

# Part III

## Multimedia Environments

- ❑ Videodisc
- ❑ the CD family and DVD
- ❑ CD-i
- ❑ DVI
- ❑ QuickTime

### Videodisc

- ❑ videodiscs come in many formats; however, LV (Laser Vision) is the most common;
- ❑ LV is read only; LV technology is a forerunner of audio CD
- ❑ LV production: a laser used to form small pits on a master disc's surface; a stamper is then created from the master and used to mold discs for distribution
- ❑ an LV disc is read by a low-powered laser beam which detects the recorded pits; unlike CDs, this binary pickup signal is converted immediately to analog within the player

### Videodisc

- ❑ LV discs are produced in 2 formats
  - ❑ CAV—constant angular velocity
  - ❑ CLV—constant linear velocity
- ❑ in each case the pits are arranged into a long spiral starting near the center of the disc
- ❑ on CAV discs a single frame extends for exactly one revolution; all frames start at the same angular position; as a result CAV discs offer flexible playback: freeze frame, play forward or reverse at various rates
- ❑ CLV have twice the capacity

### Videodisc

	CAV	CLV
capacity duration/side frames/side	30 min (NTSC) 36 min (PAL) 54 000 frames	60 min not frame addressable
access time (worst case)	3s	10s
drive revolution	1 800 rpm (NTSC) 1 500 rpm (PAL)	600 - 1 800 rpm (NTSC) 570 - 1 500 rpm (PAL)
quality	high	high

## Aspen Movie Map

- ❑ Innovatives Hypermediasystem
- ❑ Andrew Lippman MIT 1978-1980
- ❑ 2 Laserdisc Players (jeweils 1 Player für Nord/Süd bzw. West/Ost Fahrten)

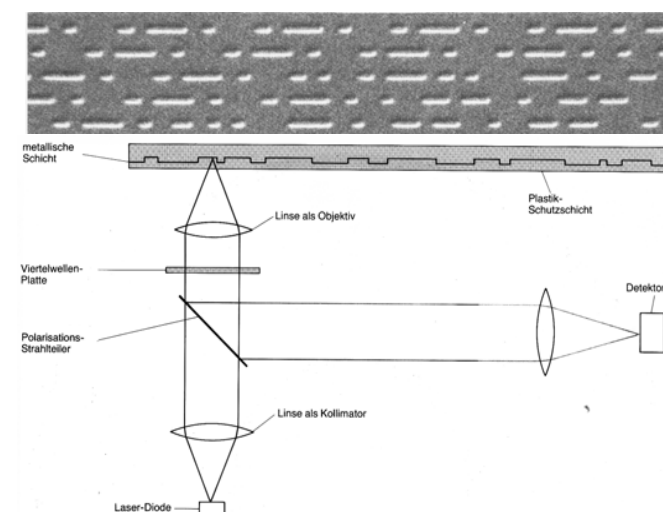
## Aspen Movie Map



## The CD Family

- ❑ CD audio developed by Philips and Sony, closely related to LV (LaserVision) videodiscs, differences:
  - ❑ true *digital medium* (LV provides only analog information)
  - ❑ CD uses *constant linear velocity* (CLV) format only
- ❑ data is encoded to assure robustness and longevity
- ❑ channel data — raw data sequence on disc
- ❑ frames — smallest unit of data independently decoded, consists of 588 channel bits, 200 bits after decoding (8 bits for subcode, 192 bits frame data)

## CD / DVD Technology



## CD Characteristics

disc diameter	12 cm
max. playing time	74 minutes
channel data rate	$4.32 \times 10^6$ bits/s
frame rate	7350 frames/s
channel bits per frame	588 bits
decoded frame size	200 bits
frame data size	192 bits
frame data rate	$1.41 \times 10^6$ bits/s
max. frame data capacity	$6.2 \times 10^9$ bits
maximum number of frames	$32.6 \times 10^6$
linear velocity	1.2 - 1.4 m/s
rotational velocity	3.5 - 8 rev/s

## The CD Family

- ❑ CD-DA — Compact Disc - Digital Audio
- ❑ CD+G, CD+MIDI
  - ❑ early CD formats storing text, simple graphics or MIDI in the 6 unused subcode bits (R-W), playback on any CD player
- ❑ CD-ROM
- ❑ Photo CD
  - ❑ Photo CD master, 35 mm, up to 100 images
  - ❑ Pro Photo CD Master, 35 mm, 70 mm, 120 mm, 6 - 100 im.
  - ❑ Photo CD Portfolio, audio, text, graphics and low resolution images, up to 800 images
  - ❑ Photo CD Catalog, up to 6,000 low-resolution images

## The CD Family

- ❑ CD-R, CD-WO — Compact Disc - Recordable / Write Once
- ❑ CD-ROM XA — Extended Architecture
  - ❑ open form of CD-i
- ❑ CD-i — Compact Disc - Interactive
- ❑ Video CD
- ❑ CD-V — Compact Disc - Video
  - ❑ hybrid digital audio / analog video

## CD-ROM

- ❑ difference between CD-ROM and CD-DA: additional layer of error detection and correction
- ❑ data block — disc is divided into consecutive data blocks (98 frames, 2352 bytes)
- ❑ 2 block formats
  - ❑ Mode 1 — 2048 bytes application data
  - ❑ Mode 2 — 2333 bytes application data

## CD-ROM Characteristics

block rate	75 blocks/s
frames per block	98
block data size	2352 bytes
application data per block	2048 bytes (Mode 1) 2336 bytes (Mode 2)
application data rate	150 Kbytes/s (Mode 1) 171 Kbytes/s (Mode 2)
max. application data capacity	645 - 735 Mbytes
max. number of blocks	330 000
average access time	1/2 s

## DVD

- ❑ generic term referring to one or more high density formats that have been proposed to support digital video data on a 5.25" disc (used to be HDCD):
  - ❑ MMCD (Multimedia Compact Disc)
  - ❑ SD (Super Density)
  - ❑ DVD (former Digital Versatile Disc or Digital Video Disc): consolidated format between SD & MMCD

