

SISTEM PAKAR UNTUK MEMBANGUN SHELL HUGIN DENGAN KEPERCAYAAN BAYESIAN SECARA MENYELURUH

Siti Munawaroh dan Hari Murti

Fakultas Teknologi Informasi, Universitas Stikubank Semarang

Abstraksi:

Penyebab jaringan probabilistik telah dibuktikan, sebuah alat perwakilan pengetahuan berguna untuk model domain, dimana penyebab hubungan yang luas di rasa jalan alami yang berkaitan dengan objek domain dan di mana hubungan ketidakpastian ini adalah warisan. Di bawah ini merupakan garis besar pelaksanaan shell hugin untuk menangani sebuah model domain yang dinyatakan oleh penyebab jaringan probabilistik. Satu-satunya bentuk batasan yang ada pada jaringan itu, tidak boleh mengandung perulangan yang diarahkan.

Pendekatan ini digambarkan langkah demi langkah untuk memecahkan, masalah keberadaan genetik. Sebuah perwakilan grafik dari model domain interaktif yang dibuat menggunakan komponen dasar simpul-simpul jaringan dan bangunan sebagai blok. Struktur ini, bersama-sama dengan jumlah hubungan antar simpul dinyatakan sebagai penyebab probabilistik, secara otomatis diwujudkan ke dalam struktur pohon, yaitu sebuah pohon persimpangan. Di sini perhitungan yang efisien dan konsep sederhana algebra kepercayaan Bayesian menyeluruh mendukung penggabungan bukti baru, penyebaran informasi, perbaikan perhitungan dan kepercayaan di terminal dari simpul dalam jaringan. Akhirnya, sebagai contoh aplikasi nyata didunia, MUNIN pakar sistem untuk electromiographi akan dibahas.

Kata Kunci : Sistem Pakar, Jaringan Probabilistik, Kepercayaan Bayesian, *shell Hugin*

PENDAHULUAN

Dalam beberapa tahun terakhir, telah dilakukan pengembangan sistem pakar berdasarkan skema probabilistik dimana struktur domain pengetahuan dilambangkan dalam sebuah grafik berarah. Bertentangan dengan upaya probabilistik sebelumnya, pendekatan ini membuat struktur domain eksplisit dan memanfaatkan kontrol di topologi grafik. Probabilistik memiliki skema sebagai aturan dan berdasarkan sistem yang berbasis logika intrinsik. Hugin digunakan untuk menangani Ketidakpastian. Hugin dan Munin adalah dua istilah, yang digunakan sebagai singkatan untuk sebuah sistem pakar untuk electromiographi.

Penyebab jaringan probabilistik, yang berdasarkan kualitas dimana simpul mewakili objek domain dan hubungan antara simpul mewakili hubungan penyebab antar objek-objek.

Pada tingkat jumlah, dinyatakan oleh hubungan link yang diwakili oleh probabilitik bersyarat. Bersama metode yang berasal dari teori keputusan Bayesian, Kunci persoalan ini adalah kemampuan untuk mengurangi perhitungan, sejumlah daerah hanya dengan

menggunakan perhitungan variabel yang didapat dari satu objek dan obyek lain dalam struktur grafik. Sistem shell MUCIN menyediakan shell :

1. satu set alat-alat untuk pembuatan penyebab jaringan probabilistik yang secara bersamaan membuat grafik untuk model domain
2. satu set alat untuk memasukkan jumlah informasi tentang keadaan dan domain penyebaran dalam jaringan.

