Description logic using Deterministic Inductive Logic (DIL)

A novel approach to leverage DIL to enable the capture of evidence encoded using semantic truth-value feature vectors that enable direct computation of inductively constructed quantified models that support negation and enable a new approach to supporting information filtering and retrieval in dynamic data environments

Allen Brewer, Ph.D.

Major topics

- Boyd cycle – OODA loop
- Semantic based description logic
- DIL data structure & binary facets' truth values
- DIL operations and tests
- Classification systems overview
- Clinical encounter evidence system

Boyd cycle

http://odsinc.us:8040/rid=1P9357MKG-287JPY4-1QD/Boyd cycle - basic functionality.cmap

The Boyd cycle:

- Observe – acquire information/knowledge
- Orient – construct and manipulate logic models
- Decide - choose
- Act – perform information and interventional acts

Which rely upon capabilities to:

- describe information/knowledge,
- organize, combine, factor and manipulate descriptive models reflecting the dynamics at hand (context)
- Compare, contrast, assess, compute relationships, make decisions and trigger actions

Model building – model use

- Observe and orient are:
  - Emergent, inductive, qualitative processes from a constructivist perspective used to build and evaluate a conceptual model that can describe a potentially highly dynamic environment
- Decide and Act are:
  - Deductive processes
  - Steps to identify one or more alternatives and make choices based upon *values* in a decision model
  - Processes that do things or take actions that may cause changes to occur

Constructing complex computable descriptions

- Concept maps – developed by Joseph Novak http://cmap.ihmc.us/
  - Constructed of propositions, formulated from two concepts connected by a linking phrase
- Statements are constructed of:
  - Propositions and chains of connected/linked propositions
- Models are constructed of:
  - Collections of *statements*

  Using an extensible vocabulary of semantically consistent terms and truth-values that are assigned in accordance with rules for their assignment and use

Meanings not forms

- Elements must reference concepts, propositions, etc., that have consistent ***expressed*** meanings not lexical text strings whose meanings are ***implied***
- Elements must be
  - Defined prior to their use so that they can be semantically consistent (intentions)
- Truth-values must be
  - assigned consistent with their rules for use so that their expression describes *reality* (extensionally consistent)

Extensibility

- Terminology
  - Terms can be added by defining them and their associated truth-values and by including new elements in truth-value feature vectors
- Evidence, experience, etc.
  - Cases can be added by including them in a repository, such as:
    - an evidentiary repository of clinical encounters (EHR/EMR repository)
    - an experiential database

Semantic descriptions

http://odsinc.us:8040/rid=1P9361KHT-4D6TLL-4NN/DIL inductive terminological hierarchy.cmap

The description (value/meaning) of any case, profile of cases, profile of profiles, etc., can be recomputed at any time to reflect any changes that have been made to:

- the underlying description system,
- how cases have been described,
- how cases have been organized into categories
- how categories have been aggregated into generalization categories
- Etc.

Critical requirements for dynamic reasoning systems

System must have a

- Knowledge representation/data structure for describing cases (profiles)
- Capability to construct intentional aggregates (combine cases) that describe categories
- Capability to quantify categories
- Capability to dynamically compute/re-compute the subsumption relationship between any two profiles
- Capability to support negation

DIL approach

- In DIL a category is defined by the characteristics of the members of that category, so a category's characteristics (profile) can be dynamic and can be recomputed cyclically such as at decision time
- Categories are described by aggregating the descriptions of the cases that are used to define the category. DIL defines a COMBINE() operation to compute aggregate descriptions (profiles) that use two new truth-values {Unknown, Indeterminate} in conjunction with the truth-values used by a faceted element, such as {True, False} as are used by binary facets.

Combine()

http://odsinc.us:8040/rid=1P9360ML5-PL9PV6-482/DIL combine operation evaluation.cmap

Element(n)	1	2	3	4	5	6	7
CASE(1)	U	U	T	F	T	F	U
CASE(2)	U	F	T	T	F	F	I
COMBINE(1,2)	U	F	T	I	I	F	I
QUANTIFIED	UNIV	UNIV	UNIV	EXIST	EXIST	UNIV	EXIST

Truth-values for facet elements

All DIL elements

- Unknown {instance, profile} – no value known
- Indeterminate {profile} – existentially quantified

Binary facet elements

- True {instance, profile} – universally quantified
- False {instance, profile} – universally quantified