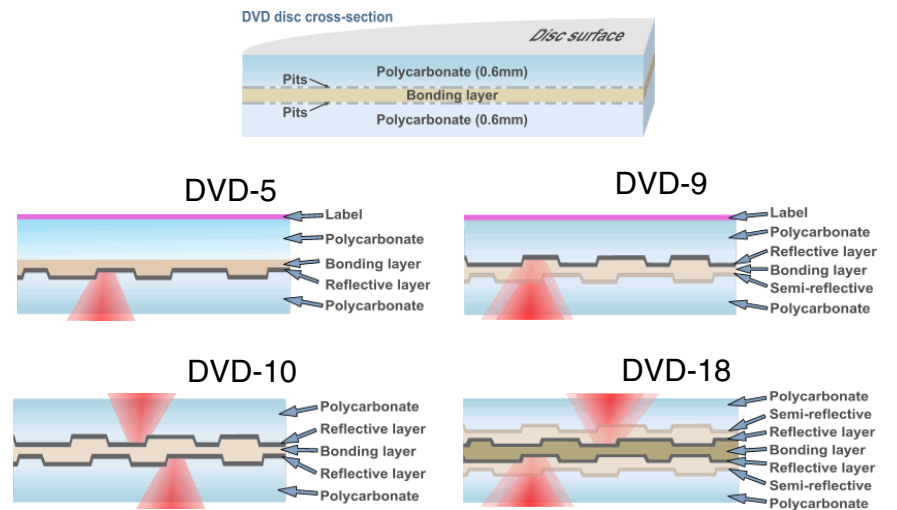
## DVD History 1/2

<b>1995</b>	Philips/Sony announce and demonstrate MMCD
	Toshiba and Warner announce and demonstrate SD
	Agreement on a single standard called DVD
<b>1996</b>	DVD-ROM and DVD-Video specifications version 1.0 published
	Digital copy protection scheme agreed
	First DVD-Video player sold in Tokyo (Nov)
<b>1997</b>	Launch in USA (August)
	DVD Consortium becomes DVD Forum, expands membership and holds first General DVD Forum Meeting with 120 members
<b>1998</b>	DVD-Video version 1.1 and DVD-ROM version 1.01 specifications released

## DVD History 2/2

<b>1998</b>	DVD Forum adopts DVD-RW as another re-writable format
	7 new members of DVD Forum Steering Committee making 17 in all
	DVD-Forum publishes DVD-Audio specification version 0.9
	Full launch of DVD in Europe. 1m DVD-Videoplayers sold in USA
	4.7 GB DVD-R and DVD-RAM version 1.9 specifications released
<b>1999</b>	DVD-Audio specification ver 1.0 released followed later by version 1.1
<b>2000</b>	copy protection for DVD-Audio agreed
	DVD-Audio players launched in USA (July)
	First DVD-Audio discs in USA (Nov)
<b>2001</b>	DVD-Audio players & discs available in Europe and elsewhere
	DVD Video Recorders launched in Europe etc.

DVD Technology



DVD Technology

Name	Capacity (GB)	Layers	Sides	Comments
DVD-5	4.7	1	1	Read from one side only
DVD-9	8.54	2	1	Read from one side only
DVD-10	9.4	1	2	Read from both sides
DVD-18	17.08	2	2	Read from both sides
DVD-R	4.7/9.4	1	1 or 2	Recordable DVD
DVD-RAM	4.7/9.4	1	1 or 2	Re-writable DVD
DVD-RW	4.7	1	1 or 2	Re-writable DVD

DVD vs. CD Characteristics

	CD	DVD	
diameter (cm)	12.0	12.0	
layers	1	single	dual
substrate thickness (mm)	1.2	0.6	
substrates/ side	1	1	2
Track pitch (microns)	1.6	0.74	
Min pit length (microns)	0.83	0.4	0.44
Scan velocity	1.3	3.49	3.84
Wavelength (nm)	780	635/650	

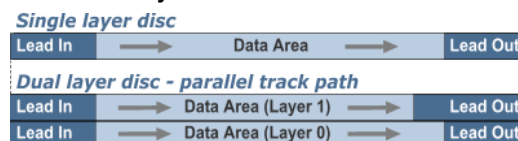
Format DVD-ROM

- ❑ DVD-ROM — for Computer applications (games, multimedia, encyclopedias, large sizes are needed...)
Backward read compatibility with existing CD-ROMs
High performance for both sequential and non-sequential data types
- ❑ DVD-ROM ↔ CD-ROM

	DVD-5	DVD-9	CD-ROM
Capacity (GB)	4.7	8.5	0.7
Drive speed	10x		48x
Data Transfer rate (Mb/s)	10 and more		2.8 to 6
Linear velocity (m/s)	21	23	up to 52

## Format DVD-ROM / DVD-R

### ❑ DVD-ROM Disc layout



### ❑ DVD-R — Write-Once;

- ❑ similar technology to CD-Rs
- ❑ allows incremental writing
- ❑ compatible with DVD-ROM, DVD-Video and DVD-Audio
- ❑ only one layer/side possible

## Format Rewritable DVD

- ❑ DVD-RAM —use CAV, data rate = 22.16 Mb/s: for high performance computer data storage and archive applications. Over 100,000 recording cycles; 30-year life or more; needs a cartridge;
- ❑ DVD-RW—use CLV: for videotape replacement, video authoring and desktop PC-backup; recordable over 1,000 times
- ❑ DVD+RW—Choice of CAV for high performance data storage applications and CLV for video recording; Compatible with DVD-ROM drives; Disadvantage: not supported by the DVD Forum.

## Format DVD-Video 1

### ❑ 4 Types of Data:

#### ❑ Video — 1 stream; encoding MPEG-1 & MPEG-2

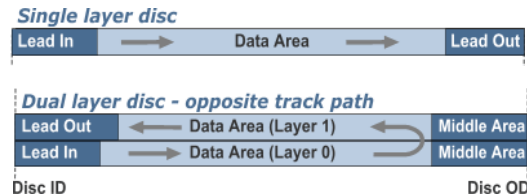
	MPEG-2	MPEG-1
<b>NTSC</b> resolution	720/704 x 480 352 x 480/240	352 x 480 351 x 240
<b>PAL/SECAM</b> resolution	720/704 x 576 352 x 576/288	352 x 576 352 x 288
<b>VBR or CBR</b>	VBR or CBR	constant bit rate
<b>PAL/SECAM</b> frame rate	25 fps	
<b>NTSC</b> frame rate (on disc)	24 or 29.97 fps	

## Format DVD-Video 2

- ❑ Audio — up to 8 streams; include Dolby Digital, MPEG-1 & MPEG-2, LPCM and DTS.
- ❑ Subpictures — graphics, overlaid: used as Subtitles (up to 32 languages), Menues
- ❑ Navigation — Program chains, Player controls, Menues
- ❑ Bit rate —
  - ❑ Maximum: 9.8 Mb/s (video+ audio+subpictures)
  - ❑ Average: for high result should be close to 4 Mb/s (depends on the number of audio streams and the encoding used)

## Format DVD-Video 3

- ❑ Playing Time— **DVD-5**: 133 minutes **DVD-9**: 240 minutes.  
(high quality MPEG-2, 3 surround sound audio channels, four subtitle channels, average data rate= 4,69 Mb/s).
- ❑ DVD-Video Disc layout:



## DVD-Audio

- ❑ sampling frequencies—48kHz, 96kHz, 192kHz, as well as 44.1kHz, 88.2kHz, 176.4 kHz
- ❑ sample size—16bit, 20bit and 24bit
- ❑ number of channels—up to six channels are available for multi-channel recording, with a transfer rate of 9.6 Mbps maximum
- ❑ compatibility with the DVD-Video and DVD-ROM formats
- ❑ recording options—2-channel stereo (>74 min on a single-sided/ single-layer disc at 192kHz/24bit); multi-channel mode (96kHz/24bit, 6-channel, >74 min)

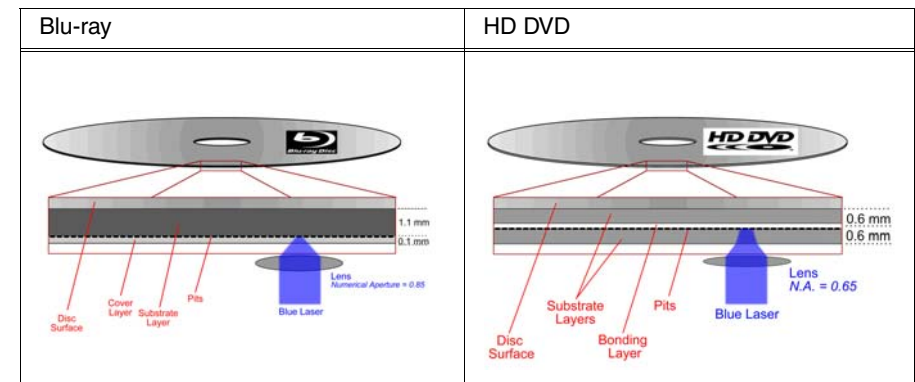
## Blu-ray vs. HD DVD



<http://www.blu-raydisc.com/>

<http://www.thelookandsoundofperfect.com/>

## Physikalischer Aufbau



## Disc Struktur

	Blu-ray	HD DVD
Schichten	<b>1 Substrate Layer</b> <b>Cover Layer</b> nur 0,1µm! Polymer „Durabis“: Von TDK Corporation neu entwickeltes, kratzbeständig- es Material	<b>2 Substrate Layers</b> Anordnung, Abmessungen und Materialien sind gleich wie bei herkömmlicher DVD
Pit - Minimale Länge	<b>0,149 µm</b>	<b>0,204 µm</b>
Spurenbreite	<b>0.32 µm</b>	<b>0,40 µm</b>
Abstand Lasereintritt/Pit	<b>0,1 µm</b>	<b>0,6 µm</b>

## Strategien zur Kapazitätsmaximierung

	Blu-ray	HD DVD	DVD
Laser	<b>Blue Laser</b> (405 nm)	<b>Blue Laser</b> (405 nm)	<b>Red Laser</b> (650 nm)
NA*	<b>0.85</b>	<b>0.65</b>	<b>0.60</b>
Zusätzliche Strategie	<b>Dünnere Cover Layer</b> vermindert die durch Lichtbeugung eingeführte Strahlen- unschärfe	-	-

\*NA - Numerische Apertur

- ☐ Kenngröße für die Bündelung eines optischen Systems
- ☐ Je höher NA, desto besser ist Lichtstrahl gebündelt

## Formate - CODECs

	Blu-ray	HD DVD
Codec Video	<b>MPEG-2</b> <b>H.264/MPEG-4 AVC</b> <b>SMPTE VC-1</b> (Populärste Implementierung: wmv9)	<b>MPEG-2</b> <b>H.264/MPEG-4 AVC</b> <b>SMPTE VC-1</b>
Codec Audio <i>verlustbehaftet</i>	<b>Dolby Digital AC-3</b> (Maximum 640kbps) <b>DTS</b> <b>Dolby Digital Plus</b> (optional, Maximum 1,7Mbps)	<b>Dolby Digital AC-3</b> (Maxi- mum 448kbps) <b>DTS</b> <b>Dolby Digital Plus</b> (Maximum 3Mbps)
Codec Audio <i>verlustfrei</i>	<b>Lineares PCM</b> (bis zu 7.1 channel) <b>Dolby TrueHD</b> <b>DTS-HD</b>	<b>Lineares PCM</b> (bis zu 7.1 channel) <b>Dolby TrueHD</b> <b>DTS-HD</b> (optional)

## Datentransfer

	Blu-ray	HD DVD	DVD
Audio+Video, maximal	<b>54 Mbps</b>	<b>36,55 Mbps</b>	<b>11,8 Mbps</b>
Nur Video, maximal	<b>40 Mbps</b>	-	-

## Digital Rights Management

Blu-ray	HD DVD	DVD
<b>Region Code</b> 3 Regionen	<b>z.Z. kein Region Code</b> Auf Druck der Hollywood Studios wird Code entwickelt	<b>Region Code</b> 8 Regionen
<b>AACS</b> <i>Advanced Access Content System</i>	<b>AACS</b> 128-bit Nachfolger von <b>CSS</b> Kontrolliert Kopiervorgänge (auch nicht digitales Kopieren!) Kontrolliert Übertragung auf andere Geräte	<b>CSS</b> (40-bit) <i>Content Scrambling System</i> Wurde geknackt, seither kein Schutz mehr vor Raubkopien
<b>BD+</b> <i>Dynamischer Schlüssel</i> Sollte Schlüssel geknackt werden, BD Hersteller können Schlüssel upgraden, nachfolgend erzeugte BDs wieder kopiergeschützt		

## Speicherkapazität und Discgrößen

Diskgröße		Blu-ray	HD DVD
ø 12cm	Single Layer	<b>25 GB</b>	<b>15 GB</b>
	Double Layer	<b>50 GB</b>	<b>30 GB</b>
	Multilayer >2	<b>X</b> (im Labor & Layers 200GB)	<b>X</b> (Im Labor 3 Layers)
	Double Side Version	<b>X</b> (Möglich bei 1, 2, 4 Layerversion)	<b>X</b>
ø 8cm	Single Layer	<b>7,8 MB</b>	<b>4,7 GB</b>
	Double Layer	<b>15,6 GB</b>	<b>9,4 GB</b>
	Multilayer >2	-	-
	Double Side Version	<b>X</b>	<b>X</b>

## Beschreibbare Discs

	Blu-ray	HD DVD
Bezeichnung	<b>BD-R</b> (1x beschreibbar) <b>BD-RE</b> (wiederbeschreibbar)	<b>HD-DVD-R</b> (1x beschreibbar) <b>HD-DVD-RW</b> (wiederbeschreibbar)
Kapazität	<b>25 GB</b> (Single Layer)	<b>15 GB</b> (Single Layer)
Datenrate bei Schreibvorgang	<b>36 Mbps</b> (Drivespeed 1x) <b>72 Mbps</b> (Drivespeed 2x)	<b>36,55 Mbps</b> (Drivespeed 1x) <b>73 Mbps</b> (Drivespeed 2x)
Schreibdauer	<b>95 min.</b> (Drivespeed 1x) <b>47 min.</b> (Drivespeed 2x)	<b>55 min.</b> (Drivespeed 1x) <b>28 min.</b> (Drivespeed 2x)
Preis	Für HD DVD-R wird ein niedriger Preis als für die BD-R erwartet	

