

RACIOCÍNIO E PERCEPÇÃO ESPACIAL: UMA ABORDAGEM LÓGICA ¹

Author(s):

Paulo E. Santos

¹This work was supported by Fapesp Project LogProb, grant 2008/03995-5, São Paulo, Brazil.

RACIOCÍNIO E PERCEPÇÃO ESPACIAL: UMA ABORDAGEM LÓGICA

Paulo E. Santos*

psantos@fei.edu.br

*Departamento de Engenharia Elétrica
Centro Universitário da FEI, São Paulo, Brazil

Resumo

O objetivo deste tutorial é apresentar a literatura clássica sobre raciocínio espacial qualitativo e sobre interpretação de cenas baseada em representação de conhecimento de alto nível. Tendo como base esta revisão bibliográfica, apresentaremos o desenvolvimento de um formalismo de raciocínio espacial cujo objetivo é a interpretação dos dados provenientes de um sistema de visão estéreo de um robô móvel.

PALAVRAS-CHAVE: Raciocínio Espacial Qualitativo, Interpretação de Cenas

Abstract

In this tutorial we present the classic literature about qualitative spatial reasoning and high-level scene interpretation. Based on this literature overview we present the building blocks for the development of a spatial reasoning formalism whose goal is the interpretation of stereo-vision data from a mobile robot.

KEYWORDS: Qualitative Spatial Reasoning, Scene Interpretation

1 INTRODUÇÃO

O objetivo deste tutorial é apresentar a literatura clássica sobre raciocínio espacial qualitativo (REQ), abordando seus fundamentos e principais aplicações (Stock, 1997; Cohn and Renz, 2008), e a literatura sobre interpretação de cenas utilizando conceitos de alto nível. A tarefa principal de um sistema de REQ é prover representações simbólicas e sistemas de inferência apropriados para entidades espaciais, em contraste com as técnicas numéricas tradicionais prevalentes, por exemplo, em robótica e visão computacional. Por outro lado, a interpretação de cenas de alto nível (Nagel, 1977), tem como objetivo a descrição de sequências de imagens por meio de palavras e, portanto, procura atingir uma interpretação qualitativa do espaço a partir de informações de baixo nível. A ligação entre formalismos de REQ e interpretação de cenas, entretanto, ainda é pouco investigada.

Tendo como base esta revisão bibliográfica, apresentaremos o desenvolvimento de um formalismo de raciocínio espacial cujo objetivo é a interpretação dos dados provenientes de um sistema de visão estéreo de um robô móvel. Este trabalho segue a metodologia da *robótica cognitiva*, cujo objetivo é equipar robôs com habilidades cognitivas de alto nível por meio

da criação e adaptação de técnicas desenvolvidas pela área de representação de conhecimento em inteligência artificial, especialmente aquelas que tem a lógica simbólica como linguagem fundamental. Há duas abordagens tradicionais para percepção espacial em robótica: em uma delas, percepção é entendida como o resultado de ações que produzem conhecimento (Scherl and Levesque, 1993), essas ações são executadas por um agente inteligente a fim e obter propriedades do seu ambiente para executar planos (Levesque, 1996); a outra abordagem (introduzida em (Shanahan, 1996) e (Shanahan, 1997)) entende percepção como um processo passivo, em que informações sensoriais são assimiladas pela base de conhecimento do agente via um processo de geração e teste de hipóteses. O formalismo para interpretação de cenas estéreo discutido aqui possui características de ambas abordagens para percepção espacial em robótica cognitiva.

A utilização da lógica em trabalhos como esse justifica-se pelo fato deste formalismo matemático ter suas propriedades amplamente conhecidas e por permitir a introdução em robótica (e visão computacional) de mais de 40 anos de resultados obtidos em lógica nas áreas de planejamento automático, raciocínio de ações, provadores de teoremas entre outros.

2 RACIOCÍNIO ESPACIAL QUALITATIVO

Raciocínio espacial, tal qual abordado neste tutorial, é parte da sub-área da Inteligência Artificial (IA) denominada Representação de Conhecimento. Neste contexto, o objetivo de se desenvolver sistemas (ou formalismos) de raciocínio espacial qualitativo é atingir um nível de manipulação de conhecimento espacial similar ao utilizado por pessoas no dia a dia. Em outras palavras, pretende-se desenvolver sistemas capazes de compreender, raciocinar e atuar sobre relacionamentos entre objetos cotidianos, como por exemplo “o copo está em frente ao vaso”; “Bonito é parte de Mato Grosso do Sul” etc. Esse tipo de conhecimento é muitas vezes impossível, e em geral computacionalmente inviável, de ser representado com os métodos numéricos utilizados nas ciências naturais. Assim, a pesquisa em raciocínio espacial em IA tem sido largamente desenvolvida de forma independente das áreas relacionadas à geometria e topologia (Bennett, 1997).

Podemos traçar as origens dos formalismos de raciocínio espacial em IA ao início do século XX, com a proposta de Russell e Whitehead para o desenvolvimento de teorias fenomenológicas capazes de descrever (de forma lógica) o mundo tal qual observado pela percepção humana. Nesta abordagem fenomenológica, objetos espaciais e eventos tornam-se os elementos primitivos de um tipo de “geometria” cujo objetivo é formalizar as leis básicas dos relacionamentos destes elementos, e calcular configurações particulares destes a partir das suas leis básicas. A construção de teorias formais a respeito de relações entre objetos espaciais é o objetivo central da área de raciocínio espacial qualitativo. Foi exatamente o fato das teorias fenomenológicas tentarem se aproximar da forma como o raciocínio humano funciona que tornou-as atraentes para o desenvolvimento de sistemas inteligentes.

Nesta seção apresentaremos em detalhes alguns dos principais formalismos de raciocínio espacial em IA. Atenção especial será dada a teorias cujo objetivo é formalizar o fenômeno visual da interposição de objetos a partir do ponto de vista de um observador. Em seguida, descreveremos brevemente algumas outras propostas de sistemas de REQ, suas aplicações básicas, bem como alguns aspectos sobre a obtenção de subconjuntos tratáveis destes sistemas. O nosso objetivo não é fazer aqui uma revisão bibliográfica completa sobre raciocínio espacial qualitativo (que pode ser encontrada em (Stock, 1997; Cohn and Hazarika, 2001; Cohn and Renz, 2008)), mas sim apresentar os conceitos fundamentais desta área.

Cálculo de conexão de regiões

O cálculo de conexão de regiões (RCC)¹ (Randell, Cui and Cohn, 1992)(Randell, Cohn and Cui, 1992)(Cohn et al., 1997) é uma axiomatização, em lógica de primeira ordem multi-sortida, de regiões espaciais baseada em uma relação primitiva binária a respeito da conexão entre duas regiões ($C/2$). Informalmente, dadas duas regiões x e y , a relação $C(x, y)$ (“ x está conectado a y ”), é verdade se e somente se o fechamento topológico de x e y tem pelo menos um ponto em comum. A relação de conexão obedece aos seguintes axiomas:

$$\begin{aligned} &\forall x C(x, x); \\ &\forall xy C(x, y) \rightarrow C(y, x); \\ &\forall xyz (C(z, x) \leftrightarrow C(z, y)) \rightarrow x = y. \end{aligned}$$

Com a relação $C/2$, dadas três variáveis para regiões espaciais x , y e z , são definidas algumas relações mereotopológicas,

¹Do inglês: *Region Connection Calculus* ou das iniciais dos três autores: Randell, Cohn e Cui.

tais como:

1. $P(x, y)$, que significa “x é parte de y”;
2. $O(x, y)$, “x se sobrepõe a y”²;
3. $PP(x, y)$, “x é parte própria de y”;
4. $Pi/2$ e $PPi/2$ são relações inversas de $P/2$ e $PP/2$ respectivamente;
5. $DC(x, y)$, cujo significado é “x é desconexo de y”;
6. $EQ(x, y)$, “x é igual a y”;
7. $PO(x, y)$, “x se sobrepõe parcialmente a y”;
8. $EC(x, y)$, “x está externamente conectado a y”;
9. $TPP(x, y)$, “x é parte própria tangencial de y”;
10. $NTPP(x, y)$, “x é parte própria não tangencial de y”;
11. $TPPi/2$ e $NTPPi/2$ são relações inversas de $TPP/2$ e $NTPP/2$ respectivamente.

O conjunto $\{DC(x, y), EQ(x, y), PO(x, y), EC(x, y), TPP(x, y), NTPP(x, y), TPPi/2, NTPPi/2\}$ é composto por relações mutuamente excludentes e disjuntas par a par. Este conjunto é usualmente conhecido por RCC8 e as definições destas relações são apresentados a seguir:

$$(\forall x \forall y (DC(x, y) \leftrightarrow \neg C(x, y))). \quad (1)$$

$$(\forall x \forall y (P(x, y) \leftrightarrow (\forall z (C(z, x) \rightarrow C(z, y)))). \quad (2)$$

$$(\forall x \forall y (PP(x, y) \leftrightarrow (P(x, y) \wedge \neg P(y, x)))). \quad (3)$$

$$(\forall x \forall y (EQ(x, y) \leftrightarrow (P(x, y) \wedge P(y, x)))). \quad (4)$$

$$(\forall x \forall y (O(x, y) \leftrightarrow (\exists z (P(z, x) \wedge P(z, y)))). \quad (5)$$

$$(\forall x \forall y (PO(x, y) \leftrightarrow (O(x, y) \wedge \neg P(x, y) \wedge \neg P(y, x)))). \quad (6)$$

$$(\forall x \forall y (DR(x, y) \leftrightarrow \neg O(x, y))). \quad (7)$$

$$(\forall x \forall y (EC(x, y) \leftrightarrow (C(x, y) \wedge \neg O(x, y)))). \quad (8)$$

$$(\forall x \forall y (TPP(x, y) \leftrightarrow (PP(x, y) \wedge (\exists z (EC(z, x) \wedge EC(z, y)))))). \quad (9)$$

$$(\forall x \forall y (NTPP(x, y) \leftrightarrow (PP(x, y) \wedge \neg (\exists z (EC(z, x) \wedge EC(z, y)))))). \quad (10)$$

$$(\forall x \forall y (Pi(x, y) \leftrightarrow P(y, x))). \quad (11)$$

$$(\forall x \forall y (PPi(x, y) \leftrightarrow PP(y, x))). \quad (12)$$

$$(\forall x \forall y (TPPi(x, y) \leftrightarrow TPP(y, x))). \quad (13)$$

$$(\forall x \forall y (NTPPi(x, y) \leftrightarrow NTPP(y, x))). \quad (14)$$

Transições entre relações RCC8 estão representadas no diagrama conceitual de vizinhança (CND)³ representado na figura 1. Diagramas conceituais de vizinhança são ferramentas fundamentais da área de raciocínio espacial qualitativo (Freksa, 1991). Informalmente, um CND é um grafo cujos vértices representam relações espaciais entre objetos e cujas arestas representam as *transições contínuas* entre estas relações. Neste contexto, transições contínuas significam que entre vértices adjacentes (ou seja, entre duas relações separadas por uma e somente uma aresta) não há uma terceira relação que caracterize o estado dos objetos presentes nos argumentos dessas relações.

²Do inglês: *overlaps*.

³Do inglês: *conceptual neighbourhood diagram*

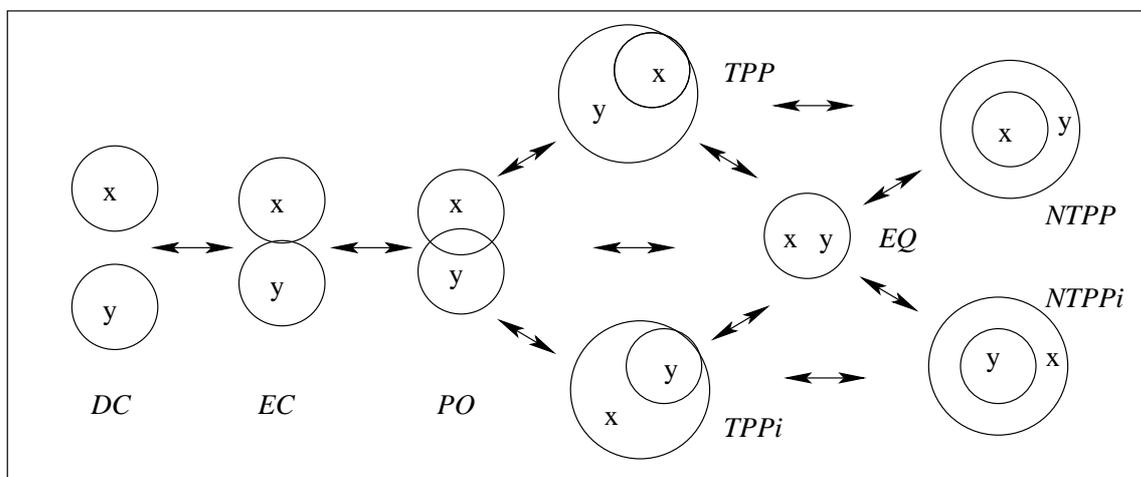


Figura 1: Conjunto de relações RCC8 e o seu diagrama conceitual de vizinhança.

