Shet, Vinay et alii

Predicate Logic based Image Grammars for Complex Pattern

International Journal of Computer Vision (2011)

Hemerson Pistori

Biotechnology Dept., Dom Bosco Catholic University

January, 2012

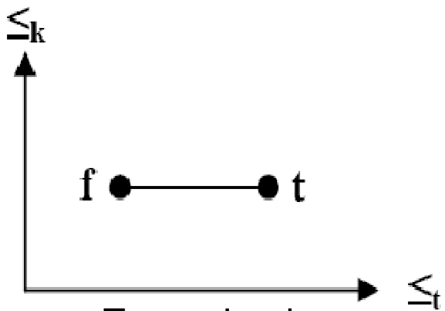
Bristol, UK

Most of the pictures used in this presentation were extracted from Shet's paper

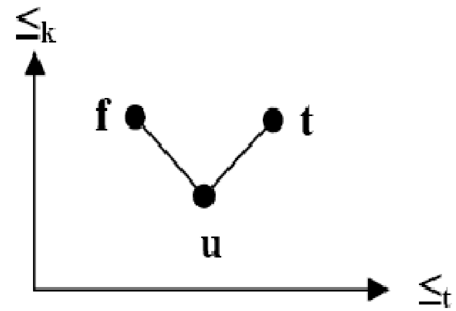
Problems with Traditional Logics Addressed

- Rules are binary (true or false) and definite
- Facts are binary (true or false0
- No support for explicit negation in the rule's head
- Not scalable (easily)
- No support for integrating evidence from multiple, possibly contradictory, sources

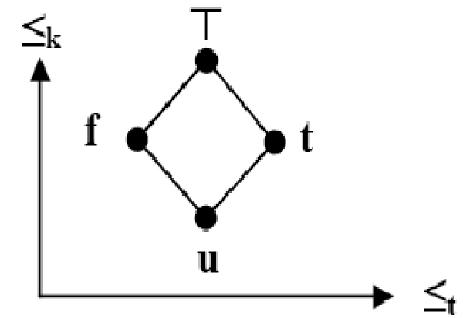
Bilattice Theory – Examples of Lattices



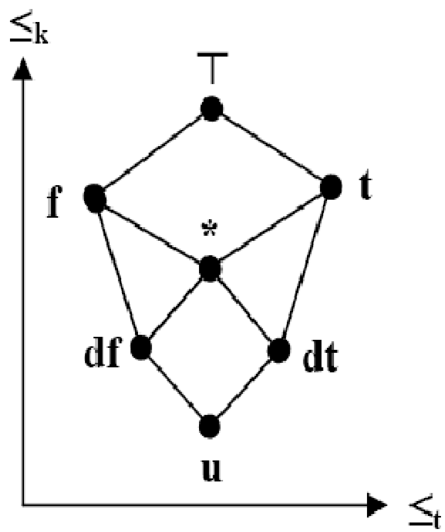
Two valued
(a)



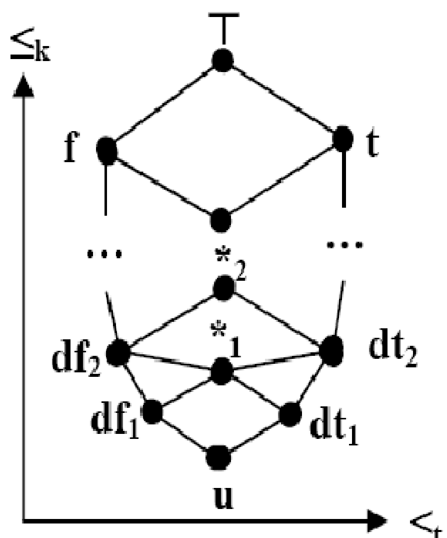
Three valued with Unknown
(b)



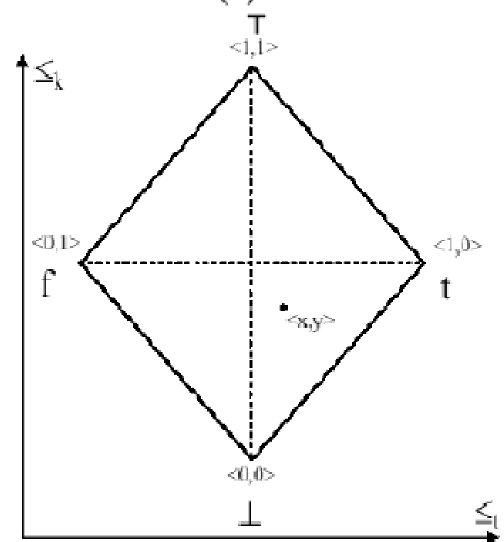
Four valued with contradiction
(c)