Quelle: DVDForum Homepage, Wikipedia

## Quellen

### Offizielle Homepages

- ☐ Blu-ray Homepage - <http://www.blu-raydisc.com/>
- ☐ HD DVD Homepage - <http://www.thelookandsoundofperfect.com/>
- ☐ DVD Forum - <http://www.dvdforum.org/forum.shtml>

## Compact Disc - Interactive (CD-i)

- ❑ first mm platform, early 90ies
- ❑ complete system: physical and logical organization of MM information on CD, HW interfaces for CD-i players, SW interfaces for CD-i programs.
- ❑ requires a video monitor and audio presentation equipment

## CD-i Media Types

- ❑ audio — CD-DA, Level A, Level B, Level C
- ❑ image — RGB 5:5:5, DYUV, CLUT (256, 128, 16 entries), RL (run-length coding, 128 or 8 entries)
- ❑ video — MPEG-1 encoded video
- ❑ text and graphics (interpretation of this data is left to applications)

## Digital Video Interactive (DVI)

- ❑ DVI technology — several components: formats and codecs for audio and video, hw and sw allowing applications to capture and playback these media
- ❑ DVI system — integration of these components into a particular host hw and sw platform
- ❑ developed at David Sarnoff Research Center, then acquired by INTEL, DVI chip and board sets (ActionMedia boards)

## DVI Media Types

- ❑ video — real-time video (RTV), production-level video (PLV)
- ❑ audio — a number of audio types, eg. FM (5 hours stereo), “mid-range” (20 hours mono), near AM mono (40 hours)
- ❑ image — number of formats, eg. RGB, YUV, monochrome, color mapped images and images with alpha channel (DVI hw offer lossless and lossy compression)
- ❑ text and graphics — rendering relies on DVI hw

## QuickTime

### General

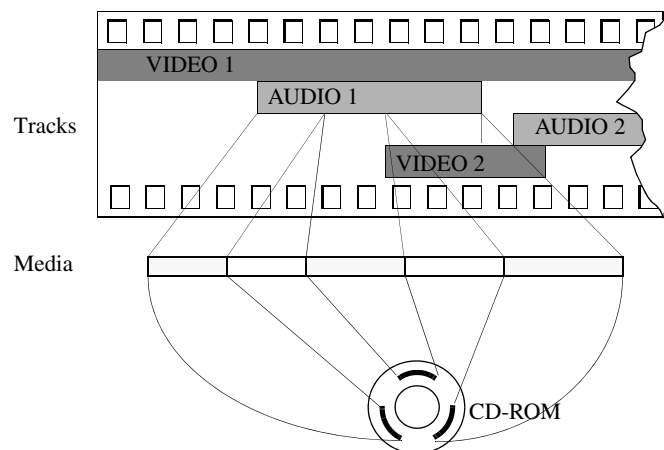
- ❑ extension to Macintosh OS supporting time-based data
- ❑ movie — multimedia type composed of tracks (audio or video)
- ❑ open environment — allowing the addition of new media types, compression components and hw devices
- ❑ component — piece of sw with an interface of a particular form (component types)

## QuickTime Media Types

- ❑ video and images — based on Apple's PICT format, several compression algorithms
- ❑ audio
  - ❑ range of sample rates (max. 65kHz)
  - ❑ mono or stereo (interleaved)
  - ❑ PCM with sample size of 8, 16 or 32 bits
- ❑ text and graphics — can be included in PICT data or movies, basically ignored by QuickTime

## QuickTime Media Organization

### Movie



## QuickTime Media Organization

### Time in QT

- ❑ timing is explicit
- ❑ values represented by unsigned integers
- ❑ time scale—units per second
- ❑ duration—maximum time value
- ❑ timebase—playback rate
- ❑ time coordinate system (TCS, scale and duration)

## QuickTime Media Organization

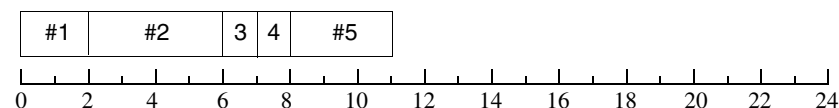
Conceptional level:

- ☐ data entity—represents the actual storage used for a/v data.
- ☐ media entity—temporal sequence measured in *media time*, each entity has a media type (either a or v), elements reference storage regions.
- ☐ track entity—reordering of media entity, each entity is contained in one movie entity, time measured in movie TCS.
- ☐ movie entity—group of track entities, specifies time scale and duration, tracks are offset within the movie.

## QT Media Organization (Examples)

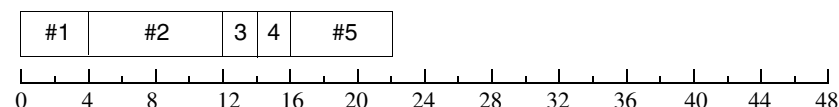
media entity

$vMedia_{25} = (<\#1,0,2>, <\#2,2,4>, <\#3,6,1>, <\#4,7,1>, <\#5,8,3>)$



track entity

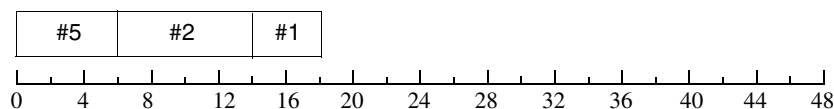
$vTrack_{50} = (<\#1,0,4>, <\#2,4,8>, <\#3,12,2>, <\#4,14,2>, <\#5,16,6>)$