	DC	EC	PO	TPP	NTPP	TPPi	NTPPi	EQ
DC	no info.	DR,PO,PP	DR,PO,PP	DR,PO,PP	DR,PO,PP	DC	DC	DC
EC	DR,PO,PPi	DR,PO,TPP	DR,PO,PP	EC,PO,PP	PO,PP	DR	DC	EC
PO	DR,PO,PPi	DR,PO,PPi	no info.	PO,PP	PO,PP	DR,PO,PPi	DR,PO,PPi	PO
TPP	DC	DR	DR,PO,PP	PP	NTPP	DR,PO,TPP	DR,PO,PPi	TPP
NTPP	DC	DC	DR,PO,PP	NTPP	NTPP	DR,PO,PP	no info.	NTPP
TPPi	DR,PO,PPi	EC,PO,PPi	PO,PPi	PO,TPP	PO,PP	PPi	NTPPi	TPPi
NTPPi	DR,PO,PPi	PO,PPi	PO,PPi	PO,PPi	O	NTPPi	NTPPi	NTPPi
EQ	DC	EC	PO	TPP	NTPP	TPPi	NTPPi	EQ

Figura 2: Tabela de composição para RCC8; a entrada *no info.* representa o caso em que nenhuma das relações do RCC8 são excluídas.

Outra ferramenta importante do raciocínio espacial qualitativo são as tabelas de composição (*composition tables*) (Allen, 1983). Dadas duas relações entre pares de objetos $\{a \text{ e } b\}$, e $\{b \text{ e } c\}$ (por exemplo $R_1(a, b)$ e $R_2(b, c)$), a célula da tabela relativa a $R_1(a, b)$ e $R_2(b, c)$ informa o conjunto minimal de disjunções $R_3(a, c)$ das possíveis relações entre a e c .

A figura 2 representa a tabela de composição para RCC8, em que a primeira linha representa as relações $R_1(a, b)$ e a primeira coluna representa $R_2(b, c)$ (para relações $R_1/2$ e $R_2/2$ do RCC8 e regiões espaciais a, b e c). A disjunção das possíveis relações entre a e c , a partir de $R_1(a, b)$ e $R_2(b, c)$ é dada pela célula que é intersecção da linha $R_1(a, b)$ e da coluna $R_2(b, c)$.

O cálculo de conexão de regiões foi aplicado, por exemplo, em simulações qualitativas (Cui et al., 1992), na representação de formas (Gotts, 1994), em teorias qualitativas de movimento (Wolter and Zakharyashev, 2000)(Muller, 1998), como teoria base para o modelamento de relações entre observador e objeto, como oclusão (Randell et al., 2001)(Randell and Witkowski, 2002)(Köhler, 2002). A formalização de oclusão com RCC8 será descrita a seguir.

Axiomatizando oclusão

Nessa seção apresentaremos algumas das teorias de raciocínio espacial que tem como objetivo a formalização da oclusão a partir do ponto de vista de um observador.

Geometria do espaço visual

A geometria do espaço visual (Petrov and Kuzmin, 1996) é uma formalização de oclusão espacial a fim de definir as bases fenomenológicas das estruturas geométricas do espaço visual.

A escolha da oclusão como elemento principal da teoria se apóia em três observações sobre a interposição de objetos: pri-

meiro, oclusão é um fenômeno puramente visual, isto é, não ocorre em outros canais sensoriais; segundo, é um fenômeno geral, pois ocorre independentemente da forma, significado ou função dos objetos. E, finalmente, oclusões múltiplas ocorrem com frequência em nossa percepção visual do mundo.

É introduzido em (Petrov and Kuzmin, 1996) um conjunto de axiomas sobre oclusão a partir de relações ternárias entre de pontos (denominadas localizações). Dado um conjunto de localizações X , uma relação de oclusão em X é definida como a relação ternária $O \subset X^3$: $\langle x, y, z \rangle$, que informalmente significa “o ponto y oclui os pontos x e z ”. No caso em que, dados três pontos, u, v e w , ocorre o caso $(u, v, w) \notin O$, então denota-se $\overline{\langle x, y, z \rangle}$.

Para quaisquer $x, y \in X$, $\langle x, x, y \rangle$ e $\langle y, x, x \rangle$, as relações da geometria do espaço visual obedecem aos seguintes axiomas (descritos informalmente):

- há pelo menos três pontos em X que não formam uma tríade.
- para quaisquer três pontos distintos $x, y \in X$ existem pontos $z_1, z_2, z_3 \in X$ tais que $\langle z_1, x, y \rangle, \langle x, z_2, y \rangle$ e $\langle x, y, z_3 \rangle$.
- se $x \neq y \neq z \in X$ e $\langle x, y, z \rangle$ então $\overline{\langle y, x, z \rangle}$ e $\overline{\langle x, z, y \rangle}$.
- para qualquer tríade $\langle x, y, z \rangle$, é verdade que $\langle z, y, x \rangle$ é também uma tríade.
- se $\langle x_1, x_2, x_3 \rangle$ e $\langle x_2, x_3, x_4 \rangle$ então $\langle x_1, x_2, x_4 \rangle$ e $\langle x_1, x_3, x_4 \rangle$.
- se $\langle x_1, a, b \rangle$ e $\langle x_2, a, b \rangle$ então ou $\langle x_1, x_2, a \rangle$ ou $\langle x_2, x_1, a \rangle$.

De acordo com (Petrov and Kuzmin, 1996), o conjunto de axiomas acima permite a definição de intervalos, raios e linhas. Um resultado interessante desta teoria é que vários teoremas da geometria Euclideana podem ser provados a partir destas relações de oclusão.

Cálculo de linhas de visão

O cálculo de conexão de regiões apresentado acima representa relações mereotopológicas entre regiões espaciais que são independentes do ponto de vista de um observador. Em contraste, (Galton, 1994) propõe o cálculo das linhas de visão (*lines-of-sight calculus*) a fim de representar as posições relativas entre pares de corpos convexos não sobrepostos.

Há 14 relações qualitativamente distintas no cálculo de linhas de visão. Dado x e y objetos no mundo, as relações são:

1. $C(x, y)$, “ x está livre de y ” (*is clear from*);
2. $JC(x, y)$, “ x acabou de ficar livre de y ” (*is just clear from*);
3. $PH(x, y)$, “ x esconde parcialmente y ” (*partially hides*);
4. $PHI(x, y)$, “ x está parcialmente escondido por y ” (*is partially hidden by*);
5. $JH(x, y)$, “ x acabou de esconder y ” (*just hides*);
6. $JHI(x, y)$, “ x acabou de ser escondido por y ” (*is just hidden*);
7. $H(x, y)$, “ x esconde y ” (*hides*);
8. $HI(x, y)$, “ x está escondido por y ” (*is hidden by*);
9. $EH(x, y)$, “ x esconde exatamente y ” (*exactly hides*);
10. $EHI(x, y)$, “ x está exatamente escondido por y ” (*is exactly hidden*);
11. $F(x, y)$, “ x está a frente de y ” (*is in front of*);

12. $FI(x, y)$, “ x tem y em sua frente” (*has y in front of it*);
13. $JF(x, y)$, “ x está imediatamente a frente de y ” (*is just in front of*);
14. $JFI(x, y)$, “ x tem y a sua frente” (*has y just in front of it*);

Este conjunto de relações, e seu diagrama conceitual de vizinhança, estão representados na figura 10.

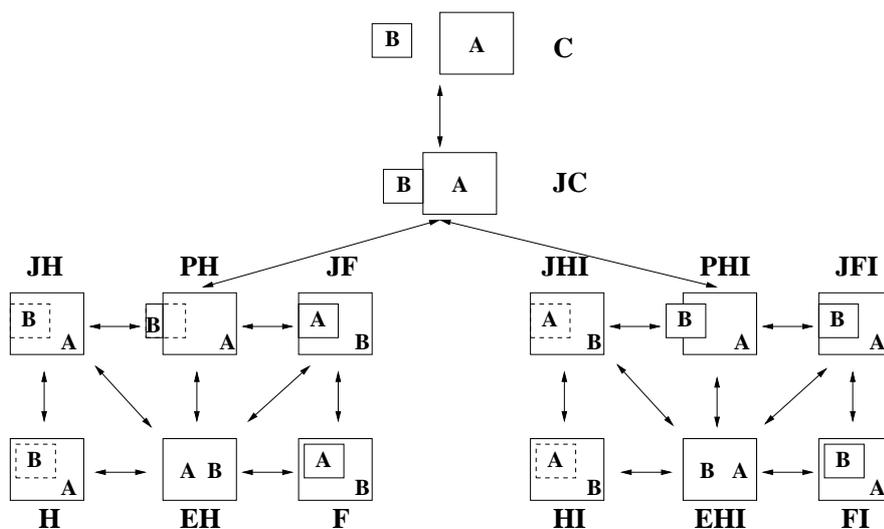


Figura 3: As relações do cálculo de linhas de visão.

O cálculo de linhas de visão e a geometria do espaço visual, entretanto, sofrem por não apresentar explicitamente o ponto de vista de algum observador. Isso é levado em consideração no cálculo de oclusão de regiões descrito a seguir.

Cálculo de oclusão de regiões

Inspirado pelo cálculo de linhas de visão, o cálculo de oclusão de regiões (ROC^4) (Randell et al., 2001) foi definido como uma extensão do RCC a fim de representar as várias possibilidades de interposição entre corpos de formas arbitrárias. As relações básicas do ROC estão representadas na figura 4.

De maneira a distinguir corpos físicos das suas regiões de ocupação e das suas projeções a respeito de um ponto de vista específico, ROC assume as seguintes funções: *region* e *image*. A função *region* pode ser entendida como um mapeamento de um corpo físico à região espacial em que ele ocupa (a sua *região de ocupação*). A função *image* é um mapeamento de um corpo físico à sua projeção bidimensional a partir de um ponto de vista específico.

Além da relação de conexão do RCC ($C/2$), ROC utiliza a relação primitiva $TotallyOcludes(x, y, \nu)$, que significa: “o objeto x oclui totalmente o objeto y a partir do ponto de vista ν ”. $TotallyOcludes/3$ é restrita pelos seguintes axiomas:

$$\forall x \forall \nu \neg TotallyOcludes(x, x, \nu) \tag{15}$$

$$\forall x \forall y \forall z \forall \nu (TotallyOcludes(x, y, \nu) \wedge TotallyOcludes(y, z, \nu)) \rightarrow TotallyOcludes(x, z, \nu) \tag{16}$$

Os seguintes axiomas introduzem o RCC8 nesta teoria.