Konsep sederhana sebagai hubungan tunggal tanpa penyebab jaringan probabilistik. Kesederhanaan ini dicapai oleh transformasi yang terhubung ke satu set jaringan kepercayaan yang terorganisir sebagai pohon, persimpangan pohon, dan dengan menyediakan operasi untuk penyebaran informasi antara kepercayaan di pohon. Shell Hugin, mengilustrasikan contoh tentang genetika, melalui langkah-langkah utama :

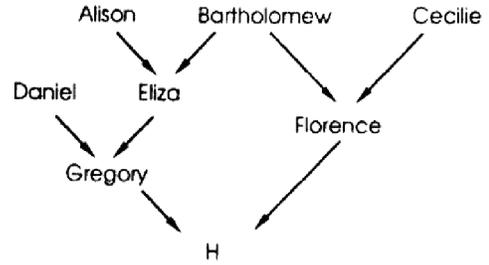
- menentukan model untuk membuat operasional,
- inferring kepercayaan berdasarkan data yang dimasukkan ke dalam model. menggunakan perhitungan dari kepercayaan Bayesian. Contoh Sistem MUNIN, sebuah sistem pakar untuk electromiografi.

PENGEMBANGAN

Dalam shell HUGIN mengambil contoh tentang masalah genetika. Florence dan Gregory melakukan hubungan perkawinan. Akan tetapi Gregory adalah keponakan dari Florence, dan dalam sejarah dari Bartholomew, ayah dari Florence dan kakek dari Gregory, penyakit mengancam kehidupan seperti hantu. Penyakit yang disebabkan dominan allele A, dan waktu munculnya penyakit ini di permukaan agak terlambat di kehidupan individu. Bartholomew menikah dua kali, maka Florence dan Gregory dari ibu yang setengah-saudara. Tidak kedua-duanya Florence Gregory, orangtua mereka, maupun Gregory dari nenek telah menunjukkan ada tanda-tanda penyakit. Apa yang dimaksud dengan risiko fatal yang akan timbul pada anak mereka yang akan mewarisi karakteristik penyakit tersebut

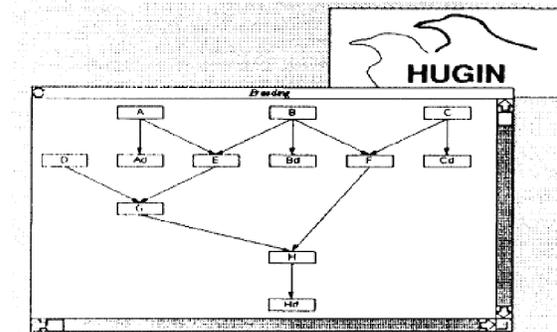
Model kualitatif

Langkah pertama dalam sesi Hugin adalah untuk merumuskan masalah dalam hal penyebab jaringan probabilistik. Contoh probabilitas adalah struktur silsilah yang akan kita ditampilkan dalam Gambar 1. Masing-masing diwakili oleh nama tersebut, dan hubungan orang tua-anak diwakili oleh panah dari orangtua ke anak. Catatan diagram yang berisi dua jalur yang berbeda dari Bartholomew ke H: melalui Eliza dan Gregory, dan melalui Florence. Secara umum, menentukan genotip dari individu-individu. Untuk masing-masing individu, mengasosiasikan sebuah node diinisialkan dan mengidentifikasi geno jenis yang diidentifikasi oleh kombinasi allele. Oleh karena itu, masing-masing individu yang tepat adalah salah satu genotipe-genotipe AA, Aa, atau aa.



Gambar 1: Contoh untuk Struktur silsilah

Di mana satu allele yang dominan, ini berarti bahwa genotipe-genotipe AA dan Aa pembawa dari penyakit, sedangkan aa bukan. Tapi dapat diamati apakah ada kehadiran penyakit atau tidak. tapi ini hanya dapat ditentukan bila individu telah mencapai usia dewasa disebabkan karakter tersembunyi dari penyakit. Untuk dapat memasukkan informasi yang relevan pada apa yang diamati dan harapan yang terjadi dari penyakit, kami menambahkan node penyakit dari kedatangan generasi. Node ini, diberi label oleh individu dengan indeks awal d, menyatakan ada dua kemungkinan, ya dan tidak ada, sesuai dengan kehadiran dan ketidakhadiran dari penyakit, masing-masing