Default
(d)

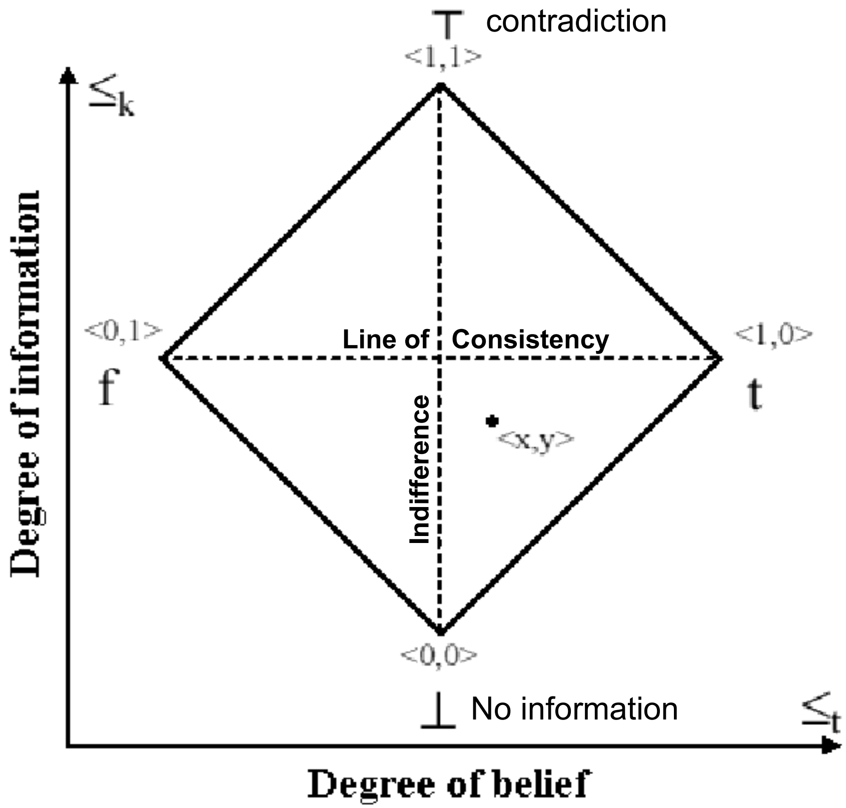


Prioritized Default
(e)



Continuous valued
(f)

Square Bilattice and Inference



< x = evidence for, y = evidence against >

Conjuntion, disjunction, combination of evidence and consensus

$$\begin{aligned}
\langle x_1, y_1\rangle \wedge \langle x_2, y_2\rangle &= \langle x_1 \wedge_L x_2, y_1 \vee_R y_2\rangle, \\
\langle x_1, y_1\rangle \vee \langle x_2, y_2\rangle &= \langle x_1 \vee_L x_2, y_1 \wedge_R y_2\rangle, \\
\langle x_1, y_1\rangle \otimes \langle x_2, y_2\rangle &= \langle x_1 \wedge_L x_2, y_1 \wedge_R y_2\rangle, \\
\langle x_1, y_1\rangle \oplus \langle x_2, y_2\rangle &= \langle x_1 \vee_L x_2, y_1 \vee_R y_2\rangle
\end{aligned}$$

Triangular norms (t-norms)

$$\mathscr{T}(a,b) \equiv a \wedge_L b = ab$$
$$\mathscr{S}(a,b) \equiv a \vee_L b = a + b - ab$$

S_q^+ Sentences that entail *q*

S_q^- Sentences that entail not *q*

$cl(\phi)(q)$ Closure of *q* given truth assignment ϕ

$$cl(\phi)(q) = \bigoplus_{U \in S_q^+} \perp \vee [\bigwedge_{p \in U} cl(\phi)(p)] \oplus \neg \bigoplus_{U \in S_q^-} \perp \vee [\bigwedge_{p \in U} cl(\phi)(p)]$$

Square Bilattice and Inference

Assume the following set of rules and facts:

Rules	Facts
$\phi(human(X,Y,S) \leftarrow head(X,Y,S)) = \langle 0.40, 0.60 \rangle$	$\phi(head(25,95,0.9)) = \langle 0.90, 0.10 \rangle$
$\phi(human(X,Y,S) \leftarrow torso(X,Y,S)) = \langle 0.30, 0.70 \rangle$	$\phi(torso(25,95,0.9)) = \langle 0.70, 0.30 \rangle$
$\phi(\neg human(X,Y,S) \leftarrow \neg scene\_consistent(X,Y,S)) = \langle 0.90, 0.10 \rangle$	$\phi(\neg scene\_consistent(25,95,0.9)) = \langle 0.80, 0.20 \rangle$

Inference is performed as follows:

$$\begin{aligned} cl(\phi)(human(25,95,0.9)) &= \langle 0,0 \rangle \vee [\langle 0.4,0.6 \rangle \wedge \langle 0.9,0.1 \rangle] \oplus \langle 0,0 \rangle \vee [\langle 0.3,0.7 \rangle \wedge \langle 0.7,0.3 \rangle] \oplus \neg(\langle 0,0 \rangle \vee [\langle 0.9,0.1 \rangle \wedge \langle 0.8,0.2 \rangle]) \\ &= \langle 0.36,0 \rangle \oplus \langle 0.21,0 \rangle \oplus \neg \langle 0.72,0 \rangle = \langle 0.4944,0 \rangle \oplus \langle 0,0.72 \rangle = \langle 0.4944,0.72 \rangle \end{aligned}$$

Table 1 Example showing inference using closure within a $([0,1]^2, \leq_t, \leq_k)$ bilattice

- human(X,Y,S) : A human with 2D center of mass at (X,Y) image coordinate and at scale S (e.g: image pyramid)
- All evidence values in the rules and facts lie on the “consistency line” (that is why they sum up to 1). This will probably not be the case in real cases.
- human(25,95,0.9) = <0.4944,0.72> : There is more evidence (0.72 > 0.49) against than in favour of the presence of a human centred at point <25,95> and scale 0.9

Rule Weight Learning

- Rules are mapped to nodes and links of an Artificial Neural Network
- Rule weights (links) are learned using a modified version of the back-propagation algorithm
- Facts are mapped to input nodes and heads are mapped to output nodes
- Facts values are assumed to be available. The facts values are detected using other object parts computer vision detectors (problem dependent)

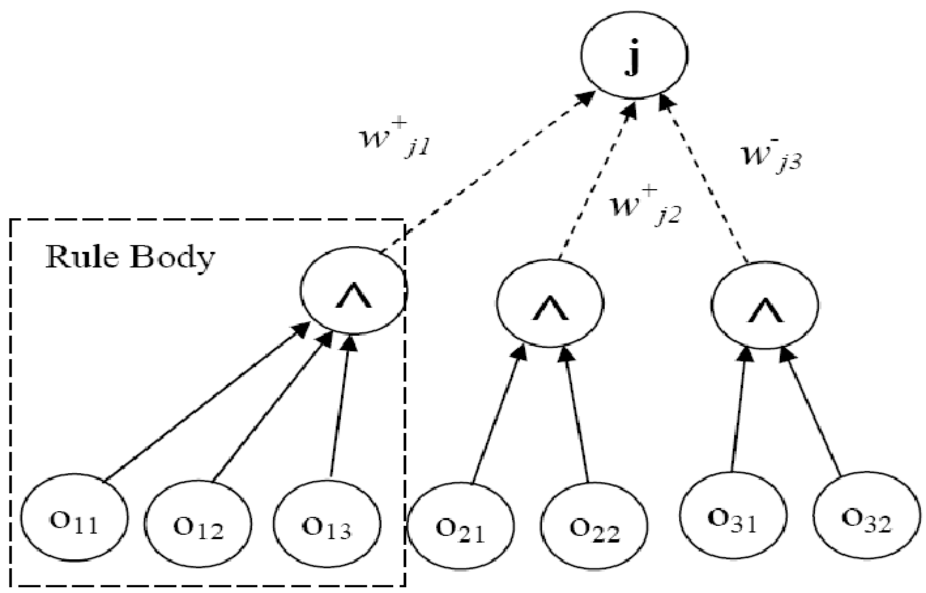


Figure 13 Example of a knowledge based artificial neural network representing rules depicted in (5).

What we can explore

- Learning rules from pure textual information
- Map logic based image grammars (LBIG) and stochastic grammars of images and try to get the best to the two formalisms
- Augment LBIG with ILP
- Propose a way to deal with temporal information