⁴Do inglês: *Region Occlusion Calculus*

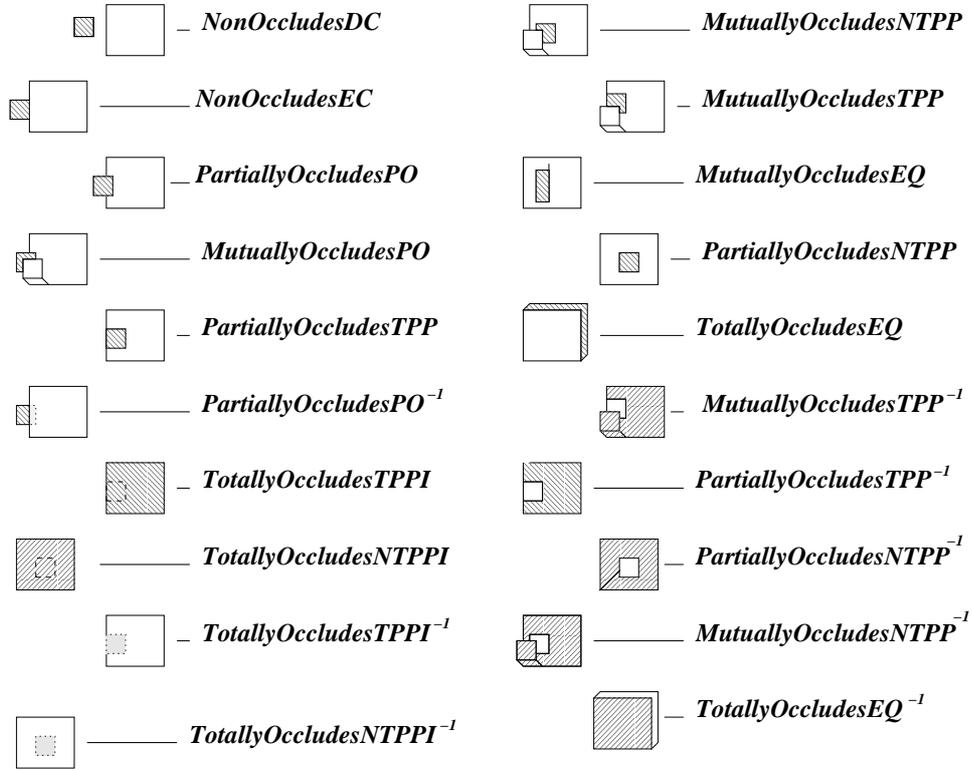


Figura 4: As relações básicas do cálculo de oclusão de regiões.

$$\forall x \forall y \forall z \forall \nu ((TotallyOccludes(x, y, \nu) \wedge P(region(z), region(y))) \rightarrow TotallyOccludes(x, z, \nu)) \quad (17)$$

$$\forall x \forall y \forall \nu (TotallyOccludes(x, y, \nu) \rightarrow \forall z (P(region(z), region(y)) \rightarrow \neg TotallyOccludes(z, x, \nu)) \quad (18)$$

$$\forall x \forall y \forall \nu (TotallyOccludes(x, y, \nu) \rightarrow \forall z \forall u (P(region(z), region(x)) \wedge \quad (19)$$

$$P(region(u), region(y)) \rightarrow \neg TotallyOccludes(u, z, \nu))$$

$$\forall x \forall \nu \exists y \exists z (P(region(y), region(x)) \wedge P(region(z), region(x)) \wedge TotallyOccludes(y, z, \nu)) \quad (20)$$

$$\forall x \forall y \forall \nu (TotallyOccludes(x, y, \nu) \rightarrow P(image(y, \nu), image(x, \nu))) \quad (21)$$

A relação *Occludes/3* (formula 22) é uma noção mais fraca de oclusão do que aquela apresentada pelos axiomas 15 ao 21, pois permite a definição de oclusão parcial e mútua

$$Occludes(x, y, \nu) \leftrightarrow \exists z \exists u (P(region(z), region(x)) \wedge P(region(u), region(y)) \wedge \quad (22)$$

$$TotallyOccludes(z, u, \nu))$$

Não oclusão, oclusão parcial e oclusão mútua podem, então, ser definidas pelas seguintes fórmulas:

$$PartiallyOccludes(x, y, \nu) \leftrightarrow Occludes(x, y, \nu) \wedge \neg TotallyOccludes(x, y, \nu) \wedge \neg Occludes(y, x, \nu) \quad (23)$$

$$MutuallyOccludes(x, y, \nu) \leftrightarrow Occludes(x, y, \nu) \wedge Occludes(y, x, \nu) \quad (24)$$

$$NonOccludes(x, y, \nu) \leftrightarrow \neg Occludes(x, y, \nu) \wedge \neg Occludes(y, x, \nu) \quad (25)$$

$$\forall x \forall y \forall \nu (NonOccludes(x, y, \nu) \rightarrow DR(image(x, \nu), image(y, \nu))) \quad (26)$$

$$\forall x \forall y \forall \nu (PartiallyOccludes(x, y, \nu) \rightarrow \quad (27)$$

$$(PO(image(x, \nu), image(y, \nu)) \vee PP(image(x, \nu), image(y, \nu))))$$

$$\forall x \forall y \forall \nu (MutuallyOccludes(x, y, \nu) \rightarrow \quad (28)$$

$$(PO(image(x, \nu), image(y, \nu)) \vee P(image(x, \nu), image(y, \nu))) \vee$$

$$PI(image(x, \nu), image(y, \nu)))$$

Axiomas para as relações inversas de *Occludes/3*, *TotallyOccludes/3*, *PartiallyOccludes/3* e *TotallyOccludes/3*, e para as demais relações apresentadas na figura 4 (mas não axiomatizados em pelas fórmulas do ROC acima) foram omitidos da presente descrição, mas suas descrições são similares às apresentadas acima.

Além de prover uma caracterização formal de relações de oclusão, ROC pode ser utilizado em um sistema de raciocínio automático sobre os estados dos objetos do domínio conforme eles (ou o observador) se movem no ambiente. Utilizando simulações qualitativas (Cui et al., 1992), é possível definir métodos para prever quando um objeto irá desaparecer atrás de outro (com relação a um observador), e para planejar as ações que o observador deverá executar para reaver a visão do objeto. Entretanto, o cálculo de oclusão de regiões define interposição de objetos como um conceito estático. Em ROC transições temporais entre as suas relações são introduzidas a partir de axiomas de visualização⁵ (Cui et al., 1992)(Weld and de Kleer, 1990). Tais axiomas são construídos em ROC a partir da introdução de variáveis para instantes de tempo e de um predicado *HoldsAt(a, t)*, definido para uma relação de oclusão *a* e um instante *t*.

O predicado *HoldsAt(a, t)* é intensamente utilizado em teorias sobre mudanças qualitativas, conforme discutido brevemente a seguir.

MUDANÇAS QUALITATIVAS EM RACIOCÍNIO ESPACIAL

Uma discussão aprofundada sobre mudanças qualitativas em raciocínio espacial é apresentada em (Galton, 2000). Em particular o raciocínio de conexão de regiões é utilizado em (Galton, 2000) para representar diversos tipos de movimento, tais como uma região entrando em outra, ou duas regiões (inicialmente distintas) entrando em contato.

O modelo temporal de (Galton, 2000) assume ambos, instantes e intervalos de tempos. Portanto, um estado do mundo pode ser verdade em um instante ou durante um intervalo de tempo, e esta distinção é incorporada na teoria por dois predicados: *Holds_T* que representa um estado verdadeiro em um instante de tempo, e *Holds_I* reservado para estados verdadeiros em um intervalo de tempo. A partir disso oito tipos distintos de transição entre pares de estados são definidos para cobrir todas as possibilidades de combinação entre dois estados no tempo.

O trabalho de Galton abriu caminho para o desenvolvimento de várias teorias qualitativas sobre o movimento.

Teorias qualitativas sobre movimento

Nas teorias qualitativas sobre movimento propostas em (Muller, 2002) e (Muller, 1998) o cálculo de conexão de regiões (apresentado na Seção 2 acima) é utilizado como linguagem básica na definição de relações espaço-temporais. Tais relações representam a história do desenvolvimento espacial dos objetos e, portanto, esta teoria trata de eventos e objetos de forma análoga.

Uma família de lógicas para representação de regiões espaço-temporais no tempo é investigado em (Wolter and Zakharyashev, 2000), em que o cálculo de conexão de regiões é definido em termos do sistema modal S4 e integrado a uma lógica

⁵Do inglês: *envisioning axioms*

temporal que formaliza relações “desde quando” (*since*) e “até quando” (*until*).

O problema de caracterizar o comportamento espaço-temporal de objetos de forma qualitativa também foi atacado em investigações sobre banco de dados relacionais (Erwig and Schneider, 2002; Erwig, 2004), em que predicados que envolvem topologia e lógica temporal são definidos a fim de garantir a integração de vários conjuntos de dados em bases de dados sobre espaço. Isso é feito a partir de um conjunto de predicados espaço-temporais (chamados predicados canônicos) que nada mais são do que as relações sobre entidades espaciais associadas ao tempo. Os predicados canônicos podem ser combinados para definir estruturas complexas chamadas *desenvolvimentos*, que representam a história dos objetos da base de dados.

ESPAÇO E DEFAULT REASONING

Nesta parte do artigo descreveremos a teoria sobre formas e espaço proposta em (Shanahan, 1995) e (Shanahan, 1996).

É *evidente* que qualquer ação, cujos efeitos são mudanças na posição de um objeto, possui como pre-condição que a destinação deste objeto esteja desocupada. Entretanto, na área de representação do conhecimento em IA, o uso da expressão “é *evidente*” usualmente implica que estamos utilizando nosso conhecimento de bom senso sobre o mundo. A representação e o raciocínio a respeito do conhecimento de bom senso é uma das principais áreas de investigação em representação de conhecimento em inteligência artificial. Um tema central desta área é o raciocínio sobre conhecimento padrão (ou *default*) e os trabalhos descritos em (Shanahan, 1995) e (Shanahan, 1996) foram os primeiros a abordar esse tema em raciocínio espacial: a fim de formalizar a pre-condição sobre a desocupação do local de destinação de um objeto, antes deste ser movido para lá, assume-se que o espaço é vazio por *default*.

Uma outra possibilidade de se representar esse padrão é descrever (à força bruta) todas as regiões do espaço que estão desocupadas. Porém isso acarreta em uma série de problemas, por exemplo:

- Uma descrição total deve ser feita para cada estado do mundo;
- Toda região desocupada deve ser descrita;
- Pode não haver uma descrição completa dos objetos e suas localizações;
- Pode ocorrer da informação de um determinado agente sobre a descrição de um objeto no mundo estar errada e, portando, deve ser revisada quando novos dados estiverem disponíveis.

Considerando estes problemas, a definição de uma regra padrão sobre ocupação espacial é a melhor opção. Esta regra é denominada *lei de bom senso sobre ocupação espacial* e é definida formalmente a partir do predicado *Occupies*/2. Data uma região espacial g e um objeto w , a sentença $Occupies(w, g)$ representa o fato de w ocupar a região g .

Occupies deve obedecer duas restrições de domínio que expressam o conhecimento de bom senso de que “nenhum objeto (w) ocupa duas regiões (g_1 e g_2) ao mesmo tempo” (formula 29) e que “não há dois objetos distintos (w_1 e w_2) ocupando a mesma região do espaço ao mesmo tempo” (formula 30).

$$HoldsAt(Occupies(w, g_1), t) \wedge \quad (29)$$

$$HoldsAt(Occupies(w, g_2), t) \rightarrow g_1 = g_2$$

$$HoldsAt(Occupies(w_1, g_1), t) \wedge \quad (30)$$

$$HoldsAt(Occupies(w_2, g_2), t) \wedge w_1 \neq w_2 \rightarrow$$

$$\neg \exists p (p \in g_1 \wedge p \in g_2)$$

A fórmula 31 representa que o espaço é vazio por *default*. Isto é, os objetos existentes no estado inicial são exceções (ou anomalias⁶) no espaço.

⁶Do inglês: *abnormalities*.

$$AbSpace(w) \leftarrow Initially(Occupies(w, g)) \quad (31)$$

O passo final para definir a *lei de bom senso sobre ocupação espacial* é garantir que só fazem parte do modelo da teoria as extensões de *AbSpace* cujos objetos foram explicitamente descritos em suas regiões de ocupação (ou seja, o padrão do espaço é ser vazio, objetos são exceções). Esta restrição nas extensões de *AbSpace* é obtida em (Shanahan, 1995) por uma política de circunscrição que minimiza este predicado, deixando o predicado *Initially* variar.

Cálculo de direções cardinais

O cálculo de direções cardinais (CDC) (Frank, 1996) é um formalismo para raciocínio automático sobre direções entre objetos espaciais. O conjunto de relações básicas deste cálculo é composto de nove relações: *norte*, *sul*, *leste*, *oeste*, *nordeste*, *noroeste*, *sudeste*, *sudoeste* e *EQ* (dados x e y , $EQ(x, y)$ significa que “ x está na mesma direção de y ”). A tarefa central do CDC é executar inferências sobre a direção relativa entre dois objetos A e B a partir da direção entre A e um outro objeto C (distinto de A e B). Por exemplo, dado $norte(A, B)$ e $nordeste(B, C)$, a tarefa é calcular as relações possíveis entre A e C .

Em (Ligozat, 1998) uma rede binária de restrições é desenvolvida a partir deste cálculo, em que direções cardinais são definidas como linhas bidimensionais entre dois objetos pontuais em um espaço 2D. Assim, o problema de raciocínio sobre direções cardinais é reduzido para um problema de propagação de restrições. Provas de complexidade computacional e sub-conjuntos tratáveis são explorados em (Ligozat, 1998).

O *double-cross calculus* (Freksa, 1992; Scivos and Nebel, 2001) é um cálculo de direções que permite definir a direção de um ponto com relação a um segmento de reta orientado. Este formalismo é definido por 15 relações ternárias sobre pontos, relacionando todas as possibilidades de composições entre as direções *frente-trás* e *esquerda-direita*.

Mais complexas formalizações de direções espaciais qualitativas envolvem a orientação de pontos em um espaço tridimensional (Pacheco et al., 2002), a definição de relações dinâmicas entre direções (Ragni and Wölfl, 2006), entre outros.

