


# 5. Ton und Klang

- 5.1 Ton: Physikalische und physiologische Aspekte
- 5.2 Kompression von Audio-Signalen: MPEG-Audio
- 5.3 Audio-Datenformate: Übersicht 
- 5.4 Klangerzeugung und MIDI

Weiterführende Literatur:

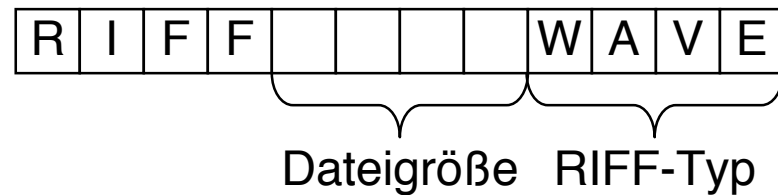
Arne Heyda, Marc Briede, Ulrich Schmidt: Datenformate im Medienbereich, Fachbuchverlag Leipzig 2003

# RIFF (Resource Interchange File Format)

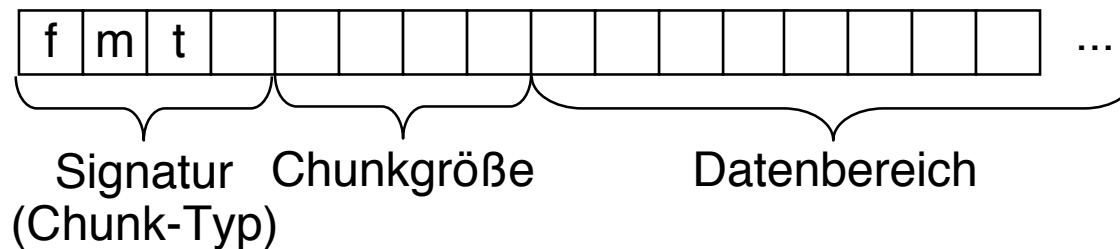
- **IFF:** 1985 von der Firma Electronic Arts eingeführt
  - Sehr einfaches Einheitsformat für verschiedene Arten von Multimedia-Daten, stark verbreitet auf AMIGA-Rechnern
  - Prinzip („Tagged File Format“):
    - » Header gibt Dateityp an
    - » Eigentliche Daten in einer Folge von ebenfalls (über Header) typisierten *chunks*
- **RIFF:**
  - Bestandteil der „Multimedia Programming Interface and Data Specifications“ von Microsoft und IBM, 1991
  - Basiert auf der Idee von IFF
  - Existiert prinzipiell in zwei Varianten:
    - » RIFF für Intel-Architektur („little-endian“)
    - » RIFX für Motorola-Architektur („big-endian“)(RIFX heutzutage auch auf Motorola-Prozessoren ungebräuchlich)

# Grundstruktur von RIFF-Dateien

RIFF-Header (in Bytes):



Chunk-Header (in Bytes):



- Verbreitete RIFF-Datentypen (als eigenständige Dateiformate bekannt):
  - WAVE (oder .wav): Audio, unkomprimiert
  - AVI: Video (Audio/Video Interlaced), unkomprimiert
  - RMI: MIDI-Daten (sh. später)
  - BND: „Bündel“ von RIFF-Dateien

# Wave-Format

- Spezialfall des RIFF-Formats (RIFF-Typ „WAVE“)
- Zwei Arten von Chunks:
  - FMT-Chunk (Signatur „fmt“)
    - » Format-Typ (z.B. MS PCM, IBM ADPCM)
    - » Anzahl Kanäle
    - » Sampling-Rate (Hz)
    - » Datenrate (Bytes/s)
    - » Größe von Datenblöcken
    - » Formatspezifische Information  
(Z.B. bei MS PCM 2 Byte Sample-Größe (bits/Sample))
  - DATA-Chunk (Signatur „data“), meist nur ein solcher Chunk vorhanden
    - » Bei mehreren Kanälen „interleaving“,  
d.h. alle Kanäle für einen Zeitpunkt in Folge