## QT Media Organization (Examples)

edited track entity

$edited-vTrack_{50} = (<\#5,0,6>, <\#2,6,8>, <\#1,14,4>)$

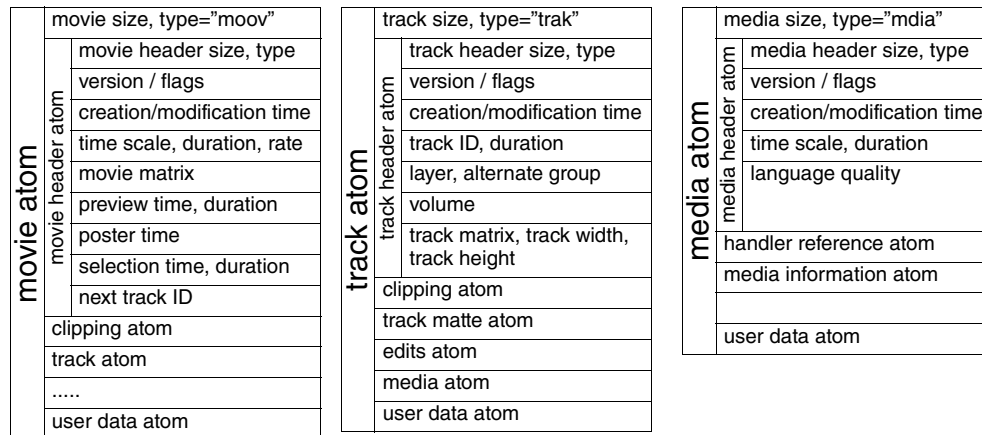


## QT Media Organization

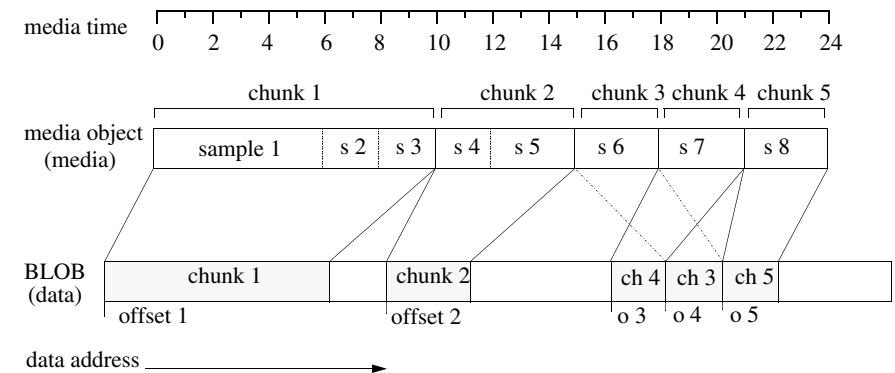
Physical organization

- ☐ atom—basic storage unit, consists of
  - ☐ size (in bytes)
  - ☐ four character code (identifies type)
  - ☐ content section (fields or other atoms)

## QuickTime Media Organization



## QuickTime Media Mapping (Example)



The mapping of media time to data address

## QuickTime Media Mapping (Example)

		Time-to-sample			
		sample number	sample span	sample start time	sample duration
Ch 1	sample 1	1	1	0	6
	sample 2	2	3	6	2
	sample 3	5	4	12	3
Ch 2	sample 4				
	sample 5				
Ch 3	sample 6				
Ch 4	sample 7				
Ch 5	sample 8				

Sample-to-chunk				
chunk number	chunk span	start sample in chunk	samples /chunk	encoding
1	1	1	3	E1
2	1	4	2	E1
3	2	6	1	E1
5	1	8	1	E2

## Part IV Multimodal Information Retrieval

- ❑ Teil 1: Einführung + Grundlagen
- ❑ Teil 2: Bildretrieval
- ❑ Teil 3: Videoretrieval
- ❑ Teil 4: Audioretrieval
- ❑ Teil 5: Distanz- und Ähnlichkeitsmaße

## Teil IV.1 MMIR Grundlagen

### Definition Information Retrieval

**Information Retrieval (IR)** (Informationswiedergewinnung, gelegentlich Informationsbeschaffung)

betrachtet Informationssysteme in Bezug auf ihre Rolle im Prozess des Wissenstransfers vom menschlichen Wissensproduzenten zum Informations-Nachfragenden.

[GI Fachgruppe IR]

beschäftigt sich mit computergestützter, inhaltsorientierter und unscharfer Suche in unstrukturierten Datenmengen

<http://de.wikipedia.org/wiki/Information-Retrieval>

### Definition Multimodal

- ❑ **Multi** [lat. multus] viel-, vielfach-, Viel-
- ❑ **Modal** [lat.: Modus] die Art und Weise bezeichnend (mit der etwas geschieht oder gedacht wird)

#### Multimodal Information Retrieval

oft synonym mit

**(Inhaltsbasiertes) Multimedia Information Retrieval**

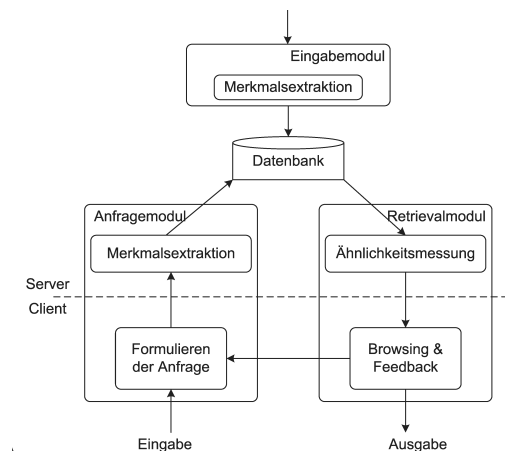
### Grundprinzip

- ❑ aus einem Medienobjekt oder Multimediaobjekt werden **inhaltsbasierte Merkmale** extrahiert (feature extraction)
- ❑ diese Merkmale werden in einem **Merkmalsvektor** zusammengefasst
- ❑ Merkmalsvektoren werden gespeichert
- ❑ die Merkmalsextraktion erfolgt für alle Objekte einer Kollektion
- ❑ Anfrage:  
welche zu einem Anfrage-Medienobjekt ähnlichen Objekte enthält die Kollektion?

## Grundprinzip (2)

- ❑ Vergleich von Merkmalsvektoren (Anfragevektor vs. Kollektionsvektoren), Ähnlichkeitssuche
- ❑ Reihung (**Ranking**) der Ergebnisse
- ❑ Iteration durch Modifikation der Anfrage oder **Relevanz-Feedback**

## Grundprinzip (3)



## IV.1.1 Retrieval Modelle

## Überblick

- ❑ IR-Systeme lassen sich nach den zugrunde liegenden IR-Modellen klassifizieren.
- ❑ IR-Modell legt die Art der Realisierung der folgenden Komponenten eines IR-Systems fest:
  - ❑ interne Dokumentdarstellung,
  - ❑ Anfrageformulierung und interne Anfragedarstellung,
  - ❑ Vergleichsfunktion zwischen jeweils zwei Dokumenten beziehungsweise zwischen Anfrage und einem Dokument.