Gambar 2. Tampilan HUGIN yang tidak berguna

Ini, masalah struktur kualitatif telah kita tentukan, dan kami berharap shell HUGIN. Shell ini adalah perantara, dan sebagian besar operasi yang dilakukan seperti pilihan petunjuk. Gambar 2 menunjukkan model layar utama yang dimasukkan. Disini spesifik domain adalah jendela label yang dibuat, dan model dibangun: hal ini dilakukan dengan memilih domain yang dibuat masuk pada

menu utama (terlampir ke simbol HUGIN). Semua node dan link dimasukkan, dan Simbol grafis dibuat. Untuk setiap node, tampilan nama dan state yang ada Memasuki state untuk node penyakit meskipun tidak diperlukan, karena menyatakan ya dan tidak berlaku standar. Jika dikehendaki demikian, model domain dapat disimpan pada titik tertentu, untuk kemudian dilanjutkan pada pemanggilan kembali.

Model Kuantitatif

Menyelesaikan model dengan menentukan bagian angka dari link. Hal ini dilakukan dengan maksud menentukan tabel kondisi kemungkinan. Untuk setiap simpul, memberikan syarat kemungkinan setiap state diberikan dari orang tua. Dengan demikian untuk node penyakit, telah kami ditetapkan dalam tabel sebagai Tabel 1.

Disease	Genotype		
	AA	Aa	aa
yes	1	1	0
no	0	0	1

Tabel 1: penyebab probabilistik untuk tabel hubungan genotip-penyakit. Setiap masukan menentukan probabilitas dari penyakit yang diberi genotipe.

P ₁	P ₂		
	AA	Aa	aa
AA	(1, 0, 0)	($\frac{1}{2}$, $\frac{1}{2}$, 0)	(0, 1, 0)
Aa	($\frac{1}{2}$, $\frac{1}{2}$, 0)	($\frac{1}{4}$, $\frac{1}{2}$, $\frac{1}{4}$)	(0, $\frac{1}{2}$, $\frac{1}{2}$)
aa	(0, 1, 0)	(0, $\frac{1}{2}$, $\frac{1}{2}$)	(0, 0, 1)

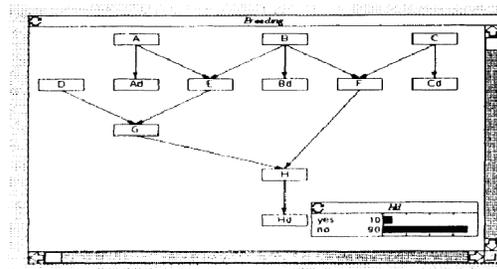
Tabel 2. Tabel penyebab probabilistik untuk hubungan orangtua anak, masing-masing spesifikasi kemungkinan masukan didistribusikan untuk genotype anak (AA,Aa,aa) diberikan dari tipe genotype orang tua P1 dan P2

Gambar 3. Tabel masukan untuk H

Tabel dibuat secara otomatis, dan tipe angka-angka dimasukkan. model domain selesai dan berbagai pilihan akan mengkompilasi sistem menjalankan waktu operasional. Walaupun langkah terakhir ini tampak sederhana, maka sebenarnya memegang kunci untuk keberhasilan metode. Apa yang terjadi adalah model state ruang difaktorkan agar perhitungan efisien.

Pemecahan masalah interaktif.

Pada node Hd, terlihat risiko sebesar 10% (Gambar 1) untuk menjadi korban dari penyakit. Ini tampaknya cukup rendah. Bartholomew, telah terjangkit penyakit. Untuk melihat apakah ini akan mempengaruhi persetujuan perkawinan, mari masukkan ke dalam kenyataan baru model kami.



