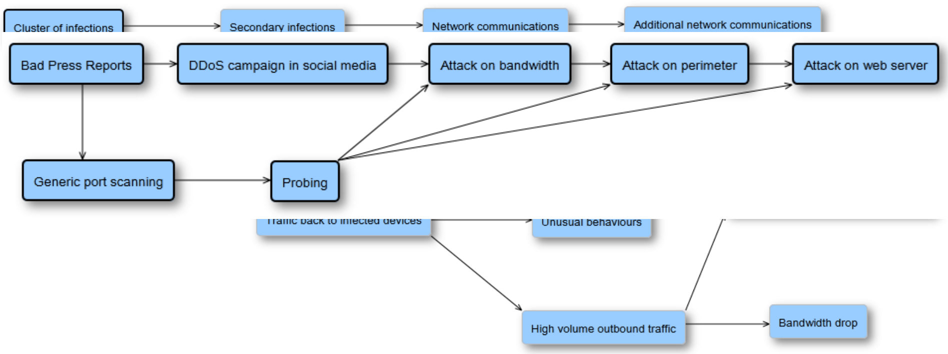


- Visualisation helps to summarise current events
  - events can be linked in sequences
  - sequences are helpful in predicting next states
  - statistics/ML not good for ***novel*** sequences – needs insight

Saturn - BT

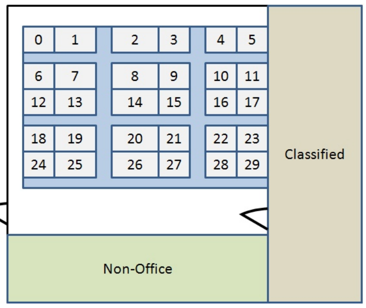
What analysts also want

- postulate sequences that haven't been seen before
- multi-stage events / soft thresholds / multiple data sources
- prediction of future events (based on specified patterns)
- interaction between multiple analysts (trust / reputation)



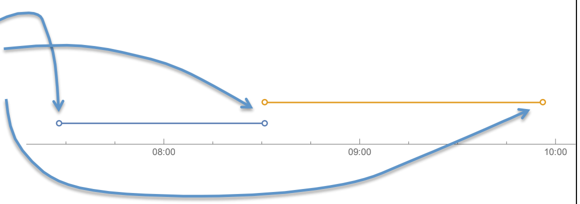
Illustration

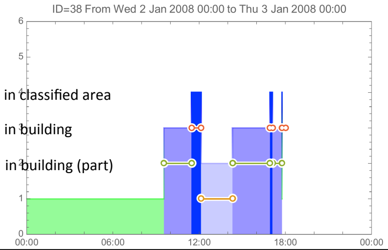
Example (VAST-2009) Movement in/out of embassy (right)
Entrance to **building** controlled by swipecard (unrestricted exit)
Swipecard needed for **Classified area** entrance and exit



Date – time	*ID*	*Type*
2008 – 01 – 01*T*07 : 28	44	*prox – in – building*
2008 – 01 – 01*T*08 : 31	44	*prox – in – classified*
2008 – 01 – 01*T*09 : 23	38	*prox – in – building*
2008 – 01 – 01*T*09 : 56	44	*prox – out – classified*
2008 – 01 – 01*T*11 : 06	44	*prox – in – building*

… *etc.*





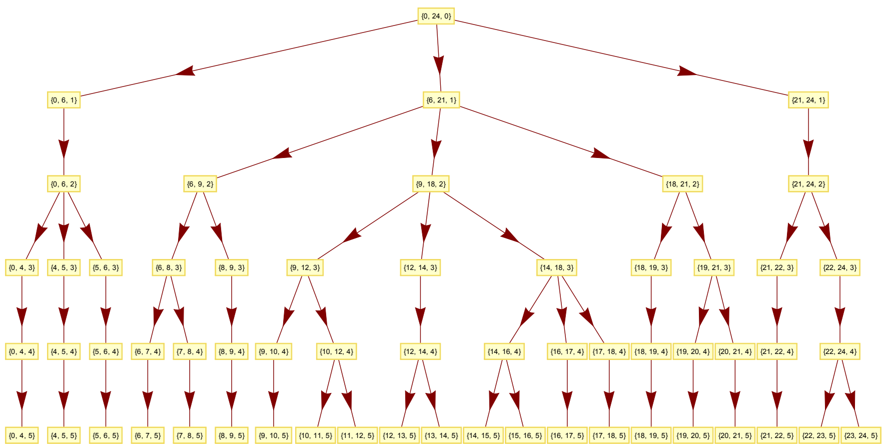
end event

start event

Status during interval	**in-building**	**in-classified**	**out-classified**
in-building	in building for part of interval	definitely in building	not valid
in-classified	not valid	not valid	Classified area
out-classified	in building for part of interval	definitely in building	not valid