# Beispiel: Hexadezimaler “Dump”

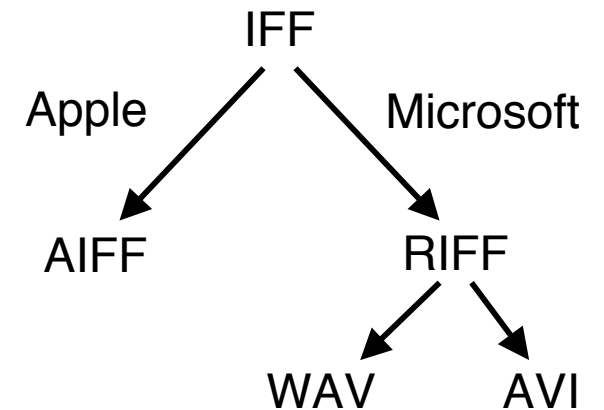
```

Hex Dump for "bach.WAV"
00000000: 52 49 46 46 24 0A 0F 00 57 41 56 45 66 6D 74 20  RIFF$....WAVEfmt
00000010: 10 00 00 00 01 00 02 00 44 AC 00 00 10 B1 02 00  .....D.....
00000020: 04 00 10 00 73 6D 70 6C 3C 00 00 00 00 00 00 00  ....smpl<.....
00000030: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00  .....
00000040: 00 00 00 00 00 00 00 00 01 00 00 00 00 00 00 00  .....
00000050: 00 00 00 00 00 00 00 00 00 00 00 00 74 01 03 00  .....t...
00000060: 00 00 00 00 00 00 00 00 64 61 74 61 2C F8 0E 00  .....data,...
00000070: B8 FA 5C 01 D5 FC 8C 01 72 F9 CD F8 5E FF A5 FC  ..\.....r...^...
00000080: 2D F8 CB F2 8E F2 E3 EC 69 FD 3D FA 85 F9 92 F5  -.....i.=.....
00000090: 0D FF D1 F7 F0 03 95 F9 0F 03 01 F8 33 03 F5 FA  .....3...
000000A0: 97 07 C7 03 31 11 86 0F 0A 14 DC 13 8E 0D B0 0E  ....1.....
000000B0: 16 00 B7 02 E7 08 8F 0B 60 07 26 0B BD FE 71 00  .....`.&...q.
000000C0: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00  .....

```

# AIFF (Audio Interchange File Format)

- Herstellerspezifische Erweiterung von IFF durch Apple für unkomprimiertes Audio
  - Format-Chunk
  - Daten-Chunks, byteweise gepackt
- Audiodaten für bis zu 6 Kanäle (Surround Sound)
- Möglichkeit zur Einstreuung von MIDI-Chunks und Instrumenten-Chunks
- Spezialvariante AIFF-C für komprimierte Audiodaten (ca. 6:1)



# AU (Audio File Format)

- Bei NeXT entwickelt, weit verbreitet im UNIX-Bereich (z.B. Sun)
- Header:
  - Abtastrate, Kanalzahl, Datenformat etc.
  - beliebig lange Textinformation
- Datenbereich:
  - Kanäle miteinander verschränkt
  - Viele Datenformate, z.B.:
    - » von 8 bis 32 Bit
    - »  $\mu$ -Law und linear
    - » Festkomma, Gleitkomma, doppelte Genauigkeit
- Unterstützung von Dateifragmentierung