Gambar 4. Tampilan kepercayaan di Hd

Dengan memasukkan fakta, kemudian memperbaharui kepercayaan dengan memilih memperkembang-tumbuhan permintaan kepercayaan Hd sekarang memiliki hasil risiko dari 36%, memaksa Florence dan Gregory mempertimbangkan kembali rencana mereka. Misalkan informasi tentang Bartholomew dituduh Bersalah atas kemalangan yang terjadi, tetapi sebagai gantinya Alison menunjukkan penyakit. Kita akan melakukan temuan di Bd (mengembalikan kepada a, priori kepercayaan untuk B) dan mengikuti prosedur yang

digambarkan di Ad. Sebagai alternatif, kita bisa memilih Inialisasi dan pengulangan seperti di atas. Di dalam kedua kasus, kita menemukan resiko untuk terjangkit menjadi 22%. Dalam posisi ini berbagai kemungkinan lain bisa diselidiki.

LATAR BELAKANG

Di dalam contoh Bagian 2, beberapa hal terjadi di balik peristiwa, tetapi pengguna itu tidak perlu cemas kecuali jika model mempunyai suatu kompleksitas bahwa memerlukan intervensi manual (lihat Bagian 5). Di dalam bagian ini, kita menguraikan secara singkat fitur yang paling penting di latar belakang. Di dalam diskusi beberapa pakar didalam artikel mereka, beberapa metoda dapat ditemukan.

Transformasi penyebab jaringan.

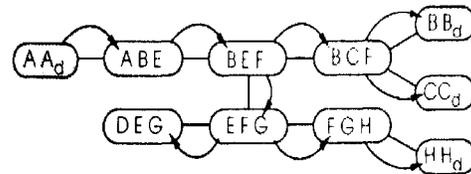
Inferensi di suatu jaringan HUGIN tidak selesai secara langsung di dalam jaringan. Sebagai gantinya, daerah yang diwakili oleh jaringan itu disekat ke dalam satu set subdomain disebut kepercayaan menyeluruh. Suatu kepercayaan menyeluruh terdiri dari satu set simpul-simpul, yang ditandai dan suatu yang dapat dipercaya. Karena pertimbangan jadinya segera membersihkan, gambar yang di set terbuka disebut pengelompokan-pengelompokan. Suatu tabel kepercayaan adalah membuat table normal untuk kemungkinan-kemungkinan yang menghubungkan ruang kepercayaan. Konstruksi dari suatu sistim kepercayaan menyeluruh, yaitu perpaduan model domain asli, terdiri dari langkah-langkah berikut:

- Bentuk grafik moral : Untuk masing-masing simpul semua di dalam jaringan, tambahkan sambungan antara orang tua .
- Grafik moral triangle: grafik adalah triangle jika semua siklus-siklus dari panjangnya >3 mempunyai suatu tali. Tambahkan sambungan dengan grafik moral sampai suatu grafik yang triangle diperoleh.
- Bentuk sistim dari belifuninerses: gambar terbuka yang di-set adalah persekongkolan-persekongkolan dari grafik yang triangulated (suatu persekongkolan adalah suatu yang maksimal set dari simpul-simpul, semua yang bersifat

memasangkan terhubung). tabel-tabel kepercayaan yang awal dihitung sebagai produk-produk yang sesuai ol tabel-tabel kementakan bersyarat

- Organisir sistim sebagai suatu pohon simpangan: Hubungan kepercayaan menyeluruh diperkenalkan.

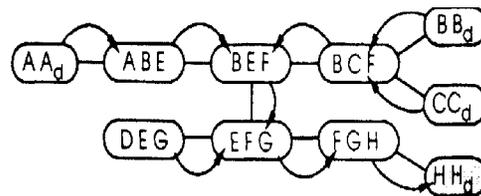
Semua langkah-langkah kecuali yang kedua bersifat deterministik: Ada saja satu grafik moral, dan set dari persekongkolan-persekongkolan dari suatu triangulated grafik adalah unik. Mungkin ada beberapa pohon-pohon simpangan, meskipun demikian, perbedaan-perbedaan itu tidak penting.



Gambar 5: Bukti didistribusikan dipanggil dari (4, Ad), dan bukti disebarakan kepada semua secara menyeluruh.