Hierarchy example - hours of the day

level 1 of discernibility : early (before 06:00) daytime (06:00-21:00) late (after 21:00)
level 2 : daytime -> core office hours (09:00-18:00), pre and post … etc
Also {*alldays*} < {*weekend, weekdays*} < {*weekend, midweek, mon, fri*} < …
or {*mixedWorkPattern*} < {*early/lateStart × severalShortActivites/fewLongActivities × early/late fin*



Datetime	ID	Type
2008-01-01T07:28	44	prox-in-building
2008-01-01T08:31	44	prox-in-classified
2008-01-01T09:23	38	prox-in-building
2008-01-01T09:56	44	prox-out-classified
2008-01-01T14:26	38	prox-in-building
2008-01-02T07:11	17	prox-in-building
2008-01-02T07:24	8	prox-in-building
2008-01-02T07:32	44	prox-in-building
2008-01-02T07:36	56	prox-in-building
2008-01-02T07:44	44	prox-in-classified
2008-01-02T07:46	17	prox-in-classified
2008-01-02T07:51	42	prox-in-building
2008-01-02T07:51	39	prox-in-building
2008-01-02T07:55	35	prox-in-building
2008-01-02T08:07	45	prox-in-building
2008-01-02T08:08	18	prox-in-building
2008-01-02T08:21	28	prox-in-building
2008-01-02T08:22	17	prox-out-classified
2008-01-02T08:23	44	prox-out-classified
2008-01-02T08:28	30	prox-in-building
2008-01-02T08:34	9	prox-in-building
2008-01-02T08:56	49	prox-in-building
2008-01-02T08:57	29	prox-in-building
2008-01-02T08:58	21	prox-in-building
2008-01-02T09:01	44	prox-in-classified
2008-01-02T09:02	16	prox-in-building
2008-01-02T09:17	47	[illegible]

Some oddities are visible, such as employees who

Selecting a different date range shows who was in the building over the weekend

information on employee location
(not full information as we don't know when they leave the building)

We can focus on individual employees on specific days
Here, for employee 38 on Wed Jan 2:
- we have no information before 09:33 (green),
- employee was in building (medium blue 9:33 – 11:28),
- in classified area (dark blue 11:28 – 12:08)
- left building and re-entered (light blue 12:08 – 14:19)
- etc

We also have IP logs
(source, time, destination, port, data in/out)
Source IP address corresponds to employee ID

We can examine all employees on a specific day

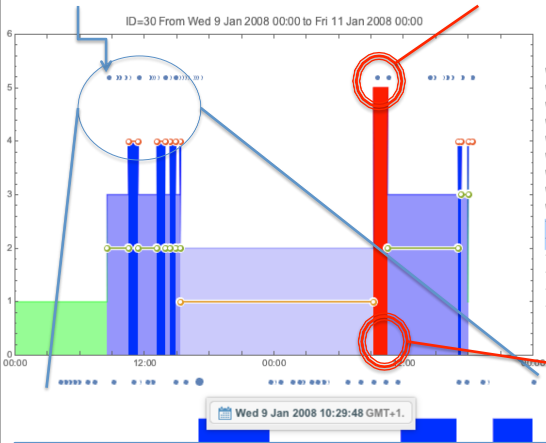
Focusing on em[illegible] …ursday morning.
Probable expla[illegible]
Hence employe[illegible] …ht similarities in sequences

Examining the whole month of data reveals groups of similar sequences AND some major anomalies (red) where there is an exit record from the classified area but no entry record.