Similar ao *double-cross calculus*, (Santos et al., 2008) propõe o sistema cardinal local (SCL) para definir sistemas de referência a respeito da região ao redor de objetos que são postos sobre uma mesa. Esse sistema foi utilizado no aprendizado automático de atenção visual a partir da observação de arranjos de blocos.

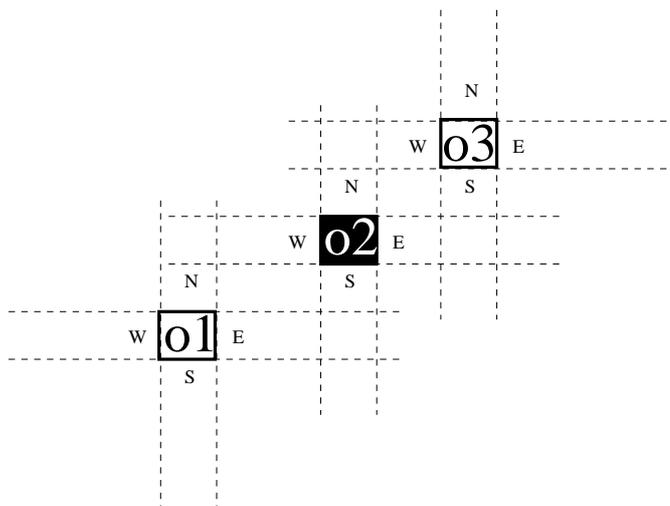


Figura 5: Sistema cardinal local.

O sistema cardinal local funciona da seguinte forma. Cada objeto que é posto sobre a mesa define o seu próprio sistema de coordenadas cardinais que é utilizado para localizar outros objetos que são colocados em sua vizinhança. Isto é, cada

objeto é localizado com relação ao seu vizinho mais próximo. As direções cardinais de cada SCL são limitadas pelos pontos extremos dos objetos de referência e são dependentes do ponto de vista do observador. Por exemplo, as regiões norte e sul de um objeto são limitadas por duas linhas paralelas, cada uma das quais passam pelos pontos extremos da direita e da esquerda deste objeto e são direcionadas verticalmente com relação ao ponto de vista do observador. Ou seja, as direções norte/sul serão sempre ortogonais à direção de observação do objeto (conforme mostra a figura 5). De forma análoga, as direções oeste e leste de um SCL são limitadas por duas linhas paralelas, cada uma das quais passando pelo pontos extremos de cima e de baixo do objeto de referência, e que são perpendiculares à linhas definindo as regiões norte e sul. As direções nordeste, sudeste, noroeste e sudoeste são definidas de forma similar. A seguinte restrição é assumida em sistemas cardinais locais: objetos já postos sobre a mesa não são descritos no SCL de novos objetos (restrição de precedência). Assim, há uma noção de precedência temporal implícita na definição de SCLs. A descrição do mundo em termos da vizinhança local dos objetos, somada à restrição de precedência, garante uma representação bastante eficiente dos objetos no espaço.

2.1 Outros cálculos espaciais

Há diversos outros cálculos espaciais cujas descrições detalhadas deixamos para um trabalho futuro. Esta seção descreve brevemente teorias qualitativas sobre tamanho e distância e teorias a respeito da forma de objeto.

Teorias qualitativas sobre tamanho e distância tratam estes aspectos, ou utilizando uma escala absoluta (em que relações entre objetos como *maior/menor do que* são definidas da forma usual), ou de forma relativa, onde a conexão relativa entre três objetos é utilizada para definir os conceitos de *equidistância* e *proximidade* (Cohn and Renz, 2008). Cálculos sobre tamanho e distância são usualmente associados a outros formalismos de raciocínio espacial a fim de estender a expressividade deles.

A representação qualitativa de formas é talvez um dos problemas menos compreendidos na área de raciocínio espacial qualitativo, conforme afirmam (Galton, 2000) e (Cohn and Hazarika, 2001). Alguns exemplos clássicos da literatura que tratam do assunto são (Schlieder, 1996; Clementini and Felice, 1997; Cohn, 1995; Meathrel and Galton, 2001; Gotts, 1994). Em geral, a descrição de formas é feita a partir da definição de algumas *primitivas* que representam características das bordas, do interior, ou de alguns aspectos globais dos objetos (Cohn and Renz, 2008).

É importante mencionar aqui que a combinação de cálculos espaciais é também um tema importante desta área.

2.2 Eficiência dos cálculos espaciais

Há diversos artigos que descrevem sub-conjuntos tratáveis de cálculos espaciais (Renz and Nebel, 1999; Jonsson and Drakengren, 1997). Em (Cohn and Renz, 2008) são apresentados quatro *ingredientes* para se encontrar sub-conjuntos tratáveis, são eles:

1. um método para provar que um dado subconjunto é tratável;
2. um método para se propor possíveis subconjuntos tratáveis (os quais devem ser verificados pelo item anterior);
3. para provar que um dado conjunto de relações é tratável, deve-se provar que a adição de qualquer relação (não contida nele), deixa o conjunto intratável;
4. deve-se, por fim, provar que o subconjunto obtido é maximal.

Os autores em (Cohn and Renz, 2008) descrevem uma série de métodos que podem ser utilizados como partes desses *ingredientes*. Em particular, vale citar aqui o trabalho proposto em (Renz, 2007), que descreve um algoritmo capaz de identificar automaticamente subconjuntos tratáveis a partir de teorias espaciais qualitativas.

2.3 Principais aplicações de raciocínio espacial qualitativo

As principais aplicações computacionais de formalismos de raciocínio espacial qualitativo estão em sistemas de informações geográficas (GIS)⁷, robótica inteligente, e visão cognitiva.

⁷Do inglês *geographical information systems*.

Em GIS, formalismos espaciais são aplicados para adicionar a estes sistemas um nível conceitual de representação e de raciocínio automático, a fim de facilitar a interação entre usuários e as bases de dados geográficos. O objetivo não é somente representar informações geográficas, para a compreensão do usuário, mas efetuar inferências a partir dessas informações (que seriam inviáveis de se executar com os dados brutos). Esse projeto, tem como principal dificuldade a representação consistente de conceitos vagos utilizados em linguagem natural (Santos et al., 2005).

Aplicações de raciocínio espacial qualitativo em robótica e visão cognitiva tem também o objetivo de incorporar a esses sistemas um nível conceitual, não somente para representar os dados brutos de forma mais legível para uma pessoa, mas para efetivamente permitir a interação entre o processamento de baixo nível e representações de bom senso dos objetos e eventos cotidianos. Exemplos destas aplicações podem ser encontrados em (Santos et al., 2008; Needham et al., 2005; Santos, 2007; Santos et al., 2009). Uma descrição mais detalhada da importância do REQ em visão computacional, para a interpretação de cenas, é dada na próxima seção. Outras aplicações de REQ são descritas em (Cohn and Renz, 2008).

3 INTERPRETAÇÃO QUALITATIVA DE CENAS

Bertrand Russell escreveu em seu célebre texto de 1914 que

(...) what is actually given in sense is much less than most people would naturally suppose, and (...) much of what at first sight seems to be given is really inferred. This applies especially in regard to our space-perceptions. For instance, we unconsciously infer the "real" size and shape of a visible object from its apparent size and shape, according to its distance and our point of view. (...). Thus, the first step in the analyses of data, namely, the discovery of what is really given in sense, is full of difficulty.

[B. Russell (1914), **Our Knowledge of the External World**, pp.75-76]

Apesar das idéias de Russell estarem nas origens dos formalismos de raciocínio espacial qualitativo, a interpretação de cenas em termos destes formalismos ainda é um problema em questão. Nesta seção descreveremos algumas propostas que (indiretamente) abordam este problema.

3.1 Avaliação de sequências de imagens

Sistemas desenvolvidos para interpretar sequências de imagens a partir de conceitos datam da década de 70 (Nagel, 1977). Desde então, várias propostas foram desenvolvidas. Exemplos são o sistema ALVEN (Tsotsos et al., 1980) para a abstração de conceitos sobre movimento a partir de sequências de imagens do coração humano; o sistema VITRA (Herzog, 1995)(Herzog and Wazinski, 1994) cujo propósito é unir interpretação de sequências de imagens com processamento de linguagem natural. A descrição de cenas de tráfego de veículos, utilizando fluxo ótico, foi investigada em (Gerber et al., 2002), (Nagel, 2000), (Nagel, 1988) e (Bouthemy and François, 1993).

O sistema descrito em (Frank et al., 1996) propõe uma solução ao rastreamento de veículos utilizando conhecimento sobre o contexto. Um conjunto de predicados a respeito de oclusão é introduzido, junto com predicados sobre o contexto. Uma cena é interpretada a partir da comparação do seu desenvolvimento temporal com um diagrama de transição que representa as possíveis mudanças entre objetos.

Algumas propostas clássicas de interpretação de imagens são caracterizadas pelo uso de atributos físicos das cenas observadas. Um desses sistemas (Brand, 1997) utiliza a estrutura causal dos elementos das cenas, isto é, como eles interagem e respondem a forças. Causalidade física é obtida pela análise da conectividade e do espaço livre entre os objetos da cena. A maior parte desta pesquisa foi aplicada à interpretação do funcionamento de máquinas (Brand, 1997)(Brand et al., 1993), ou ao reconhecimento de imagens para a manipulação de objetos.

De forma análoga, o sistema desenvolvido em (Siskind, 1995; Mann et al., 1997; Siskind, 2000) descreve movimento de objetos observados por vídeo a partir de noções como *suporte*, *contato* e *ligação*. Este sistema segmenta cenas em eventos distintos e as classifica em termos de tipos de eventos (ou verbos de movimento): *jogando*, *atirando*, *apanhando* e *deposi-tando*. Tais relações são assimiladas por um processo de simulação contra-factual, isto é, eventos futuros são previstos a partir

da simulação de efeitos qualitativos de mudanças hipotéticas aplicadas nas descrições do mundo. Quatro restrições de bom senso (inspiradas por (Hayes, 1984)) são assumidas neste processo: a restrição de *substancialidade* impõe que um objeto não pode atravessar por dentro de outro, a restrição de *continuidade*, diz que qualquer mudança na posição de um objeto é devida somente a movimento contínuo, a restrição de *gravidade*, indica que objetos não suportados por uma base caem, e a restrição de *base no solo* representa o fato de que o solo suporta todos os objetos.

Um arcabouço para a compreensão de cenas baseado em espaços conceituais (Gärdenfors, 2000) é proposto em (Chella et al., 2000), em que a tarefa de interpretação de cenas é reduzida a três processos básicos. Em um primeiro nível, a *área sub-conceitual* é responsável pelo processamento dos dados dos sensores para gerar uma descrição da cena dinâmica observada. O resultado deste processo é entrada para o segundo nível, o *nível conceitual*, cujo objetivo é gerar descrições de alto-nível da cena. Este estágio utiliza o cálculo de situações (que será descrito na Seção 4.2) (Reiter, 2001), cujas definições básicas constituem o terceiro módulo do sistema de interpretação de cenas: a *área linguística*.

3.2 Interpretação de cenas em lógica

A primeira abordagem para a interpretação de imagens em lógica foi proposta em (Reiter and Mackworth, 1989), descrita brevemente a seguir.

A abordagem de Reiter e Mackworth

Na abordagem descrita em (Reiter and Mackworth, 1989) são introduzidos três conjuntos de axiomas: *axiomas de imagem*, que representam restrições do domínio, *axiomas da cena*, reservados para objetos na cena, e *axiomas de representação* os quais restringem o mapeamento entre elementos na imagem e objetos na cena. Assim, o processo de interpretação de imagem é definido como a tarefa de determinar todos os modelos proposicionais dos axiomas, via um sistema de satisfação de restrições. A interpretação de uma determinada cena é, portanto, um dos modelos dos axiomas escolhido a partir de algum critério de preferência entre o conjunto de modelos possíveis.

Esse arcabouço foi aplicado ao problema da interpretação de esboços de mapas a partir de um ponto de vista global. Nesse contexto, o domínio da imagem é constituído por duas primitivas: *cadeias (chains)*, representando curvas bidimensionais no espaço, e *regiões*, para regiões espaciais. Um conjunto de relações foram então definidas a respeito destas primitivas. Dadas duas cadeias c e c' , e uma região r , as seguintes relações primitivas foram definidas (tal qual representado na figura 6):

- $tee(c, c')$, representa a intersecção das cadeias c e c' em uma junção “ T ”;
- $chi(c, c')$, c e c' se encontram em uma junção “ χ ”;
- $bounds(c, r)$, significa que a cadeia c limita a região r ;
- $closed(c)$, representa que a cadeia c é uma figura fechada simples;
- $interior(c, r)$: o interior da cadeia fechada c é a região r ; e
- $exterior(c, r)$, representa que o exterior da cadeia fechada c é a região r .