Pengembangan bukti

Suatu tabel kepercayaan adalah suatu penilaian yang tidak dinormalisir atas kemungkinan-kemungkinan yang menghubungkan suatu persekongkolan. Dimungkinkan V adalah suatu alam semesta kepercayaan dengan kepercayaan



Gambar 6 . Kumpulan aliran bukti

Di HUGIN, satu obyek mengorientasikan gaya yang digunakan: Masing-masing kepercayaan secara menyeluruh adalah satu data pemilikan obyek sendiri dan metoda-metoda sendiri. Mereka berkomunikasi hanya dengan tetangga-tetangga dalam simpangan pohon. Kalibrasi

adalah metoda perkebangbiakan lokal. Untuk perkebangbiakan global, bukti itu disebarkan kepada semua kepercayaan secara menyeluruh melalui simpangan pohon. Ada dua jalan arah bukti penyebaran. Suatu pemain depan perkebangbiakan yang dilaksanakan oleh operasi penyebaran bukti, dan suatu perkebangbiakan mundur dilaksanakan oleh operasi kumpulan Emdence.

Penyebaran bukti digunakan ketika bukti dari suatu kepercayaan secara menyeluruh harus menyebar ke seluruh sistim: Bukti disebarkan kepada tetangga-tetangga, dan metoda ini mereka sebut penyebaran Evulence (lihat Gambar 5). Pengumpulan bukti digunakan ketika bukti irom seluruh sistim harus disebarkan kepada kepercayaan secara menyeluruh

Metoda-metoda di HUGIN mempunyai keuntungan manipulasi-manipulasi grafik dilaksanakan kali ini dan terakhir; akhirnya adalah grafik triangle. Ketika pohon simpangan sudah dibangun, satu batas atas dan kebutuhan ruang untuk perkebangbiakan bukti dapat dengan mudah diperkirakan.

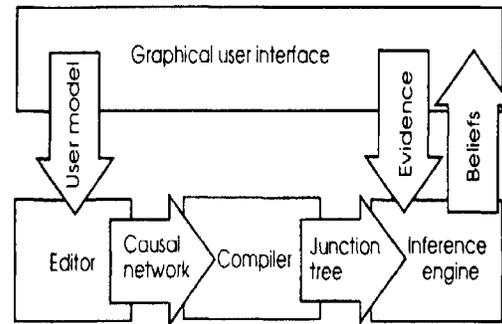
Secara umum, tugas dari inferensi di dalam jaringan probabilistik adalah menyebabkan NP-hard. Metoda di HUGIN buat nya tak perlu memecahkan satu masalah NP-hard setiap bukti waktu disebarkan. Sebagai gantinya masalah NP-hard dikerjakan di dalam tahap konstruksi, di mana banyak usaha dapat diabdikan untuk menemukan pemecahan optimal.

RINGKASAN SKEMATIS HUG I N

Tipe khas sesi hugin dibagi menjadi tiga bagian tugas:

1. Ciptaan dari domain model,
2. dari generasi runtime sistem
3. masalah yang sebenarnya.

Tugas ini harus dilakukan secara berurutan, tetapi dapat ditunda sesi pada titik tertentu, dan kemudian dilanjutkan. Shell MUCIN yang ditunjukkan di modul pada Gambar 7.



Gambar 7: Struktur Hugin

Struktur Hugin mengalir melalui sistem dari model domain untuk menjalankan sebuah sistem-waktu Antarmuka grafis, tidak semua modul pedoman diperlukan, memelihara dan menyediakan fasilitas akses untuk penciptaan, menyimpan, pemanggilan, dan penghapusan domain. Perbaikan modul menyediakan fasilitas untuk membuat dan memperbaiki penyebab jaringan probabilistik, misalnya, penciptaan dan penghapusan simpul dan penyebab link, penciptaan dan perbaikan dari tabel probabilitas bersyarat dll yang membuat penterjemah modul persimpangan pohon di mana seluruh Perhitungan dilakukan.