# QuickTime

- Bibliothek von systemnahen Programmen für MacOS und Windows für die Bearbeitung von zeitbasierten Medien („movies“)
  - Entwickelt von Apple ca. 1991-heute
- Sehr allgemeines Konzept für Medienstrukturen
  - „Atom“: Allgemeiner Container für Mediendaten
  - Mehrere Tracks je Präsentation
  - Pro Track:
    - » Medienstruktur (Referenzen zu Medien verschiedenen Typs)
    - » „Edit List“ für Zeitsynchronisation
- QuickTime wurde als Basis für die MPEG-4 Dateistruktur gewählt.
- Viele verschiedene Dateitypen von QuickTime unterstützt
  - Wichtiges spezifisches QuickTime-Format: „Movie“ (MOV)



```

'moov' - Movie
  'mvhd' - Movie Header
  'trak' - Track
    'tkhd' - Track Header
    'edts' - Edits
      'elst' - Edit List
    'mdia' - Media
      'mdhd' - Media Handler Header
      'hdlr' - Handler Description
      'minf' - Media Information
        'vmhd' - Video Media Header
        'hdlr' - Handler Description
        'dinf' - Data Handler Information
          'dref' - Data Reference
        'stbl' - Sample Table
          'stsd' - Sample Descriptions
          'stts' - Sample to Time
          'stsc' - Sample to Chunk
          'stsz' - Sample Sizes
          'stco' - Chunk Offset Table
      'udta' - User Data
  'trak' - Track
    'tkhd' - Track Header
    'edts' - Edits
      'elst' - Edit List
    'mdia' - Media
      'mdhd' - Media Handler Header
      'hdlr' - Handler Description
      'minf' - Media Information
        'smhd' - Sound Media Header
        'hdlr' - Handler Description
        'dinf' - Data Handler Information
          'dref' - Data Reference
        'stbl' - Sample Table
          'stsd' - Sample Descriptions
            flags  $00000000 numEntries  1
            descSize  52 numChannels
            dataFormat  s0wt sampleSize
            dataRefIndex  1 sampleRate
            packetSize  0 compressionID
            bytesPerPacket  2 bytesPerFrame
          'stts' - Sample to Time
          'stsc' - Sample to Chunk
          'stsz' - Sample Sizes
          'stco' - Chunk Offset Table
      'udta' - User Data
  'udta' - User Data
  'VLOC'

```

# Beispiel: QuickTime Dateistruktur

2	version
16	revlevel
0.032000	vendor
-1	samplesPerPacket
4	bytesPerSample

1	reserved	0
0		0
0		
1		
2		

# Verlustfreie Audio-Kompression: Beispiele

- MPEG-4 Audio Lossless Coding (ALS)
  - TU Berlin, Real Networks, NTT
  - Basiert auf LPC-Codierung (und Golomb-Rice-Codierung)
- FLAC (Free Audio Lossless Coding)
  - Josh Coalson, jetzt bei Xiph.org Stiftung
  - Ca. 50% Reduktion der Dateigröße
  - Lineare Prädiktion, Lauflängen-Codierung, Golomb-Rice-Codierung
- ALAC (Apple Lossless Audio Codec)
  - 40-60 % Reduktion
  - In MPEG-4 Container gespeichert

# 5. Ton und Klang

- 5.1 Ton: Physikalische und physiologische Aspekte
- 5.2 Kompression von Audio-Signalen: MPEG-Audio
- 5.3 Audio-Datenformate: Übersicht
- 5.4 Klangerzeugung und MIDI



Literatur:

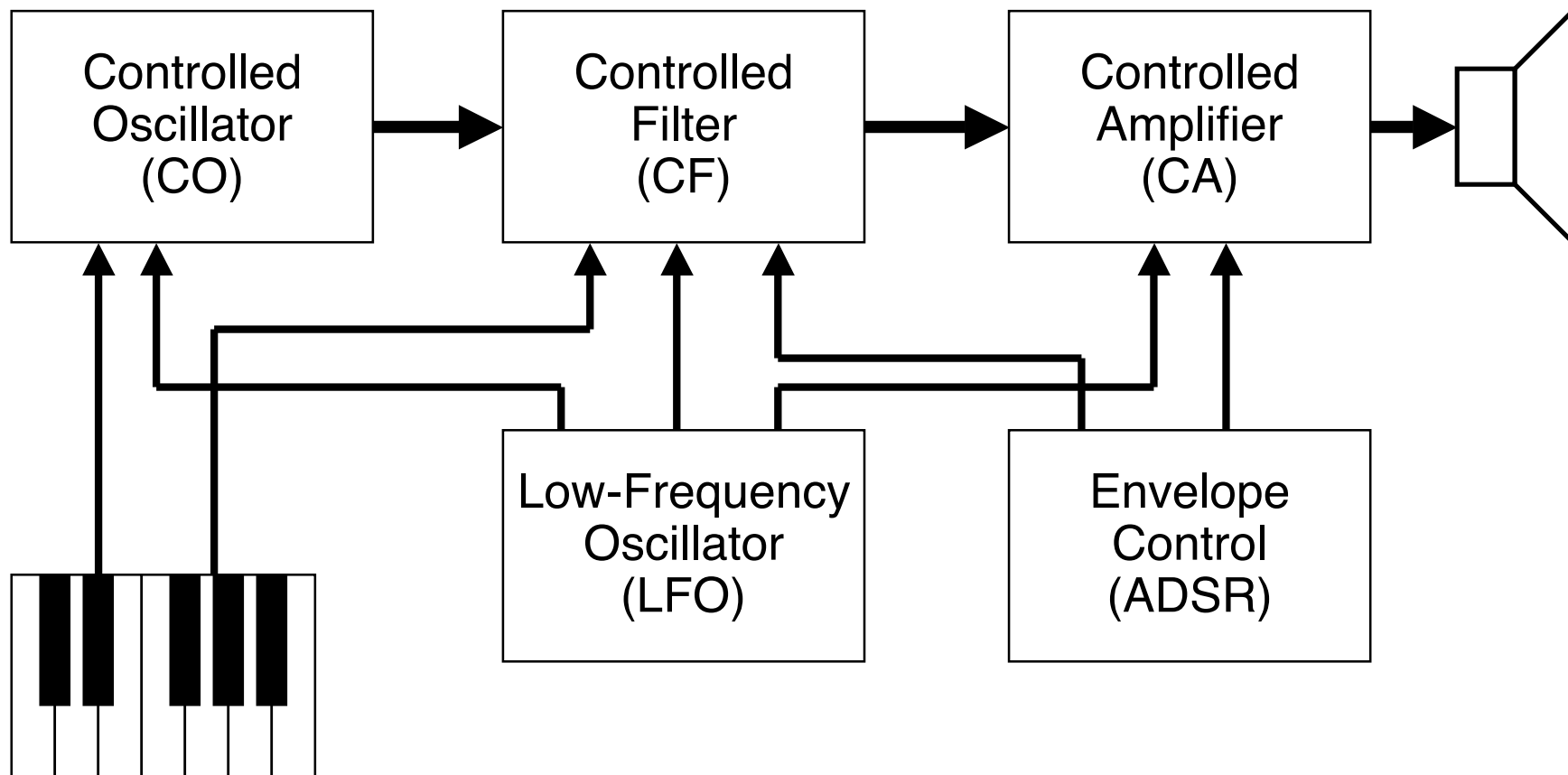
Hannes Raffaseder: Audiodesign, Fachbuchverlag Leipzig 2002

# Elektronische Klangerzeugung

- Klänge für Musik oder Sprache können künstlich produziert werden
  - Tonhöhe, Lautstärke, Klangfarbe (*timbre*) einstellbar
- Klangerzeuger:
  - Einfache Klangerzeuger in Soundkarten enthalten (Frequenzmodulation einfacher Wellenformen)
  - Hochwertige Klangerzeuger z.B. in elektronischen Musikinstrumenten („Synthesizer“, MIDI-Keyboards)
    - » mehrstimmig (z.B. 128)
    - » multitimbral (z.B. 64 Klangfarben)
- Historisch gesehen:
  - 1900 Dynamophone (Thaddeus Cahill), 1920 Termenvox (Lew Termen), 1930 Trautonium (Friedrich Trautwein), 1960 Mellotron
  - Anfang der 60er Jahre (Robert Moog): Moderne Synthesizer-Architektur
  - 1968 Walter Carlos „Switched-on Bach“