O arcabouço para interpretação de imagem proposto por Reiter e Mackworth funciona sob as seguintes hipóteses básicas: hipótese de mundo fechado (*closed world assumption*) e a hipótese da unicidade dos nomes (*unique names assumption*). A primeira rege que os predicados do domínio da imagem são todos conhecidos (isto é, todos as relações verdadeiras sobre elementos da imagem são descritas, as não descritas são falsas), a segunda indica que dois símbolos iguais sempre referem-se aos mesmos objetos.

Analogamente ao domínio da imagem, o domínio da cena considera dois tipos de objetos primitivos: *objeto linear* e *objeto área*. No domínio da interpretação de esboços de mapa, objetos lineares representam *estradas*, *rios* ou *margem*, enquanto objetos área são *terreno* ou *corpo-d’água*. Fatos gerais sobre este domínio são representados por um conjunto de axiomas de bom-senso representando coisas como “dois rios nunca se cruzam”, “rios não formam laços”, “se uma estrada ou um rio margeam uma área, então esta área é um terreno”, entre outras.

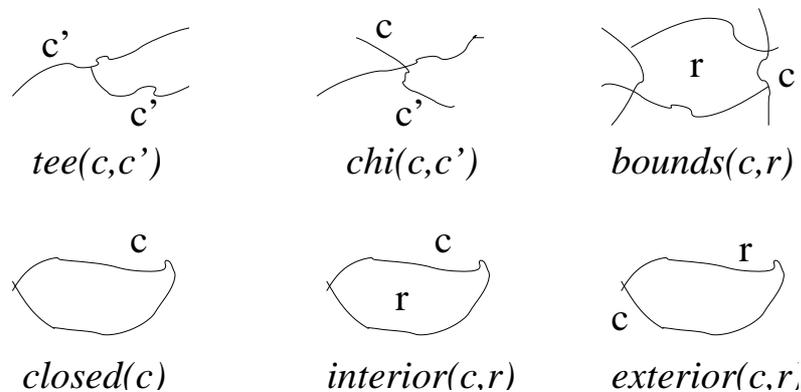


Figura 6: Relações primitivas do domínio da imagem.

Com isso, o domínio da imagem é relacionado ao domínio da cena pelos axiomas *de mapeamento imagem-cena*. Os axiomas de mapeamento *imagem-cena* dizem informalmente o seguinte:

1. A *representação* consiste somente de objetos da imagem e objetos de cena;
2. Toda imagem i refere-se a um e somente um objeto;
3. Todo objeto da cena é representado por um e somente um objeto de imagem;
4. Qualquer região (ou cadeia) na imagem relaciona-se a uma área (ou objeto linear) na cena.

Conforme mencionam os próprios autores, as hipóteses 2 e 3 são muito fortes. De fato, 2 implica que ruídos na imagem representam objetos físicos, enquanto que 3 ignora oclusão de objetos na cena. Em uma seção posterior de (Reiter and Mackworth, 1989), a hipótese 3 é relaxada a fim de formalizar o significado de uma ponte ocluindo parcialmente parte de uma estrada (ou rio) no domínio da interpretação de mapas. Assim, a hipótese 3 é modificada para expressar que todo objeto na cena seja representado por pelo menos um objeto na imagem.

Sistemas baseados em Hipótese-teste

Inspirado pelo trabalho de Reiter e Mackworth descrito acima, e pelo raciocínio baseado em hipótese proposto em (Poole et al., 1987), o sistema SIGMA (Matsuyama and Hwang, 1990) foi desenvolvido para interpretar imagens aéreas de zonas urbanas a fim de localizar regiões específicas, como estradas e ocupações residenciais. Uma linguagem para expressar propriedades geométricas de objetos para este sistema foi proposta em (Schroeder and Neumann, 1996). Este trabalho deu origem a sistemas recentes para a representação e assimilação de fatos em multimídia (Möller and Neumann, 2008).

Há também sistemas baseados em um procedimento de hipótese-teste sobre teorias de raciocínio espacial qualitativo, o RCC em particular (Hazarika and Cohn, 2002; Cohn et al., 2002). Em (Hazarika and Cohn, 2002) é proposto um modelo para a assimilação de descrições qualitativas do mundo (observações) a partir de informações qualitativas sobre histórias espaço-temporais (como descrito na seção 2). Este processo de interpretação de cenas é guiado por restrições de continuidade (com relação às histórias espaço-temporais de cada objeto do domínio) que são baseadas no diagrama conceitual de vizinhança do RCC, e por uma biblioteca de padrões específicos de comportamento para objetos rígidos. Esta biblioteca compõe parte da base de conhecimento para o processo de formação de hipótese. Do conjunto de hipóteses possíveis para explicar uma cena qualquer, a hipótese preferida é aquela que prevê o menor número de mudanças no mundo.

O trabalho descrito em (Cohn et al., 2002) propõe uma arquitetura para aprendizagem de modelos de eventos, a partir de dados de visão computacional. Tais modelos são posteriormente utilizados em um processo de interpretação de imagens. Esta arquitetura é baseada em propostas anteriores para a obtenção automática de modelos de eventos a partir de visão computacional (Ferryhough et al., 1998; Galata et al., 2002). Mais recentemente estas idéias deram origem a um sistema

capaz de interpretar imagens a partir da consistência das histórias espaço-temporais dos objetos (Bennett et al., 2005) e a um outro sistema capaz de aprender (por meio de programação em lógica indutiva) as regras e o comportamento de jogadores envolvidos em um jogo de cartas (Needham et al., 2005).

4 DESENVOLVENDO UM CÁLCULO ESPACIAL SOBRE INFORMAÇÕES DE SENSORES

Nesta seção descreveremos as idéias principais do desenvolvimento de um sistema de raciocínio espacial qualitativo baseado em informações provenientes de um sistema de visão estéreo.

4.1 Cálculo de perfis de profundidade

Em (Santos, 2007) é introduzida uma representação simbólica dos dados provenientes de um sistema de visão estéreo de um robô móvel. Esta formalização está baseada em uma simplificação dos dados, em que o domínio (tal qual visto pelo robô móvel) é representado por um *corte horizontal* da cena. Cada um destes cortes horizontais é um perfil bidimensional da cena vista pelo robô. Objetos no mundo aparecem como *picos de profundidade* nesses perfis. Esses picos (como regiões espaciais em RCC) são os objetos primitivos do cálculo de perfis de profundidade (CPP). Perfis e picos de profundidade estão representados na figura 7.

O eixo PIXELS nos perfis de profundidade representa os pixels presentes nos cortes horizontais das cenas, enquanto que o eixo vertical representa a disparidade⁸ de cada ponto por onde passa o corte. Este eixo é limitado pela distância L (figura 7b), que representa o ponto máximo de confiabilidade dos dados da visão estéreo.

A figura 7(a) representa uma visão global de uma cena que contém dois objetos (A e B), onde o ponto de vista do robô é representado pelo símbolo ν e o seu campo de visão pelas linhas tracejadas. Os símbolos α , β e L na figura 7(a) representam, respectivamente, as disparidades dos objetos A e B e o limite máximo da visão estéreo do robô. A figura 7(b) representa o perfil de profundidade extraído da cena na Figura 7(a). P e Q são picos relativos aos objetos A e B.

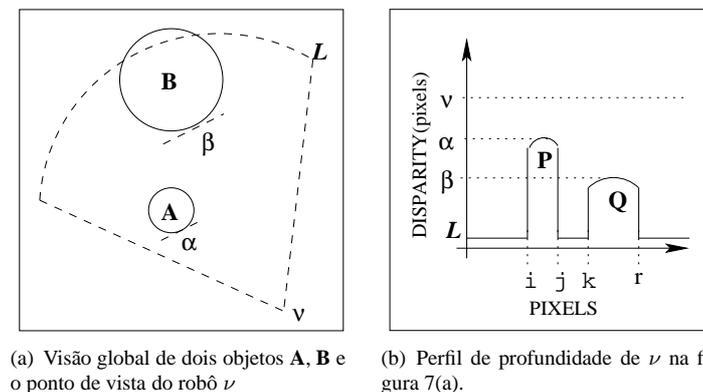


Figura 7: Perfil de profundidade.

As bordas dos picos nos perfis de profundidade representam as bordas dos objetos pelos quais passa o corte horizontal (como representado pelos símbolos i, j, k e r na figura 7(b)). Além disso, o perfil de profundidade também contém informações sobre o tamanho relativo dos objetos visuais e a distância relativa entre objetos, visto da perspectiva do observador. O valor de tamanho é dado pelo módulo da diferença entre os valores das bordas dos picos, enquanto que a distância entre picos é dada pelo módulo da diferença entre as bordas mais próximas de dois picos distintos. A partir disso, assumindo a geometria da câmera estéreo, é possível obter uma aproximação do valor real destas grandezas. Isso é feito em (Souchanski and Santos, 2008), porém a teoria desenvolvida em (Santos, 2007) baseia-se nos valores *subjetivos* (isto é, dependentes do observador) destas grandezas.

⁸Disparidade é inversamente proporcional à profundidade.

A teoria CPP é construída a partir da função $p_o/2$, que representa o pico de um objeto em particular⁹. Esta função mapeia um símbolo para um corpo físico e um instante de tempo à um símbolo para um pico de profundidade ($p_o: \text{physical body} \times \text{time point} \rightarrow \text{peak}$). Portanto, dado um símbolo para pico p e um objeto (corpo físico) b , a expressão $p = p_o(b, t)$ representa que “ p é um pico do objeto b no instante t ”.

CPP também possui funções para atributos de picos: *disparity* (que retorna a disparidade de um pico), *size* (que retorna o tamanho de um pico) e *distance* (que retorna a distância relativa entre pares picos).

Similar ao *default* sobre ocupação espacial (discutido na seção 2 acima), este trabalho assume que *todo pico em um perfil representa um objeto no mundo, a menos que haja evidência do contrário*.

A partir destes elementos básicos, a teoria CPP formaliza as seguintes relações sobre picos de profundidade:

1. *extending*(p, t), representa que a disparidade do pico p está aumentando no tempo t ;
2. *shrinking*(p, t), a disparidade do pico p está diminuindo no tempo t ;
3. *vanishing*(p, t), a disparidade de p , no instante t , está diminuindo até p não ser mais detectado pelos sensores;
4. *appearing*(p, t) é a relação complementar a *vanishing*/2, e representa que um pico (em um instante t) está aparecendo no fundo do perfil de profundidade;
5. As próximas duas relações representam que dois picos p e q estão se aproximando um ao outro no instante t , a partir do ponto de vista do observador:
 - *peak_approaching_DC*(p, q, t) representa o evento de dois picos p e q se aproximarem um ao outro mas mantendo-se *desconectados*; e,
 - *peak_approaching_EC*(p, q, t) representa o evento de dois picos p e q se aproximarem um ao outro, até se tornarem *externamente conectados*;
6. analogamente à *peak_approaching*, define-se *peak_receding_DC*/3 e *peak_receding_EC*/3. O primeiro representa o caso de dois picos *desconectados* se afastarem um do outro, enquanto que o segundo representa o caso de dois picos, inicialmente conectados, se afastarem um do outro;
7. *peak_coalescing*(p, q, t), significa que os picos p e q estão se fundindo em um no instante t ;
8. *going_out_sight*(p, q, t), significa que o pico p (ou o q) desaparece do campo de visão no instante t ;
9. *peak_splitting*(p, q, t), representa o fato de um pico se dividir em dois;

As relações sobre picos descritas acima são axiomatizadas em (Santos, 2007) em cláusulas de Horn, sujeitos corpos são transições específicas entre os atributos dos picos, dadas por mudanças nos valores das funções *disparity*, *size* e *distance* entre dois instantes de tempo subsequentes. As fórmulas 32 e 33 são exemplos de axiomas CPP para as relações *extending* e *approaching* respectivamente, em que a e b são objetos e t_1, t_2 e t são instantes de tempo.

$$\text{extending}(p_o(a), t) \leftrightarrow \exists t_1 t_2 (t_1 < t) \wedge (t < t_2) \wedge (\text{disp}(p, t_1) < \text{disp}(p, t_2)). \quad (32)$$

A fórmula 32 diz que se um pico p (representando um objeto a no instante t) teve um aumento de seus valores de disparidade entre dois perfis de profundidade subsequentes, então este pico está se *expandindo* (*extending*).