JARINGAN M U N I N

Jaringan besar MUNIN sedang dibangun untuk mendiagnosa otot dan penyakit syaraf. jaringan yang terdiri dari ratusan simpul. Penanganan jaringan ini menimbulkan masalah baru. Masalah utama keprihatinan ruang dan waktu. Otomatis triangulasi dicapai dalam Hugin oleh maksimum cardinality algoritma pencarian. Metode ini sudah cukup untuk masalah kecil, namun tidak memadai untuk jaringan besar karena memproduksi kepercayaan yang besar. Sebagai contoh, salah satu di jaringan MUNIN terdiri dari 57 node. Dalam jaringan ini membutuhkan tabel kepercayaan yang besar, hal ini berarti merupakan ruang perambatan.

KESIMPULAN

Sistem pakar shell model domain baru oleh penyebab probabilistik probabilistik. Prestasi utama adalah penanganan banyak dihubungkan oleh transformasi jaringan pada struktur tabel kepercayaan. Shell, yang telah membuktikan kemampuan untuk menangani model domain non-mudah seperti ukuran disederhanakan ukuran oleh aplikasi MUNIN. HUGIN sekarang alat yang baik untuk

pembangunan penelitian seperti sistem pakar khusus, berikut ide-ide dan fungsionalitas berada di bawah penyelidikan dan pengembangan:

- Deteksi yang bertentangan dengan bukti
- Penjelasan fasilitas
- Mencari temuan penting
- Perencanaan
- Belajar
- Ekstensi untuk variabel kontinyu

DAFTAR PUSTAKA

- S. K. Andersen, F. V. Jensen, and K. C. Olesen. The HUGIN core preliminary considerations on systems for fast manipulations of probabilities. In *Proceedings of Workshop on Inductive Reasoning: Managing Empirical Information in AI-Systems*, Risø National Laboratory, Roskilde, Denmark, April 1987.
- S. Andreassen, M. Woldbye, B. Falck, and S. K. Andersen. MUNIN—A Causal Probabilistic Network for Interpretation of Electromyographic Findings. In *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, pages 366-372, Milan, Italy, August 1987.
- S. Andreassen, F. V. Jensen, S. K. Andersen, B. Falck, A. Rosenfalck, and F. Jensen. MUNIN -An Expert EMG Assistant. In *Computer-Aided Surface or Needle Electromyography and Expert Systems in Diagnosis*, J. E. Desmedt (editor), Elsevier Science Publishers, Amsterdam, The Netherlands, 1989.
- G. F. Cooper. NESTOR: A Computer Based Medical Diagnostic Aid that integrates Causal and Probabilistic Knowledge. Technical Report HHP-84-48, Medical Computer Science Group, Stanford University, Stanford, California, 1984.
- G. F. Cooper. Probabilistic Inference Using Belief Networks is NP-Hard. Research Report KLS-87-27, Medical Computer Science Group, Stanford University, Stanford, California, 1987.
- [Dechter *et al*, 1988] R. Dechter, A. Dechter, and J. Pearl. Optimization in Constraint Networks. In *Proceedings of Conference on Influence Diagrams for Decision Analysis, Inference, and Prediction*, UCB, Berkeley, California, 1988.
- F. Gavril. Some AfV -complete problems on graphs. In *Proceedings of the Eleventh Conference on Information Sciences and Systems*, Johns Hopkins University, Baltimore, Maryland, 1977.
- E. J. Horvitz, J. S. Breese, and M. Henrion. Decision Theory in Expert Systems and Artificial Intelligence. To appear in *Journal of Approximate Reasoning, Special Issue on Uncertain Reasoning*, 1988.
- Howard and Matheson, 1981] R. A. Howard and J. E. Matheson. Influence Diagrams. In *Readings in Decision Analysis*, R. A. Howard and J. E. Matheson (editors), Strategic Decisions Group, Menlo Park, California, chapter 38, "763-771, 1981.
- F. V. Jensen. Junction Trees and Decomposable Hypergraphs. Judex Research Report, Aalborg, Denmark,
- F. V. Jensen. Discussion published in [Lauritzen and Spiegelhalter, 1988], 1988.
- F. V. Jensen, S. K. Andersen, U. Kjaerulff, and S. Andreassen. MUNIN: on the case for probabilities in medical expert systems—a practical exercise. In *Proceedings of the First Conference of Artificial Intelligence in Medicine*, J. Fox, M. Fieschi, and R. Engelbrecht (editors), pages 149–160, Springer-Verlag, Heidelberg, 1987.
- F. V. Jensen and U. Kjaerulff. Triangulation of graphs—Algorithms giving small total clique size, 1989. In preparation.
- F. V. Jensen, K. G. Olesen, and S. K. Andersen. An Algebra of Bayesian Belief Universes for Knowledge Based Systems. Research Report R.-88-25, Institute of Electronic Systems, Aalborg University, Denmark, 1988. To appear in *Networks*, 1989.
- J. I. Kim and J. Pearl. A computational model for causal and diagnostic reasoning in inference systems. In *Proceedings of the Joint Conference on Artificial Intelligence*, pages 190-193, Karlsruhe, West Germany, 1983.
- C. T. A. Kong. Multivariate Belief Functions and Graphical Models. Ph.D. Thesis, Harvard University, Research Report S-107, 1986.