Anomalies

IP activity for *ID=30*

IP activity whilst (possibly) in classified area



Wed 9 Jan 2008 08:32:00	30	prox-in-building
Wed 9 Jan 2008 10:30:00	30	prox-in-classified
Wed 9 Jan 2008 11:25:00	30	prox-out-classified
Wed 9 Jan 2008 13:10:00	30	prox-in-classified
Wed 9 Jan 2008 13:53:00	30	prox-out-classified
Wed 9 Jan 2008 14:23:00	30	prox-in-classified
Wed 9 Jan 2008 14:53:00	30	prox-out-classified
Wed 9 Jan 2008 15:17:00	30	prox-in-classified
Wed 9 Jan 2008 15:21:00	30	prox-out-classified
Thu 10 Jan 2008 09:14:00	30	prox-in-building
Thu 10 Jan 2008 10:33:00	30	prox-out-classified
Thu 10 Jan 2008 17:05:00	30	prox-in-classified
Thu 10 Jan 2008 17:22:00	30	prox-out-classified
Thu 10 Jan 2008 17:59:00	30	prox-in-classified
Thu 10 Jan 2008 18:23:00	30	prox-out-classified

exit from classified area without record of entry

We use a graph representation of event sequences with hierarchical fuzzy similarity
Indiscernibility arises from initial definitions and from explorations (e.g. groups with similar behaviour)

Recap – the need for a collaborative approach

- typically the aim of the investigation changes as the investigation progresses

  "here's some data. There might be something odd going on. Tell us if there is, and how to detect it / stop it / … "

- need to
  - visualise data at different levels of resolution
  - represent sequences, assumptions
  - extract patterns, propose "underlying process"
  - record, possibly repeat, processing steps
  - trace output data back to : inputs, parameter values, analyst choices, …

  i.e. - build up an approximate ontology – entities, constraints, relations etc.

Summary

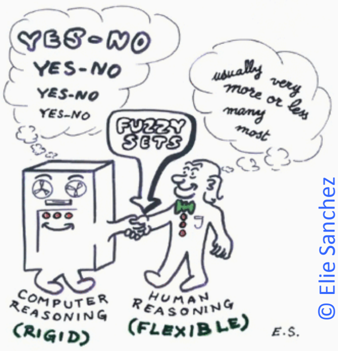
collaborative intelligent systems : human plus computer

machine to do "simple" tasks, human to provide analysis, insight and direction
visualisation is good for machine to human communication
fuzzy humans need fuzzy machines (or machines that can be fuzzy)
(where fuzzy = graded tolerance)
Never completely trust the data

application – situation awareness / cyber-security

model and detect event sequences
(including previously unseen sequences)
widely applicable

work is needed to combine
data (integration of multiple sources),
visual analytics, AI models and human creativity



© Elie Sanchez

Fig. 90.1. Human reasoning versus computer reasoning

Some related publications

Papers

Finding Fuzzy Concepts for Creative Knowledge Discovery
T. P. Martin and A. Majidian
International Journal of Intelligent Systems 28 (2013) 93 – 114

Change mining in evolving fuzzy concept lattices
T. P. Martin
Evolving Systems 5 (2014) 259-274

Representation and Identification of Approximately Similar Event Sequences
T. P. Martin and B. Azvine
In Proc. *Flexible Query Answering Systems*, Krakow, (2015) 24-29

The X-mu representation of fuzzy sets
T. P. Martin
Soft Computing 19 (2015) 1497 – 1509

An Incremental Fuzzy Approach to Finding Event Sequences
T. P. Martin and B. Azvine
In *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, (2016) 525-536

A Virtual Machine for Event Sequence Identification using Fuzzy Tolerance
T. P. Martin and B. Azvine
In Proc. *FUZZ-IEEE*, (2016) 1080-1087

Patents

SIFT - Sequence Identification using Fuzzy Tolerance : Construction
T. P. Martin and B. Azvine
(2013)

A Virtual Machine for SIFT - Sequence Identification using Fuzzy Tolerance
T. P. Martin and B. Azvine
(2013)

Multi-level tolerance relations to detect anomalies in physical security
T. P. Martin and B. Azvine
(2014)

Efficient event filter
B. Azvine and T. P. Martin
WO2015044630 A1 (2015)

Sequence identification
B. Azvine and T. P. Martin
WO2015044629 A1 (2015)

Thank you for your attention

Any questions?

- The colour of truth is grey

André Gide, French author and winner of the 1947 Nobel Prize (Literature)