⁹Do inglês: *peak of*

$$\begin{aligned}
\text{peak_approaching}(p_o(a), p_o(b), t) \leftrightarrow & \\
& \exists t_1 t_2 (t_1 < t) \wedge (t < t_2) \wedge DC(p_o(a), p_o(b), t_1) \wedge \\
& DC(p_o(a), p_o(b), t_2) \wedge \\
& (\text{dist}(p_o(a), p_o(b), t_2) < \text{dist}(p_o(a), p_o(b), t_1))
\end{aligned}
\tag{33}$$

O evento de dois picos se aproximando um do outro (representado por *peak_approaching/3*) em um instante t é definido em CPP pela diminuição da distância angular entre eles em dois perfis de profundidade subsequentes, conforme formalizado na fórmula 33.

Diagramas conceituais de vizinhança para as relações do CPP estão representados nas figuras 8 e 9. Na prática, estes diagramas conceituais podem ser implementados como uma tabela. Assim, acessar as relações e transições entre relações CPP reduz-se a uma simples busca em tempo constante.

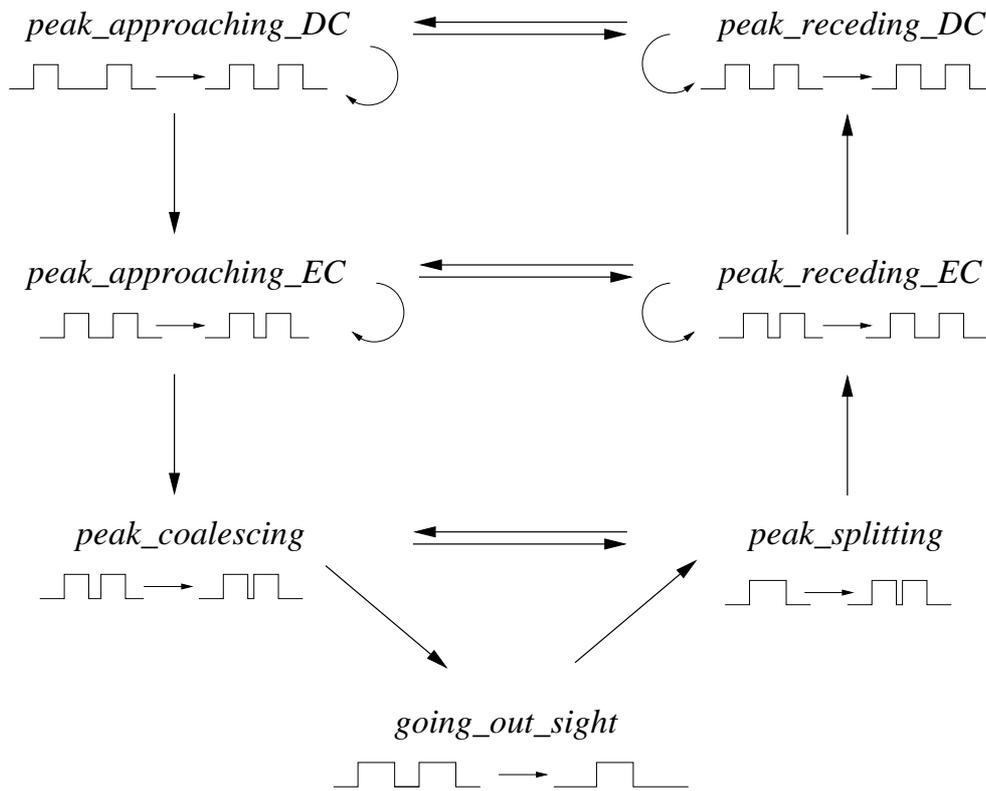


Figura 8: Diagrama conceitual de vizinhança para relações sobre pares de picos.

De fato, cada aresta conectando dois predicados nos diagramas das figuras 8 e 9 deveriam ser teoremas da teoria subjacente. A prova destes teoremas, entretanto, não é trivial e exige uma teoria mais forte, que trate rigorosamente das dependências temporais dos predicados. A conjectura 1 apresenta o enunciado formal destes teoremas. A teoria necessária para prová-los será descrita na próxima seção.

Conjectura 1 (Transições no diagrama conceitual de vizinhança) Seja ϕ_i e ϕ_j duas relações distintas do conjunto de relações do CPP; p e q quaisquer picos distintos e t um instante de tempo. Existe uma aresta direcionada entre $\phi_i(p, q, t)$ e $\phi_j(p, q, t')$ nos diagramas conceituais de vizinhança (figuras 8 e 9) em um instante t' (subsequente a t) se e somente se em qualquer modelo para os axiomas do CPP existir uma transição contínua de $\phi_i(p, q, t)$ a $\phi_j(p, q, t')$.

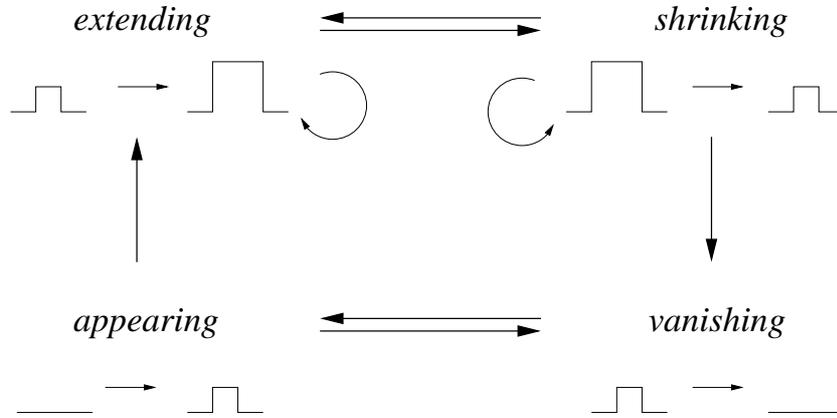


Figura 9: Diagrama conceitual de vizinhança para relações sobre picos únicos.

4.2 Perfis de profundidade em Cálculo de situações

Conforme mencionado na seção anterior, para provar a conjectura 1 é necessário formalizar o cálculo de perfis de profundidade em uma teoria capaz de lidar rigorosamente com ações e mudança de estado. Esses resultados foram desenvolvidos em (Souchanski and Santos, 2008), em que CPP foi formalizado em cálculo de situações, conforme brevemente introduzido a seguir.

Cálculo de Situações

O cálculo de situações (CS) é uma formalização em lógica de primeira ordem de aspectos do raciocínio sobre ações e mudanças e, portanto, possui como elementos fundamentais *ações* e *fluentes* (Reiter, 2001).

Ações são termos representados por símbolos funcionais e seus argumentos. Por exemplo, o termo $endMove(R, loc(3, 8), loc(5, 10), 105.7)$ denota a ação de um agente móvel finalizando sua locomoção da posição $loc(3, 8)$ para a posição $loc(5, 10)$ no instante 105.7. Além de ações que mudam as propriedades do mundo, há também ações que alteram o conhecimento do agente sobre o mundo (*ações sensoriais*), por exemplo, a ação $sense(p, loc(x_r, y_r), t)$, executada no instante t , com a posição do agente $loc(x_r, y_r)$, obtém uma observação do mundo como o perfil de profundidade p .

Uma *situação* denota uma sequência de estados. Estas sequências são representadas por uma função binária $do/2$: $do(\alpha, s)$ denota a sequência de estados resultantes da aplicação da ação α à sequência s . O símbolo S_0 denota o estado inicial.

Relações ou funções cujos valores variam de situação em situação são chamados de *fluentes*.

O cálculo de situações distingue um predicado $Poss(a, s)$ que, associado à um conjunto de axiomas, caracteriza as condições para uma ação a ser executada no estado s .

Além dos axiomas de condição de ações, o SC possui axiomas que formalizam mudanças de estado no domínio: os *axiomas de estados sucessores* (AES). Há um AES para cada fluente relacional $F(\vec{x}, s)$:

$$F(\vec{x}, do(a, s)) \equiv \Phi_F(\vec{x}, a, s), \text{ onde } \Phi_F(\vec{x}, a, s) \text{ é uma fórmula com variáveis livres } a, s, \vec{x} \text{ e expressa da seguinte maneira:}$$

$$a = PositiveAction \wedge \gamma^+(\vec{x}, s) \vee \dots$$

$$\vee F(\vec{x}, s) \wedge \neg(a = NegativeAction \wedge \gamma^-(\vec{x}, s) \vee \dots),$$

em que *PositiveAction* é uma ação que tem um efeito positivo no fluente F (isto é, faz o fluente ficar verdadeiro) e $\gamma^+(\vec{x}, s)$ é uma fórmula que define o contexto em que esse efeito ocorre. De maneira análoga, *NegativeAction* é uma ação negativa sobre F se $\gamma^-(\vec{x}, s)$ é verdade na situação s . Esse axioma define o valor verdade de F na situação $do(a, s)$ em termos da situação anterior s . Os axiomas de estados sucessores são elementos fundamentais para umas das mais elegantes

soluções para o problema do quadro (*frame problem*) (Reiter, 2001).

4.3 Perfis de profundidade em cálculo de situações

O aspecto mais importante da formalização do CPP em cálculo de situações são os axiomas de estados sucessores sobre os predicados do cálculo de perfis de profundidade. Nestes axiomas, as relações representando transições na percepção de profundidade são combinadas com ações que causaram tais mudanças. Assim um agente inteligente, munido desta teoria, é capaz de descrever observações do mundo, bem como raciocinar sobre os efeitos de suas próprias ações e aquelas executadas por outros agentes. Uma apresentação completa de todos esses axiomas, porém, esta fora do escopo deste artigo, entretanto (conforme descrito em (Souchanski and Santos, 2008)) descreveremos o pseudo-código do AES para *extending* e *approaching*, este último pode ser tanto *peak_approaching_DC* quanto *peak_approaching_EC* (cf. seção 4.1).

Dado $extending(peak, viewpoint, do(a, s))$ representando que o pico *peak*, relativo a algum objeto, observado pelo ponto de vista *viewpoint*, é percebido como *em expansão* (*extending*) na situação $do(a, s)$. O AES para esta relação (em pseudo-código) é o seguinte:

$extending(peak, viewpoint, do(a, s))$ **se e somente se**
a é uma ação sensorial que mediu o tamanho angular de *peak* como maior na situação corrente do que no estado anterior *s* **ou**
a é uma ação *endMove* que termina o movimento do agente resultando em um ponto de vista em que o tamanho computado do pico *peak* é maior no estado corrente, do que em *s* **ou**
a é uma ação *endMove* terminando o movimento do objeto para uma nova posição tal que, visto do ponto de vista do agente, o tamanho computado do pico *peak* é maior do que era na situação *s* **ou**
 $extending(peak, viewpoint, s)$ e % frame axiom %
a não é nenhuma das ações capazes de alterar o tamanho de *peak*

Em outras palavras, um pico de profundidade é percebido como *extending* na situação $do(a, s)$ se e somente se houve uma ação sensorial que percebeu o tamanho angular do pico maior em $do(a, s)$ do que esta mesma grandeza em *s*; ou o robô (ou o observador) moveu-se para uma posição tal que o tamanho angular (calculado a partir da ação de movimento) do objeto é maior em $do(a, s)$ do que em *s*.

O AES para *approaching* expressa que dois picos de profundidade estão se aproximando um ao outro se e somente se o ângulo aparente entre eles, obtido por uma ação sensorial, é menor na situação $do(a, s)$ do que na situação anterior *s* ou, o observador (ou o objeto) se moveu para uma posição tal que a distância entre os objetos (calculada a partir da nova posição) é menor em $do(a, s)$ do que em *s*. Nesse caso, a distância entre os picos de dois objetos b_1 e b_2 é calculada por simples manipulações trigonométricas. O pseudo-código deste axioma é o seguinte:

$approaching(peak1, peak2, viewpoint, do(a, s))$ **se e somente se**
a é uma ação sensorial que mede o ângulo entre *peak1* e *peak2* e obtém um valor menor do que o valor do estado anterior *s* **ou**
a é uma ação *endMove* que termina o movimento do agente resultando em um ponto de vista cuja distância angular calculada entre *peak1* e *peak2* é menor no estado atual do que era em *s* **ou**
a é uma ação *endMove* terminando o movimento do objeto para uma nova posição tal que, do ponto de vista do agente, a distância angular calculada entre os picos decresceu em relação ao seu valor em *s* **ou**
 $approaching(peak1, peak2, viewpoint, s)$ e % frame axiom %
a não é nenhuma das ações que tem efeito sobre a aparente distância angular entre *peak1* e *peak2*.

A conjectura 1 é teorema do cálculo de perfil de profundidade formalizado em cálculo de situações, cuja prova está resumida a seguir (Souchanski and Santos, 2008).