## Überblick

- ❑ Modelle
  - ❑ boolesches Modell: Dokumente werden als Mengen von Indextermen repräsentiert. Suche über Terme in Termmengen; boolescher Junktoren
  - ❑ Fuzzy-Modell: Erweiterung von booleschem Modell
  - ❑ Vektorraummodell

## Boolesches Modell

- ❑ Konzepte der Mengentheorie und der booleschen Algebra
- ❑ klare Semantik, sehr einfaches Modell
- ❑ Gewicht eines Terms bezogen auf ein Text-Dokument binär
- ❑ Jedes Dokument wird intern durch die Menge von Indextermen mit Gewicht „1“ repräsentiert
- ❑ in Anfrage werden Terme angegeben, die durch boolesche Junktoren, also durch »and«, »or« und »not«, kombiniert werden

## Boolesches Modell

- ❑ Innerhalb der **Vergleichsfunktion** werden die durch die Anfrage spezifizierten Anfrageterme in den jeweiligen Dokumenten auf Enthaltensein getestet.
- ❑ Ergebnis eines Termtests ist ein boolescher Wert
- ❑ boolesche Junktoren zur Kombination einsetzbar

## Beispiel

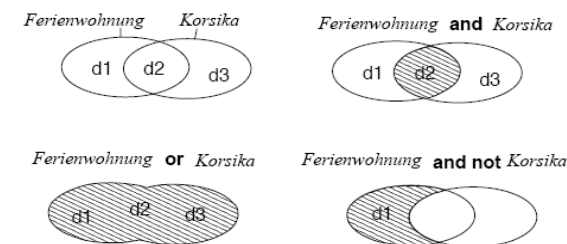
- ❑ Menge von Text-Dokumenten, die Beschreibungen zu Urlaubsorten enthalten:
  - ❑ Indexvokabular = {Korsika, Sardinien, Strand, Ferienwohnung, Gebirge}
- ❑ drei Dokumente
  - ❑ Dokument d1 : {Sardinien, Strand, Ferienwohnung}
  - ❑ Dokument d2 : {Korsika, Strand, Ferienwohnung}
  - ❑ Dokument d3 : {Korsika, Gebirge}

## Beispiel - 2

### ❑ Anfragen und Ergebnisse:

- ❑ Korsika liefert {d2, d3}
- ❑ Ferienwohnung liefert {d1, d2}
- ❑ Ferienwohnung *and* Korsika liefert {d2}
- ❑ Ferienwohnung *or* Korsika liefert {d1, d2, d3}
- ❑ Ferienwohnung *and not* Korsika liefert {d1}

## Beispiel - 3



- ❑ Problem bei der Verwendung des not-Junktors:  
„but“ statt „and not“

## Normalformen

- ❑ Für die Berechnung des Anfrageergebnisses wird die Anfrage normalisiert:
- ❑ disjunktive (DNF) oder konjunktive (KNF) Normalform:
  - ❑ Ferienwohnung *and* ((Sardinien *and* Strand) *or* Korsika)
  - ❑ (Ferienwohnung *and* Sardinien *and* Strand) *or* (Ferienwohnung *and* Korsika) (DNF)
  - ❑ Ferienwohnung *and* (Sardinien *or* Korsika) *and* (Strand *or* Korsika)

## Nachteile

- ❑ **exaktes Modell**: aufgrund binärer Gewichte keine Ähnlichkeitssuche
- ❑ **Größe des Ergebnisses**: oft zu viele Dokumente oder keine
- ❑ **boolesche Junktoren**: Schwierigkeiten vieler Anwender

## Milderung

- ❑ **exaktes Modell:** Umwandlung von Konjunktionen in Disjunktionen; Stufen der Relevanz
- ❑ **Größe des Ergebnisses:** einige Systeme haben zweistufiges Suchverfahren (faceted query)
  - ❑ Anfrage formuliert und verfeinert, ohne jedoch das Ergebnis anzuzeigen.
  - ❑ vollständiges Ergebnis
- ❑ **boolesche Junktoren:** *all* und *any* statt *and* und *or*

## Fuzzy-Modell

- ❑ Erweiterung des booleschen Modells
- ❑ Anfragen mit Hilfe von booleschen Junktoren
- ❑ **Fuzzy-Theorie** zur Milderung der zu scharfen Enthaltenseinsbedingung von Termen in Dokumenten
- ❑ graduellen Zugehörigkeit von Dokumenten zu Termen

## Fuzzy-Menge

Definition:

Eine **Fuzzy-Menge**  $A = \{ \langle u; m_A(u) \rangle \}$  über einem Universum  $U$  ist durch eine **Zugehörigkeitsfunktion**  $m_A : U \rightarrow [0, 1]$  charakterisiert, welche jedem Element  $u$  des Universums  $U$  einen Wert  $m_A(u)$  aus dem Intervall  $[0, 1]$  zuordnet.

- ❑ Menge aller gespeicherten Dokumente: Universum;  
Term: Fuzzy-Menge
- ❑ Ein Fuzzy-Wert  $\mu_t(d1)$  des Dokuments  $d1$  bezüglich des Terms  $t$  drückt aus, wie stark der Term das Dokument charakterisiert.

## Beispiel

- ❑ Die Fuzzy-Mengen **Korsika** bzw. **Strand** drücken die folgenden Zugehörigkeiten dreier Dokumente zum Term »Korsika« bzw. »Strand« aus:

$$\begin{aligned} \text{Korsika} &= \{ \langle d1; 0,1 \rangle, \langle d2; 0,6 \rangle, \langle d3; 1 \rangle \} \\ \text{Strand} &= \{ \langle d1; 0,3 \rangle, \langle d2; 0,2 \rangle, \langle d3; 0,8 \rangle \} \end{aligned}$$

- ❑ Zugehörigkeitsfunktionen

$\mu$	$d1$	$d2$	$d3$
$\mu_{\text{Korsika}}$	0,1	0,6	1
$\mu_{\text{Strand}}$	0,3	0,2	0,8

## Mengenoperationen

$$\begin{aligned}\mu_{A \cap B}(u) &= \min(\mu_A(u), \mu_B(u)) \\ \mu_{A \cup B}(u) &= \max(\mu_A(u), \mu_B(u)) \\ \mu_{\bar{A}}(u) &= 1 - \mu_A(u)\end{aligned}$$

- ☐ Anfrage analog zum booleschen Modell
- ☐ jeder Suchterm erzeugt Fuzzy-Menge
- ☐ Anfrage in disjunktive Normalform übergeführt
- ☐ Fuzzy-Mengenoperationen auf Fuzzy-Mengen ausgeführt
- ☐ Ergebnis entsprechend den Zugehörigkeitswerten sortiert

## Beispiel

Anfrage	$\mu$	d1	d2	d3
	$\mu_{\text{Korsika}}$	0,1	0,6	1
	$\mu_{\text{Strand}}$	0,3	0,2	0,8
1	$\mu_{\text{Korsika} \cap \text{Strand}}$	0,1	0,2	0,8
2	$\mu_{\text{Korsika} \cup \text{Strand}}$	0,3	0,6	1
3	$\mu_{\overline{\text{Korsika}}}$	0,9	0,4	0

## Zugehörigkeitswerte

- ☐ viele Ansätze für die Berechnung
- ☐ z. B. Term-zu-Term-Korrelationsmatrix

$$c_{i,j} = \frac{n_{i,j}}{n_i + n_j - n_{i,j}}$$

- ☐  $n_{i,j}$  Anzahl der Dokumente, die die Terme  $t_i$  und  $t_j$  gemeinsam enthalten
- ☐  $n_i$  und  $n_j$  jeweils die Anzahl der Dokumente, die den entsprechenden Term enthalten