# Grundstruktur eines Synthesizers



# Grundelemente bei der Klangerzeugung

- Oszillator
  - Erzeugt mehr oder weniger obertonreiches Signal, das die Grundfrequenz und auch wesentlich den Klangcharakter bestimmt
- Filter
  - Z.B. Hochpass, Tiefpass, Bandfilter
- Verstärker (*Amplifier*)
  - Kann über zeitabhängigen Pegelverlauf Klangempfindung wesentlich beeinflussen
- Hüllkurvengenerator (*Envelope Control*)
  - Zeitlicher Verlauf eines Klangereignisses auf ein einmaliges erzeugendes Ereignis hin (z.B. Tastendruck), meist ADSR (siehe nächste Folie)
- *Low Frequency Oscillator LFO*
  - Dient zur kontinuierlichen Veränderung eines klangbestimmenden Parameters innerhalb eines bestimmten Zeitintervalls
  - Frequenzen typischerweise zwischen 0 und 20 Hz

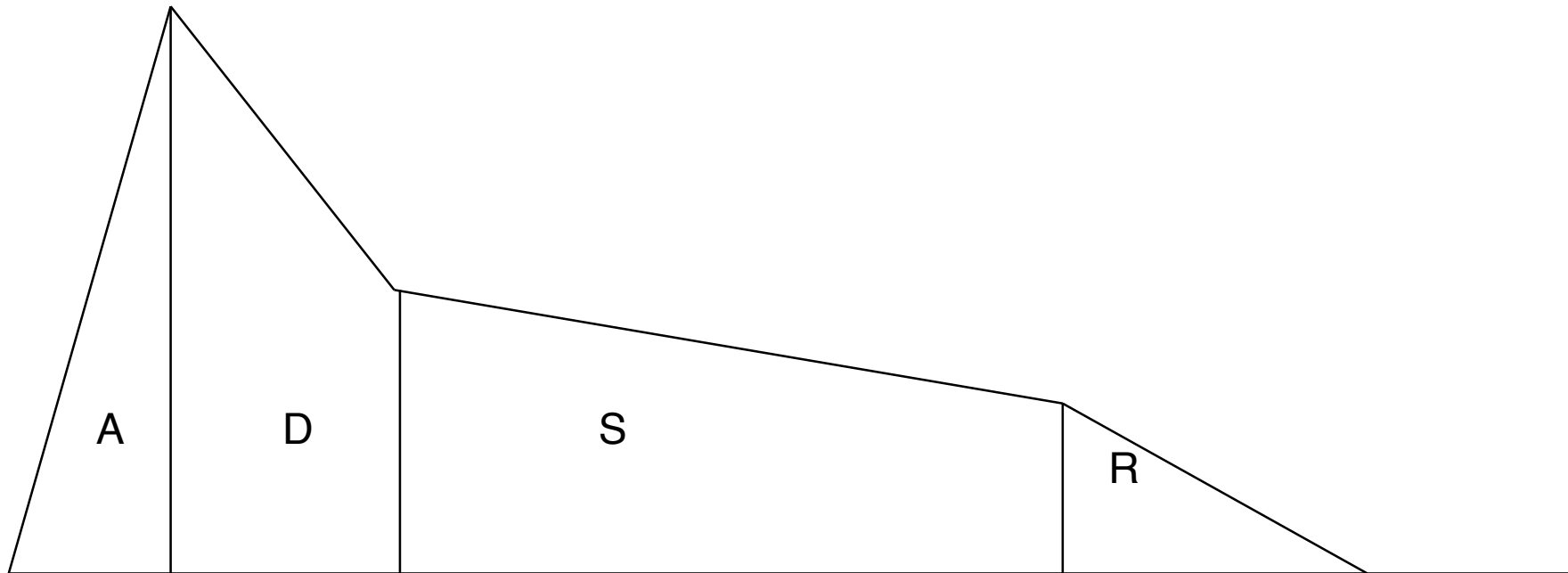
# Beispiel: Software-Synthesizer



Software: Propellerhead Reason

# ADSR-Modell

- Modulation nach dem ADSR-Modell
  - Attack (A), Decay (D), Sustain (S), Release (R)
  - Höhe und Breite der vier Parameter variabel
  - Verbreitet: Einstellung der A-, D-, S- und R-Zeiten über Regler