Greater than binary facet elements

- May contain *n* mutually exclusive and collectively exhaustive truth-values

DIL critical functionality test

http://odsinc.us:8040/rid=1P9362618-8ZM31G-4V7/DIL tests.cmap

- Capability to compute subsumption relationships
  - The case is subsumed if its characteristics are within the scope of the category's characteristics
  - A case(n) is subsumed by a profile that is defined by a case-list(X) IFF combine(case-list(X)) = combine(case-list(X), case(n))
- Capability to support negation – if a case profile is not subsumed by a category profile, the case is NOT a member of the category

DIL capabilities

http://odsinc.us:8040/rid=1P935YGRS-KC2RSJ-42R/Boyd cycle with DIL overlay.cmap

To build models inductively and to reason about those models one needs to be able to:

- Combine() - aggregate values
- Compare() - similarity analysis
- Contrast() - difference analysis
- Factor() - divide descriptions into subsets or sub-classes
- Compute hierarchical subsumption
- Identify any universally quantified discriminants of rules, decisions, models, etc.
- Evaluate whether a representation is sufficient to deterministically distinguish between two categories, compute the degree of similarity or difference between two descriptions, etc.

Classification systems

http://odsinc.us:8040/rid=1P934VXF5-1JNSPJV-SL/Classification models and methods.cmap

- **Aristotelian classification** – a priori, pre-coordinated, deterministic, static, rigid, authoritative, uses top-down organization by division, single perspective, etc.
- **Faceted classification** – a priori, post-coordinated, deterministic, static, description is authoritative, query formulation is more dynamic, multiple perspectives, etc.
- **Probabilistic** clustering approach – more dynamic, organized bottom-up (inductive) by similarity, quantitative feature space is defined by term frequency distributions
- **DIL** – dynamic, feature space defined by semantic truth-values, qualitative, supports quantification and negation, can maintain cycle-time-coordination, inductive, flexible perspective, user defined descriptions

Encoding clinical evidence

http://odsinc.us:8040/rid=1P934Z7KS-MP2NW9-1CB/GAHMJ paper BrewerSep2012.pdf

- Cmaps are used to define evidentiary patterns of propositions and propositional chains to be used to describe measurements, observations, elicited signs, symptoms, experience, etc.
- Clinical statements expressed in the DIL description logic are used to record encounter evidence
- A profile of an encounter is created by combining all of the statements associated with that encounter

Objects and descriptive surrogates

- Original evidentiary statements, expressed in the DIL description logic, are packaged into a digital object for management and tracking. A clinical encounter is an example of a package
- The metadata surrogate for a DIL encoded collection of statements (cases) is formed by computing the combination of all statements (cases) associated with one or more digital objects or packages

Using profiles to construct cohort models

- Selected patient encounters may be used to construct a quantified model inductively for a patient, cohort, etc., using the DIL combine(). The aggregation is expressed as a *profile*.
- A profile may be used as a query formulation to retrieve encounters that are subsumed by that profile
- The cases retrieved may be combined with the originally selected cases to construct a quantified cohort profile inductively

Natural language translation

- A DIL encoded proposition can be translated into natural language using the definitions and translations specified for that proposition
- A model can be converted into a navigable Cmap for human use/review by building a text file of all propositions and passing it to IHMC's CmapTools for visualization.

Original description

- Using DIL description logic assures a coherent, semantically consistent, computable framework to record, combine, compare, contrast, factor, evaluate subsumption, etc. It enables computable semantic clinical documentation
- The “evidentiary language” is formed using universally unique semantic identifiers (uuSIDs) that reference concepts, linking phrases, propositions and propositional chains

Technical implementation

Semantic identifiers can be mapped to a digital truth-value feature vector that:

- Is dense, enabling models employing millions of truth-values to be stored in a digital object the size of a high resolution photograph
- Can be computed at CPU speeds
- Can be partitioned or segmented
- Can be processed in parallel
- Can leverage hardware accelerator ASICs in certain applications and implementations

Published papers

- Dissertation – Deterministic inductive logic: a multi-valued logic for reasoning about categories
  http://odsinc.us:8040/rid=1P937PPTS-4HLKG6-6LS/Dissertation Final Version.PDF
- IHMC conference paper - The use of Cmaps in the description of clinical information structure and logic
  http://odsinc.us:8040/rid=1P937P7ZQ-1GRSLYL-6LM/cmc2010-b7.pdf
- GAHMJ paper - An Application of Cmaps in the Description of Clinical Information Structure and Logic in Electronic Health Records
  http://odsinc.us:8040/rid=1P934Z7KS-MP2NW9-1CB/GAHMJ paper BrewerSep2012.pdf