A prova da conjectura 1 parte da observação de que os vértices do diagrama conceitual de vizinhança (e as arestas que os conectam) nas figuras 8 e 9 representam todos os perceptos que podem ser sentidos dadas as definições do cálculo de perfis de profundidade em um domínio em que observador e objetos podem se mover. Assim, pode-se dizer que percepção em CPP é correta e completa em relação ao movimento, isto é, os vértices e arestas dos diagramas conceituais de vizinhança do CPP (figuras 8 e 9) são resultado somente do movimento dos objetos (*percepção é correta*) e que qualquer movimento de objetos no mundo é representado por um vértice ou uma aresta dos diagramas conceituais de vizinhança (*percepção é completa*).

Teorema 1 (*Percepção é correta em relação ao movimento dos objetos*).

Esboço da prova:

Para cada vértice dos diagramas conceituais de vizinhança do CPP, se a última ação do robô não for uma ação sensorial, então o fluente representado por este vértice só pode mudar seu valor por uma ação *endMove*, de acordo com os axiomas de estado sucessor do CPP em cálculo de situações. De forma análoga podemos mostrar que para todas as arestas conectando duas relações distintas F e F' nas figuras 8 e 9, a transição é definida por uma ação de movimento cujo efeito resulte em uma situação tal que a relação (fluente) deixe de valer, mas F' torne-se verdade.

O próximo teorema expressa que qualquer movimento no domínio pode ser representado por um vértice ou aresta dos grafos das figuras 8 e 9.

Teorema 2 (*Percepção é completa com relação ao movimento dos objetos*).

Esboço da prova:

A prova segue da observação de que as 12 regiões definidas pelas bitangentes entre dois objetos (regiões numeradas na figura 10) definem todas os possíveis pontos de vista para observar estes dois objetos que são qualitativamente distintos (isto é, um observador localizado em uma dessas regiões sempre verá os objetos na mesma relação, enquanto que o mesmo não ocorre em regiões distintas). É fácil ver, portanto, que para qualquer movimento de um observador dentro de uma região, ou cruzando regiões adjacentes, na figura 10 existe uma ação A mencionada nos axiomas de estados sucessores do CPP que corresponde a este movimento. Assim, é consequência dos AESs que qualquer vértice ou aresta dos grafos da figuras 8 e 9 descreve a percepção resultante de movimento do mundo. Por exemplo, considere um robô localizado na região 5 (figura 10) observando os dois objetos a e b , mas se afastando deles. Os axiomas de estados sucessores permitiriam a conclusão que os picos de profundidade relativos a a e b estariam se aproximando (*approaching*) e diminuindo de tamanho (*shrinking*). Por outro lado, se este mesmo robô estivesse cruzando para a região 6 (a partir da região 5), os axiomas permitiriam inferir a transição de *approaching* para *coalescing*.

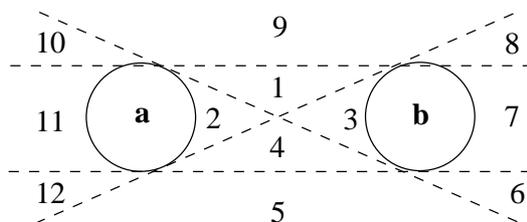


Figura 10: Bitangentes entre dois objetos visíveis

5 POSSÍVEIS DIREÇÕES FUTURAS

A seguir descrevemos as possíveis direções futuras para o desenvolvimento de sistemas de raciocínio espacial qualitativo em visão computacional e robótica.

Raciocínio com e sobre incerteza

Conforme apresentado na seção 3, durante os anos 70 e 80 houve grande interesse no desenvolvimento de sistemas de visão de alto nível, onde processamento numérico (ou quantitativo) alimenta um nível simbólico (ou qualitativo) de conhecimento a partir do qual um agente é capaz de interpretar o mundo, e agir de acordo com estas interpretações. Essas tentativas iniciais, porém, foram frustradas pela inexistência de algoritmos eficientes para tratamento de incertezas, de sistemas de representação de conhecimento tratáveis e também pelo estágio prematuro de desenvolvimento dos algoritmos de processamento de imagem.

Desde então, entretanto, avanços importantes em inteligência artificial indicam que esta área atingiu um estado de maturidade em que talvez seja possível a integração de sistemas de representação e raciocínio em IA com sistemas de visão computacional. Alguns desses avanços são: o desenvolvimento de redes Bayesianas (Pearl, 1988) (representações gráficas das variáveis de um domínio que garantem métodos eficientes de tratamento de incertezas) e o advento das lógicas de descrição (Baader et al., 2002), que são uma família de formalismos (sub-conjuntos da lógica de primeira ordem) que possuem um balanço favorável entre expressividade e complexidade. Recentemente tem havido um interesse crescente no desenvolvimento de lógicas de descrição probabilísticas, algumas das quais integram o poder de representação relacional das lógicas de descrição, com a eficiência no tratamento de incertezas das redes Bayesianas (Cozman and Polastro, 2008; Cozman and Polastro, 2009).

Portanto, a investigação de formalismos de raciocínio espacial qualitativo em lógicas de descrição probabilísticas deve ser o caminho a seguir para o desenvolvimento de sistemas de visão capazes de executar inferências em um nível de conhecimento simbólico e raciocinar com (e a respeito de) incerteza. Já há alguns trabalhos nessa direção (Santos, Hummel, Fenelon and Cozman, 2010).

Raciocínio espacial com objetos flexíveis

Em (Cabalar and Santos, 2010) foi proposto um primeiro arcabouço lógico para representar os estados, ações e mudança de estados em domínios espaciais contendo objetos não triviais como cordas e orifícios. Apesar do sucesso em aplicar este formalismo em domínios com diferentes arranjos de objetos, a prova formal de que as soluções obtidas eram consequência semântica da teoria proposta ficou para trabalhos futuros, bem como a investigação da aplicação prática óbvia do arcabouço proposto, que é o desenvolvimento de sistemas de raciocínio para a execução de cirurgias robóticas.

Acreditamos que esta prova resultará em um lema essencial para que sejam provados formalmente os limites do arcabouço lógico proposto, isto é, qual é exatamente a classe de problemas que este formalismo é capaz de representar e sobre a quais métodos de inferência podem ser executados.

Em particular, o desenvolvimento de formalismos de REQ sobre objetos flexíveis, como cordas, é um tema recente na área e deverá ser observado com maior interesse pela comunidade no futuro.

Sistemas de raciocínio automático para veículos autônomos

Uma futura aplicação importante para sistemas que combinam raciocínio espacial e visão computacional é a interpretação de cenas (e o controle autônomo) de um veículo automotivo. Nesse contexto, investigações futuras deverão se concentrar em uma série de questões, algumas delas listamos a seguir:

- a interpretação do comportamento de agentes móveis em um ambiente de tráfego de veículos real, tal qual observado por um dos veículos;
- tomada de decisões a partir do reconhecimento de sinais de trânsito;
- reconhecimento das propriedades funcionais das pistas (McCall and Trivedi, 2006) de uma rodovia (como as direções de tráfego permitidas, curvas possíveis, veículos permitidos em cada pista);

O desenvolvimento de sistemas que incluam soluções para estas questões abriria o caminho para a criação de novos dispositivos de auxílio ao condutor, como sistemas de navegação de grande precisão ou sistemas automáticos de mudança de faixas em estradas urbanas.

Raciocínio espacial para neuroimagem em doenças psiquiátricas

Há um grande número de sistemas de suporte à decisão em psiquiatria e, em particular, para o auxílio do diagnóstico da esquizofrenia (First et al., 1993; Bronzino et al., 1989; Spitzer and Endicott, 1968). Todos eles, entretanto, baseiam-se na formalização da sintomatologia da doença. Um desafio para trabalhos futuros em raciocínio espacial é desenvolver um sistema de representação de conhecimento que possua (como entidades elementares) regiões espaciais em neuroimagens que representam estruturas descritas na literatura médica como relacionadas à distúrbios psiquiátricos. Investigações iniciais nessa direção foram propostas em (Santos, Thomaz, dos Santos, Freire, Sato, Louza, Sallet, Busatto and Gattaz, 2010; Santos, Freire, Nunes, Thomaz, Sallet, Busatto and Louzã, 2010).

6 CONCLUSÃO

Raciocínio espacial está presente em quase todas as interações humanas no mundo real, desde dar um simples laço em um tênis até a execução de complicadas tarefas de navegação, ou a interpretação de cenas visuais. Entretanto, sistemas de visão computacional tem negligenciado desenvolvimentos recentes em raciocínio espacial qualitativo, sub-área de representação de conhecimento em inteligência artificial. A fim de fomentar a colaboração entre estas duas áreas, este artigo apresenta os principais formalismos de raciocínio espacial qualitativo (REQ) e discute alguns dos trabalhos centrais sobre a interpretação automática de cenas utilizando conceitos de alto nível. Em seguida apresenta em linhas gerais a construção de um formalismo de REQ baseado em informações de imagens de visão estéreo. Esperamos que este formalismo sirva como exemplo para futuros desenvolvimentos de teorias efetivas que, partindo de algum conjunto de primitivas provenientes de um sistema de visão, sejam capazes de interpretar e raciocinar sobre o conteúdo de cenas.

Agradecimentos

Partes deste trabalho foram discutidas largamente com Murray Shanahan, Mikhail Soutchanski e Hannah Dee durante visitas do autor ao Imperial College, London (UK), Ryerson University (Canada) e University of Leeds (UK) e foram motivo de publicações anteriores deste autor (Santos, 2007; Souchanski and Santos, 2008). O autor também agradece a Renata Favalli e à Valquiria Fenelon pelos comentários a respeito de uma versão preliminar deste texto.

O autor agradece suporte financeiro à FAPESP (Projeto temático 2008/03995-5 (LogProb)) e ao CNPq (bolsa Pq 303555/2008-4).

REFERÊNCIAS

- Allen, J. (1983). Maintaining knowledge about temporal intervals, *Communications of the ACM* **26**(11): 832–843.
- Baader, F., Calvanese, D., McGuinness, D., Nardi, D. and Patel-Schneider, P. (2002). *Description Logic Handbook*, Cambridge University Press.
- Bennett, B. (1997). *Logical Representations for automated reasoning about spatial relationships*, PhD thesis, School of computing, University of Leeds, UK.
- Bennett, B., Cohn, A. and Magee, D. (2005). Enforcing global spatio-temporal consistency to enhance reliability of moving object tracking and classification, *Künstliche Intelligenz* **2**: 32–35.
- Bouthemy, P. and François, E. (1993). Motion segmentation qualitative dynamic scene analysis from an image sequence, *IJCV* **10**(2): 157–182.
- Brand, M. (1997). Physics-based visual understanding, *Computer Vision and Image Understanding* **65**(2): 192–205.
- Brand, M., Birnbaum, L. and Cooper, P. (1993). Sensible scenes: Visual understanding of complex structures through causal analysis, *Proc. of AAAI*, Washington DC, U.S., pp. 588–593.
- Bronzino, J., Morelli, R. A., and Goethe, J. (1989). Overseer: a prototype expert system for monitoring drug treatment in the psychiatric clinic, *IEEE Transactions on Biomedical Engineering* **36**(5): 533–540.