## Beispiel

	$t_{\text{Sardinien}}$	$t_{\text{Strand}}$	$t_{\text{Ferienw.}}$	$t_{\text{Korsika}}$	$t_{\text{Gebirge}}$
$t_{\text{Sardinien}}$	1	$\frac{1}{2}$	$\frac{1}{2}$	0	0
$t_{\text{Strand}}$	$\frac{1}{2}$	1	1	$\frac{1}{3}$	0
$t_{\text{Ferienw.}}$	$\frac{1}{2}$	1	1	$\frac{1}{3}$	0
$t_{\text{Korsika}}$	0	$\frac{1}{3}$	$\frac{1}{3}$	1	$\frac{1}{2}$
$t_{\text{Gebirge}}$	0	0	0	$\frac{1}{2}$	1

## Dokumentzugehörigkeitswert

- ❑ Zugehörigkeitswert eines Dokumentes  $d_j$  zu einem Term  $t_i$ :
  - ❑ Aggregation aller Korrelationswerte zu allen im Dokument auftretenden Termen durch die Funktion

$$\mu_{t_i}(d_j) = 1 - \prod_{t_k \in d_j} (1 - c_{i,k})$$

## Vektorraummodell

- ❑ sehr weit verbreitetes Retrieval-Modell
- ❑ Dokumente werden als Vektoren eines Vektorraums aufgefasst
- ❑ Überführung des Retrieval-Problems in das Gebiet der Linearen Algebra
- ❑ kann überall dort eingesetzt werden, wo Medienobjekte durch eine feste Anzahl numerischer Merkmalswerte dargestellt werden können und sich Ähnlichkeit auf dieser Basis berechnet läßt

## Vektorraummodell

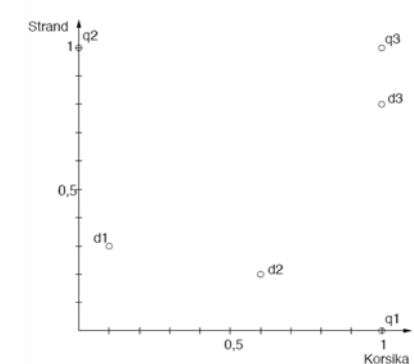
- ❑ unterstützt im Gegensatz zum booleschen Modell das Konzept der Ähnlichkeit
- ❑ Die Ähnlichkeit wird zwischen zwei Vektoren berechnet
- ❑ Anfrage wird durch einen Vektor repräsentiert
- ❑ viele Möglichkeiten zur Berechnung der Ähnlichkeit (z. B. Cosinusmaß, siehe 5.2)
- ❑ auch Ähnlichkeit über Distanzfunktionen (siehe 5.1)

## Beispiel

- ❑ drei Textdokumente mit 2 Termen; drei Anfragen

Dimension	d1	d2	d3
Korsika	0,1	0,6	1
Strand	0,3	0,2	0,8

Dimension	q1	q2	q3
Korsika	1	0	1
Strand	0	1	1



Multimedia

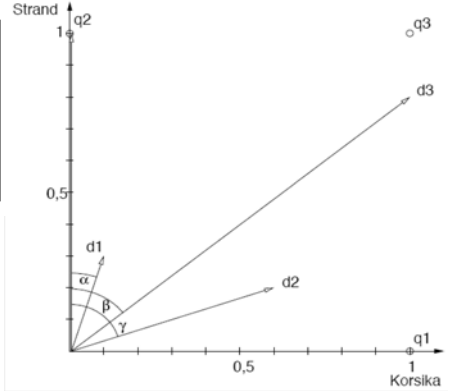
IV.1.1 Retrieval Modelle

Vektorraummodell

Beispiel - 2

□ Ähnlichkeitsmaß — Cosinusmaß

<i>sim<sub>cos</sub></i>	<i>d1</i>	<i>d2</i>	<i>d3</i>
<i>q1</i>	0,3162	0,9487	0,7809
<i>q2</i>	0,9487	0,3162	0,6247
<i>q3</i>	0,8944	0,8944	0,9939



367

Christian Breiteneder

Multimedia

IV.1.1 Retrieval Modelle

Vektorraummodell

Beispiel - 3

□ Distanzfunktion — Euklidische Distanz

<i>dissim<sub>L2</sub></i>	<i>d1</i>	<i>d2</i>	<i>d3</i>
<i>q1</i>	0,9487	0,4472	0,8
<i>q2</i>	0,7071	1	1,0198
<i>q3</i>	1,1402	0,8944	0,2

368

Christian Breiteneder

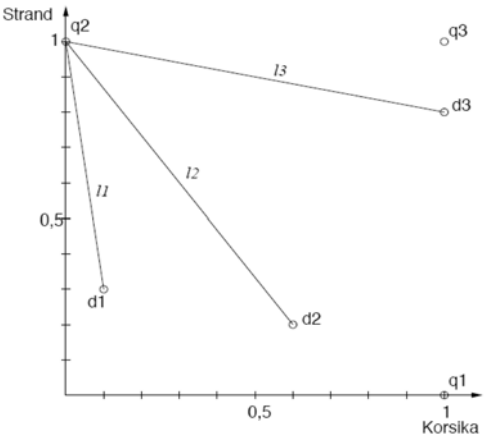
Multimedia

IV.1.1 Retrieval Modelle

Vektorraummodell

Beispiel - 3

□ Distanzfunktion — Euklidische Distanz



369

Christian Breiteneder

IV.1.2 Relevance Feedback

## Motivation

- ☐ IR-Prozess enthält häufig Iterationen
- ☐ Anfrageverfeinerung durch Nutzerinteraktion
- ☐ erste Ergebnisliste für den Suchenden oft nicht zufriedenstellend

## Anfragemodifikation — Gründe

- ☐ vage Vorstellung über Suchergebnis
- ☐ schlechte Anfrageformulierung
- ☐ unbekannte Datenkollektion
- ☐ keine relevanten Dokumente verfügbar

## Arten der Nutzerreaktion

- ☐ Browsing — sequentielle Suche, meist wenig sinnvoll
- ☐ manuelle Anfragemodifikation
- ☐ Relevance Feedback
  - ☐ Bewertung der Dokumente der Ergebnisliste entsprechend ihrer Relevanz zur Anfrage durch Nutzer
  - ☐ das IR-System modifiziert die Anfrage

## Beispiel

- ☐ ideale Anfrage  $q$  (dem Nutzer nicht bekannt);  $q_0, q_1, q_2$   
Anfrage-Iterationen mit Bewertungen durch Nutzer

Anfrage	Ergebnisdokumente			
	1	2	3	...
$q$	$d_0$	$d_1$	$d_2$	...
$q_0$	$d_4$	$d_1 (+)$	$d_5 (-)$	...
$q_1$	$d_1 (+)$	$d_3 (+)$	$d_4 (-)$	...
$q_2$	$d_3$	$d_1$	$d_0$	...