- S. L. Lauritzen and D.J. Spiegelhalter. Local Computations with Probabilities on Graphical structures and Their Application to Expert Systems. *J. Royal Statistical Society B*, 50(2): 157-224, 1988.
- K. G. Olesen and S. K. Andersen. Discussion published in [Lauritzen and Spiegelhalter, 1988], 1988.
- F. V. Jensen, B. Falck, S. Andreassen, and S. K. Andersen.
- A MUNIN Network for the Median Nerve- A Case Study on Loops. *Applied Artificial Intelligence, Special Issue: Towards Causal AI Models in Practice*, 1989.
- J. Pearl. Fusion, Propagation, and Structuring in Belief Networks. *Artificial Intelligence*, 29:241-288, 1986.
- J. Pearl. A Constraint-Propagation Approach to Probabilistic Reasoning. In *Uncertainty in Artificial Intelligence*, L. N. Kanal and J. F. Lemmers (editors), pages 357-369, Elsevier Science Publishers, Amsterdam, The Netherlands, 1986.
- J. Pearl. Embracing Causality in Default Reasoning. *Artificial Intelligence*, 35:259-271, 1988.
- J. Pearl and A. Paz. Graphoids A graph-based logic for reasoning about relevance relations. In *Proceedings of the Seventh European Conference on Artificial Intelligence*, pages 55-61, Brighton, UK, 1986.
- R. I. Shachter. Evaluating Influence Diagrams. *Operational Research*, 34:871-882, 1986.
- R. D. Shachter and D. Heckerman. A Backwards View for Assessment. In *AI magazine*, 8(3):55-62, 1987.
- G. Shafer and P. Shenoy. Bayesian and Belief-Function Propagation. School of Business, Working Paper 121, University of Kansas, Lawrence, 1988.
- D. J. Spiegelhalter. Fast algorithms for probabilistic reasoning in influence diagrams with application in genetics and expert systems. In *Influence Diagrams*, R. Barlow *et al* (editors), Wiley, Chichester, 1989.
- R. E. Tarjan and M. Yannakakis. Simple Linear-time Algorithms to Test Chordality of Graphs, Test Acyclicity of Hypergraphs, and Selectively Reduce Acyclic Hypergraphs. *SIAM J. Computing*, 13:566-579, 1984.
- M. Yannakakis. Computing the minimal fill-in is N'P-complete. *SIAM J. Algebraic Discrete Methods*, 2:77-79, 1981.