# Verfahren zur Klangsynthese

- Additive Klangsynthese
  - Fourier-Analyse in der Praxis, zur Synthese umgewandelt
  - Realisierung eines Klangs als Überlagerung von Sinustönen
- Subtraktive Klangsynthese
  - Erzeugung obertonreicher Grundsignale (z.B. Sägezahn, Dreieck, ...)
  - Steuerung der Spektren mit Filtern und der Amplitude mit Verstärkern
- Wavetable-Synthese
  - Vordefinierte, gespeicherte Wellenformen
  - Oszillator durchläuft Wavetable in programmierter Weise (z.B. LFO)
- Sampling
  - Wiedergabe digital aufgezeichneter akustischer Ereignisse
  - Multisampling: Viele Aufnahmen mit verschiedenen Parameterwerte
- Granularsynthese
  - Zerlegung von Schallsignalen in *Grains* (wenige ms lange Abschnitte)
  - Entkopplung von Wiedergabegeschwindigkeit und Tonhöhe

# MIDI: Geschichte und Überblick

- Synthesizer: Revolutionäres Musikinstrument in den 70er Jahren
  - Beatles (White Album), Carlos (Switched-on Bach), ...
  - Technische Probleme:  
Polyphonie, Kombination verschiedener Geräte, Synchronisation
- 1983: Erste Interoperabilitäts-Vorführung
- MIDI (Musical Instrument Digital Interface) Standard
  - International MIDI Association (IMA)
  - MIDI Manufacturers Association (MMA)
- Bedeutung für Multimedia:
  - Standardisierte Sprache für
    - » Übernahme von Daten aus Endgeräten, die Musikinstrumenten entsprechen (insb. Keyboard)
    - » Ansteuerung von Peripheriegeräten (Synthesizer, Beleuchtung)
    - » Abstrahierte Darstellung von gespielter Musik

# MIDI-Grundbegriffe

- Ereignis (*event*):
  - Musikalische Aktion, z.B. Musiker drückt Taste auf Keyboard mit bestimmter Anschlagsstärke (*velocity*)
    - » etwa: „NOTE ON C3 velocity 100“
  - Jedes Ereignis findet zu einem bestimmten Zeitpunkt statt (Zeitstempel)
- Nachricht (*message*):
  - Binäre Codierung der in einem Ereignis enthaltenen Information
  - Kann gespeichert, weitergegeben, vervielfältigt, modifiziert werden
- Befehl (*command*):
  - Anweisung an ein externes Gerät, bestimmte musikalische Aktionen auszuführen
- Klangfarbe (*timbre*):
  - Charakteristik eines bestimmten wiederzugebenden Instruments
  - „Multitimbral“
- Kanal (*channel*):
  - Identifikator für bestimmten Empfänger (traditionell 16 Kanäle)
  - „Musikinstrument“ bzw. entsprechender Klangerzeugungsprozess