## Bewertung von Dokumenten

Berücksichtigung folgender Aspekte:

- ☐ Anzahl der zu bewertenden Dokumente: <10
- ☐ reduzierte Darstellung der Ergebnisdokumente
- ☐ Art der Bewertung:
  - ☐ relevant und keine Bewertung
  - ☐ relevant, irrelevant und keine Bewertung
  - ☐ gestufte Relevanzwerte: gestufte Relevanzwerte

## Bewertung von Dokumenten

- ☐ Bewertungsgranulat:
  - unterschiedliche Bewertungen eines Dokumentes bezüglich verschiedener Dokumenteigenschaften; erhöhter Bewertungsaufwand
  - ☐ mehrere Anfrageobjekte
  - ☐ Ähnlichkeit aufgrund verschiedener Eigenschaftswerte
- ☐ **Pseudorelevanz**: automatische Bewertung

## Bewertungsauswertung

Dokumentenbewertung kann auslösen:

- ☐ Anfragemodifikation
- ☐ Modifikation von Nutzerprofilen
- ☐ Modifikation der Dokumentbeschreibungen
- ☐ Modifikation des Suchalgorithmus
- ☐ Modifikation von Anfragetermgewichten

## Verfahren von Rocchio

- ☐ eigentlich entwickelt für Textdokumente, aber auf andere Medientypen übertragbar
- ☐ Modifikation von Termgewichten des Anfragevektors im Vektorraummodell
- ☐ Termgewichte relevanter Dokumente werden verstärkt und die Termgewichte irrelevanter Dokumente abgeschwächt
- ☐ Verschiebung des Anfragepunktes innerhalb des Vektorraums in Richtung der relevanten Dokumente

## Verfahren von Rocchio

- ❑ Menge » $D_r$ « enthält alle bezüglich der Anfrage » $q_{alt}$ « als relevant markierten Dokumente und

- ❑ » $D_i$ « alle diesbezüglich irrelevanten Dokumente

- ❑ Modifikation des Anfragevektors:

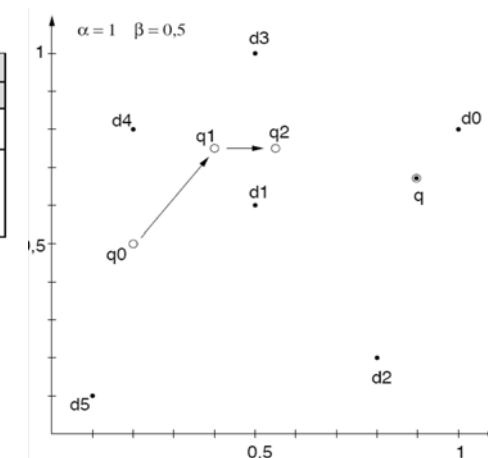
$$q_{neu} = q_{alt} + \frac{\alpha}{|D_r|} \sum_{d_r \in D_r} d_r - \frac{\beta}{|D_i|} \sum_{d_i \in D_i} d_i$$

- ❑ » $\alpha$ « und » $\beta$ « sind Koeffizienten und gewichten den Einfluss der relevanten und irrelevanten Dokumente

## IV.1.3 Bewertung von Retrieval Systemen

## Beispiel

Anfrage	Ergebnisdokumente			
	1	2	3	...
$q$	$d_0$	$d_1$	$d_2$	...
$q_0$	$d_4$	$d_1 (+)$	$d_5 (-)$	...
$q_1$	$d_1 (+)$	$d_3 (+)$	$d_4 (-)$	...
$q_2$	$d_3$	$d_1$	$d_0$	...

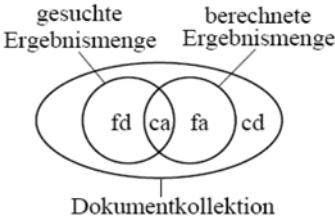


## Definitionen

- ❑ **false alarms** — Dokumente, die vom Retrieval-System irrtümlicherweise als relevant zurückgeliefert wurden.
- ❑ **false dismissals** — Dokumente, die fälschlicherweise vom Retrieval-System als irrelevant eingestuft wurden und damit im Ergebnis nicht erscheinen.
- ❑ **correct alarms**
- ❑ **correct dismissals**
- ❑ Variablen gelten jeweils gegenüber einer Anfrage  $q$

Definitionen

$| \text{gesuchte Ergebnismenge} | = fd + ca$   
 $| \text{berechnete Ergebnismenge} | = ca + fa$   
 $| \text{Dokumentkollection} | = fd + ca + fa + cd$



Definitionen

Precision

$P_q = \frac{ca}{ca + fa}$

Recall

$R_q = \frac{ca}{ca + fd}$

Fallout

$F_q = \frac{fa}{fa + cd}$

Precision-Recall-Paare

- Abhängigkeit des Precision- und Recall-Wertes von der Größe der Ergebnismenge kann in Form einer Linie in einem Precision-Recall-Diagramm dargestellt werden.
- inkrementelles Vergrößern der Ergebnismenge

Beispiel 1

Anzahl	1	2	3	4	5	6	7	8	9	10
$P_1$	$\frac{1}{1}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{3}{4}$	$\frac{3}{5}$	$\frac{3}{6}$	$\frac{3}{7}$	$\frac{4}{8}$	$\frac{5}{9}$	$\frac{5}{10}$
$R_1$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{3}{5}$	$\frac{3}{5}$	$\frac{3}{5}$	$\frac{3}{5}$	$\frac{4}{5}$	$\frac{5}{5}$	$\frac{5}{5}$
$P_2$	$\frac{1}{1}$	$\frac{2}{2}$	$\frac{2}{3}$	$\frac{2}{4}$	$\frac{3}{5}$	$\frac{4}{6}$	$\frac{4}{7}$	$\frac{5}{8}$	$\frac{5}{9}$	$\frac{5}{10}$
$R_2$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$	$\frac{2}{5}$	$\frac{3}{5}$	$\frac{4}{5}$	$\frac{4}{5}$	$\frac{5}{5}$	$\frac{5}{5}$	$\frac{5}{5}$
$P_3$	$\frac{0}{1}$	$\frac{0}{2}$	$\frac{0}{3}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{2}{6}$	$\frac{2}{7}$	$\frac{2}{8}$	$\frac{3}{9}$	$\frac{3}{10}$
$R_3$	$\frac{0}{3}$	$\frac{0}{3}$	$\frac{0}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{3}{3}$	$\frac{3}{3}$	$\frac{3}{3}$

## Beispiel 2