-
- Cabalar, P. and Santos, P. E. (2010). Formalising the fisherman's folly puzzle, *Artificial Intelligence In Press, Corrected Proof*.
- Chella, A., M.Frixione and Gaglio, S. (2000). Understanding dynamic scenes, *Artificial Intelligence* **123**(1-2): 89–132.
- Clementini, E. and Felice, P. D. (1997). A global framework for qualitative shape description, *GeoInformatica* **1**(1): 11–27.
- Cohn, A. (1995). A hierarchical representation of qualitative shape based on connection and convexity, in A. M. Frank (ed.), *Proc. of COSIT*, pp. 311–326.
- Cohn, A. G., Bennett, B., Gooday, J. and Gotts, N. (1997). Representing and reasoning with qualitative spatial relations about regions, in O. Stock (ed.), *Spatial and Temporal Reasoning*, Kluwer Academic Publishers, pp. 97 – 134.
- Cohn, A. G. and Hazarika, S. M. (2001). Qualitative spatial representation and reasoning: An overview, *Fundamenta Informaticae* **46**(1-2): 1–29.
- Cohn, A. G., Magee, D. R., Galata, A., Hogg, D. C. and Hazarika, S. M. (2002). Towards an architecture for cognitive vision using qualitative spatio-temporal representations and abduction, *Proc. of Spatial Cognition III*, Lake Starnberg, Germany.
- Cohn, A. G. and Renz, J. (2008). Qualitative spatial representation and reasoning, *Handbook of Knowledge Representation*, Elsevier, pp. 551–596.
- Cozman, F. G. and Polastro, R. (2009). Complexity analysis and variational inference for interpretation-based probabilistic description logics, *Proc. of the Conference on Uncertainty in Artificial Intelligence*.
- Cozman, F. G. and Polastro, R. B. (2008). Loopy propagation in a probabilistic description logic, *SUM*, pp. 120–133.
- Cui, Z., Cohn, A. and Randell, D. (1992). Qualitative simulation based on a logic of space and time, *Proc. of AAAI*, California, U.S., pp. 679–684.
- Erwig, M. (2004). *Spatio-temporal Databases*, Springer, chapter Toward Spatiotemporal Patterns, pp. 29–54.
- Erwig, M. and Schneider, M. (2002). Spatio-temporal predicates, *IEEE Transactions on Knowledge and Data Engineering* **14**(4): 881–901.
- Fernyhough, J., Cohn, A. G. and Hogg, D. C. (1998). Building qualitative event models automatically from visual input, *Proc. of ICCV*, IEEE Computer Society Press, Bombay, India, pp. 350–355.
- First, M., Opler, L., Hamilton, R., Linder, J., Linfield, L., Silver, J., Toshav, N., Kahn, D., Williams, J., and Spitzer, R. (1993). Evaluation in an inpatient setting of dtree, a computer-assisted diagnostic assessment procedure, *Comprehensive psychiatry* **34**(3): 171–75.
- Frank, A. U. (1996). Qualitative spatial reasoning: Cardinal directions as an example, *International Journal of Geographical Information Science* **10**(3): 269–290.
- Frank, T., Haag, M., Kollnig, H. and Nagel, H.-H. (1996). Characterization of occlusion situations occurring in real-world traffic scenes, *Proc. of the Workshop on Conceptual Descriptions from Images, ECCV*, Cambridge, UK, pp. 43–57.
- Freksa, C. (1991). Conceptual neighbourhood and its role in temporal and spatial reasoning, *Decision Support Systems and Qualitative Reasoning*, Elsevier Science, pp. 181 – 193.
- Freksa, C. (1992). Using orientation information for qualitative spatial reasoning, *Theories and Methods of Spatial-Temporal Reasoning in Geographic Space*, Vol. 629 of LNCS, Springer-Verlag.
- Galata, A., Cohn, A., Magee, A. G. and Hogg, D. (2002). Modelling interaction using learnt qualitative spatio-temporal relations and variable length markov models, *Proc. of ECAI*, Lyon, France, pp. 741–745.
- Galton, A. (1994). Lines of sight, *Proc. of the Seventh Annual Conference of AI and Cognitive Science*, Dublin, Ireland, pp. 103–113.

- Galton, A. (2000). *Qualitative Spatial Change*, Oxford University Press.
- Gärdenfors, P. (2000). *Conceptual Spaces: the geometry of thought*, The MIT Press.
- Gerber, R., Nagel, H.-H. and Schreiber, H. (2002). Deriving textual descriptions of road traffic queues from video sequences, *Proc. of ECAI*, Lyon, France, pp. 736–740.
- Gotts, N. (1994). How far can we ‘C’? Defining a ‘doughnut’ using connection alone, *Proc. of KR*, Bon, Germany, pp. 246–257.
- Hayes, P. J. (1984). The second naïve physics manifesto, in J. Hobbs and R. C. Moore (eds), *Formal Theories of the Common Sense World*, Ablex.
- Hazarika, S. M. and Cohn, A. G. (2002). Abducing qualitative spatio-temporal histories from partial observations, *Proc. of KR*, Toulouse, France, pp. 14–25.
- Herzog, G. (1995). From visual input to verbal output in the visual translator, *Technical report*, Universitat des Saarlandes. Technical Report 124.
- Herzog, G. and Wazinski, P. (1994). VISual TRanslator: Linking perceptions and natural language descriptions, *Artificial Intelligence Review* 8(2/3): 175–187.
- Jonsson, P. and Drakengren, T. (1997). A complete classification of tractability in rcc-5, *J. Artif. Int. Res.* 6(1): 211–221.
- Köhler, C. (2002). The occlusion calculus, *Proc. of Cognitive Vision Workshop*, Zürich, Switzerland.
- Levesque, H. (1996). What is planning in the presence of sensing?, *Proc. of AAAI*, Vol. 2, Oregon, U.S., pp. 1139–1146.
- Ligozat, G. (1998). Reasoning about cardinal directions, *J. Vis. Lang. Comput.* 9(1): 23–44.
- Mann, R., Jepson, A. and Siskind, J. M. (1997). The computational perception of scene dynamics, *Computer Vision and Image Understanding: CVIU* 65(2): 113–128.
- Matsuyama, T. and Hwang, V. S. (1990). *SIGMA: A Knowledge-Based Image Understanding System*, Plenum Press, New York, U.S.
- McCall, J. C. and Trivedi, M. M. (2006). Video-based lane estimation and tracking for driver assistance: Survey, system, and evaluation, *IEEE Transactions on Intelligent Transportation Systems* 7:1: 20–37.
- Meathrel, R. C. and Galton, A. P. (2001). A hierarchy of boundary-based shape descriptors, *Proc. of IJCAI*, pp. 1359–1364.
- Möller, R. and Neumann, B. (2008). Ontology-based Reasoning Techniques for Multimedia Interpretation and Retrieval, *Semantic Multimedia and Ontologies : Theory and Applications*, Springer, pp. 55–98.
- Muller, P. (1998). A qualitative theory of motion based on spatio-temporal primitives, in A. G. Cohn, L. Schubert and S. C. Shapiro (eds), *Proc. of KR*, Morgan Kaufmann, California, U.S., pp. 131–141.
- Muller, P. (2002). Topological Spatio-Temporal Reasoning and Representation, *Computational Intelligence* 18(3): 420–450.
- Nagel, H.-H. (1977). Analysing sequences of tv-frames: System design considerations, *Proc. of IJCAI*, Cambridge, U.S., p. 626.
- Nagel, H.-H. (1988). From image sequences towards conceptual descriptions, *Image and Vision Computing* 6(2): 59–74.
- Nagel, H.-H. (2000). Image sequence evaluation: 30 years and still going strong, *Proc. of ICPR*, Barcelona, Spain, pp. 1149–1158.
- Needham, C., Santos, P., Magee, D., Devin, V., Hogg, D. and Cohn, A. (2005). Protocols from perceptual observations, *Artificial Intelligence* . accepted, pending minor revisions.
- Pacheco, J., Escrig, M. T. and Toledo, F. (2002). Qualitative spatial reasoning on three-dimensional orientation point objects, *Proceedings of the QR2002. 16th International Workshop on Qualitative Reasoning. Editors : Nuria Agell and*

-
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: networks of plausible inference*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
- Petrov, A. P. and Kuzmin, L. (1996). Visual space geometry derived from occlusion axioms, *Journal of Mathematical Imaging and Vision* **6**: 291–308.
- Poole, D., Goebel, R. and Aleliunas, R. (1987). Theorist: A logical reasoning system for defaults and diagnosis, in N. Cercone and G. McCalla (eds), *The Knowledge Frontier – Essays in the Representation of Knowledge*, Springer-Verlag, pp. 331–352.
- Ragni, M. and Wöfl, S. (2006). Temporalizing cardinal directions: From constraint satisfaction to planning, *Proc. of KR*, pp. 472–480.
- Randell, D. A., Cohn, A. G. and Cui, Z. (1992). Computing transitivity tables: A challenge for automated theorem provers, in D. Kapur (ed.), *Proc. of CADE*, LCNS, Springer Verlag, Saratoga Springs, U.S., pp. 786–790.
- Randell, D., Cui, Z. and Cohn, A. (1992). A spatial logic based on regions and connection, *Proc. of KR*, Cambridge, U.S., pp. 165–176.
- Randell, D. and Witkowski, M. (2002). Building large composition tables via axiomatic theories, *Proc. of KR*, Toulouse, France, pp. 26–35.
- Randell, D., Witkowski, M. and Shanahan, M. (2001). From images to bodies: Modeling and exploiting spatial occlusion and motion parallax, *Proc. of IJCAI*, Seattle, U.S., pp. 57–63.
- Reiter, R. (2001). *Knowledge in Action: Logical Foundations for Specifying and Implementing Dynamical Systems*, MIT Press, U.S.
- Reiter, R. and Mackworth, A. (1989). A logical framework for depiction and image interpretation, *Artificial Intelligence* **41**(2): 125–155.
- Renz, J. (2007). Qualitative spatial and temporal reasoning: Efficient algorithms for everyone, *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI-07)*, Hyderabad, India.
- Renz, J. and Nebel, B. (1999). On the complexity of qualitative spatial reasoning: a maximal tractable fragment of the region connection calculus, *Artif. Intell.* **108**(1-2): 69–123.
- Santos, P., Bennett, B. and Sakellariou, G. (2005). Supervaluation semantics for an inland water feature ontology, in L. P. Kaelbling and A. Saffiotti (eds), *Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI-05)*, Professional Book Center, Edinburgh, pp. 564–569.
- Santos, P., Dee, H. and Fenelon, V. (2009). Qualitative robot localisation using information from cast shadows, *IEEE Int. Conf. on Robotics and Automation. ICRA '09.*, pp. 220–225.
- Santos, P. E. (2007). Reasoning about depth and motion from an observer’s viewpoint, *Spatial Cognition and Computation* **7**(2): 133–178.
- Santos, P. E., Hummel, B., Fenelon, V. and Cozman, F. (2010). Probabilistic encoding of spatial domains, *Prof. of First International Workshop on Uncertainty in Description Logics (UniDL)*.
- Santos, P. E., Thomaz, C. E., dos Santos, D., Freire, R., Sato, J. R., Louza, M., Sallet, P., Busatto, G. and Gattaz, W. F. (2010). Exploring the knowledge contained in neuroimages: Statistical discriminant analysis and automatic segmentation of the most significant changes, *Artificial Intelligence in Medicine* **In Press, Corrected Proof**.
- Santos, P., Freire, R., Nunes, D., Thomaz, C., Sallet, P., Busatto, G. and Louzã, M. (2010). *Qualitative Spatio-Temporal Representation and Reasoning: Trends and Future Directions*, IGI Publishing, chapter A region-based ontology of the brain ventricular system and its relation to schizophrenia. to appear.
- Santos, P., Needham, C. and Magee, D. (2008). Inductive learning spatial attention, *Sba Controle & Automação* **19**(3): 316–326.

-
- Scherl, R. and Levesque, H. (1993). The frame problem and knowledge-producing actions, *Proc. of AAI*, Washington DC, U.S., pp. 689–695.
- Schlieder, C. (1996). Qualitative shape representation, in P. A. Burrough and A. U. Frank (eds), *Geographic Objects with Indeterminate Boundaries*, Taylor & Francis Inc., pp. 123–140.
- Schroeder, C. and Neumann, B. (1996). On the logics of image interpretation: Model construction in a formal knowledge representation framework, *International Conference on Image Processing*, Vol. 2, Switzerland, pp. 785–788.
- Scivos, A. and Nebel, B. (2001). Double-crossing: decidability and computational complexity of a qualitative calculus for navigation, in D. Montello (ed.), *Spatial Information Theory: Foundations of GIS*, Vol. 2205 of *LNCIS*, Springer, pp. 431 – 446.
- Shanahan, M. (1995). Default reasoning about spatial occupancy, *Artificial Intelligence* **74**(1): 147–163.
- Shanahan, M. (1996). Robotics and the common sense informatic situation, *Proc. of ECAI*, Budapest, Hungary, pp. 684–688.
- Shanahan, M. (1997). *Solving the frame problem*, MIT press, U.S.
- Siskind, J. M. (1995). Grounding language in perception, *Artificial Intelligence Review* **8**(5-6): 371–391.
- Siskind, J. M. (2000). Visual event classification via force dynamics, *Proc. of AAI*, Austin, U.S., pp. 149–155.
- Souchanski, M. and Santos, P. (2008). Reasoning about dynamic depth profiles, *Proc. of the 18th European Conference on Artificial Intelligence (ECAI)*, IOS, pp. 30–34.
- Spitzer, R. L. and Endicott, J. (1968). Diagno: computerized program for psychiatric diagnosis utilizing the differential diagnostic procedure, *Archives of General Psychiatry*.
- Stock, O. (ed.) (1997). *Spatial and Temporal Reasoning*, Kluwer Academic Publishers.
- Tsotsos, J. K., Mylopoulos, J., Covvey, H. D. and Zucker, S. W. (1980). A framework for visual motion understanding, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2**(6): 563–573. Special Issue on Computer Analysis of Time-Varying Imagery.
- Weld, D. and de Kleer, J. (eds) (1990). *Readings in Qualitative Reasoning about Physical Systems*, Morgan Kaufmann, San Mateo, U.S.
- Wolter, F. and Zakharyashev, M. (2000). Spatio-temporal representation and reasoning based on RCC-8, *Proc. of KR*, San Francisco, U.S., pp. 3–14.