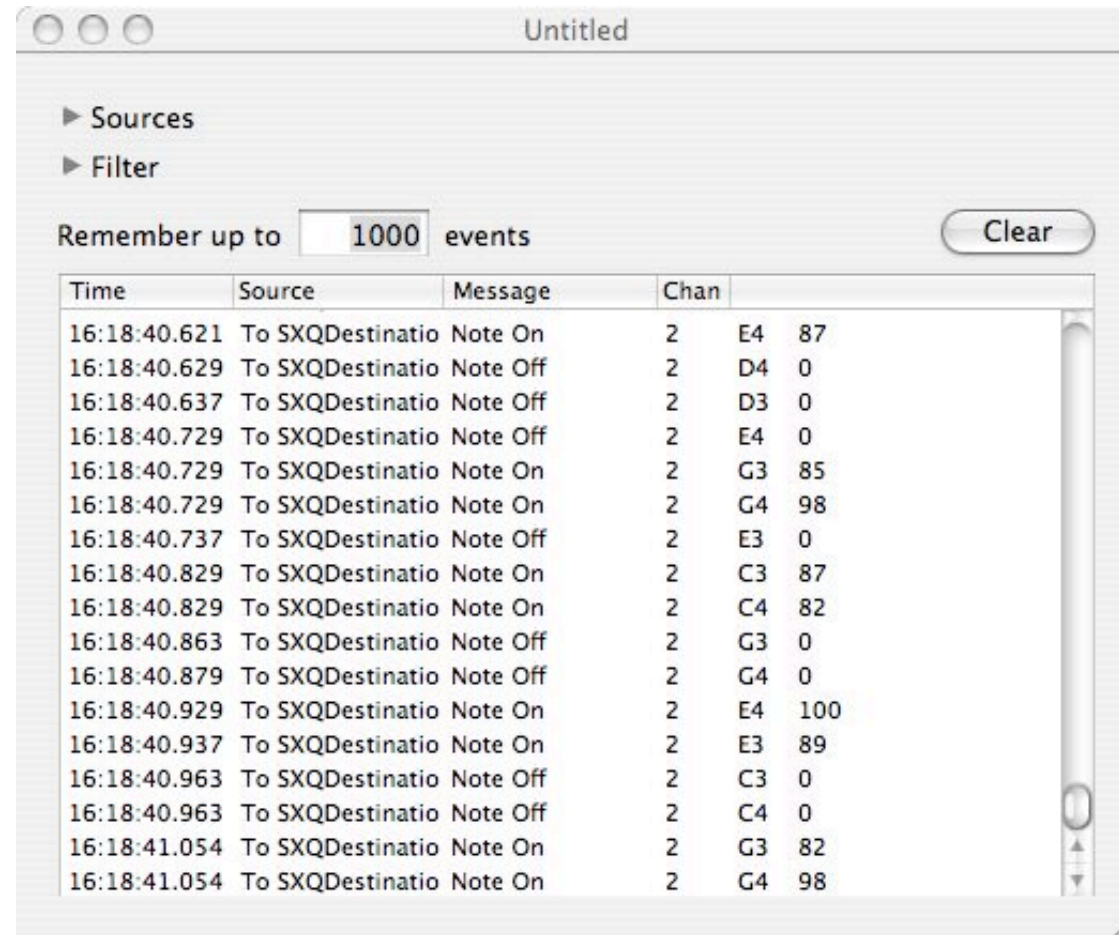
# MIDI-Nachrichten

- Channel Voice Messages
  - Eigentliche Musikdaten (sh. nächste Folie)
- Channel Mode Messages
  - Steuerung des Synthesizers
    - » Ein-/Ausschalten der eigenen Tastatur (z.B. bei Keyboard/Synthesizer)
    - » Testmodus
    - » Polyphonie-Steuerung
- System Real-Time Messages
  - Synchronisationstakt
  - Synchronisierte Sequenzen
  - Überprüfung der Verfügbarkeit von Geräten
- System Exclusive Messages (SysEx)
  - Weitergabe herstellerepezifischer Information an individuelle Geräte

# Inhalt einer MIDI-Datei: MIDI-Ereignisse

- Header-Information
- Track-Information
  - *Track* = Separat abspielbare und bearbeitbare Musikspur
- Track-Information Teil 1: Metainformation
  - Track-Nummer, -Name
  - Angaben zum Instrument (z.B. aus *General Midi*-Instrumenten)
  - Zeitbasis
- Track-Information Teil 2: Melodie
  - Folge von Channel Voice Messages, jeweils mit Zeitstempel relativ zur Zeitbasis
  - Note On (Parameter Notenwert, Anschlagstärke)
  - Note Off (Parameter Notenwert, Anschlagstärke)
  - Polyphonic Key Pressure (Parameter Notenwert, Anschlagstärke)  
(Änderung der Anschlagstärke über die Zeit)
  - Pitch Bend Change (Parameter Verschiebung)  
(Tonhöhenverstellung)

# MIDI Ereignisse: Beispiel



Remember up to  events Clear

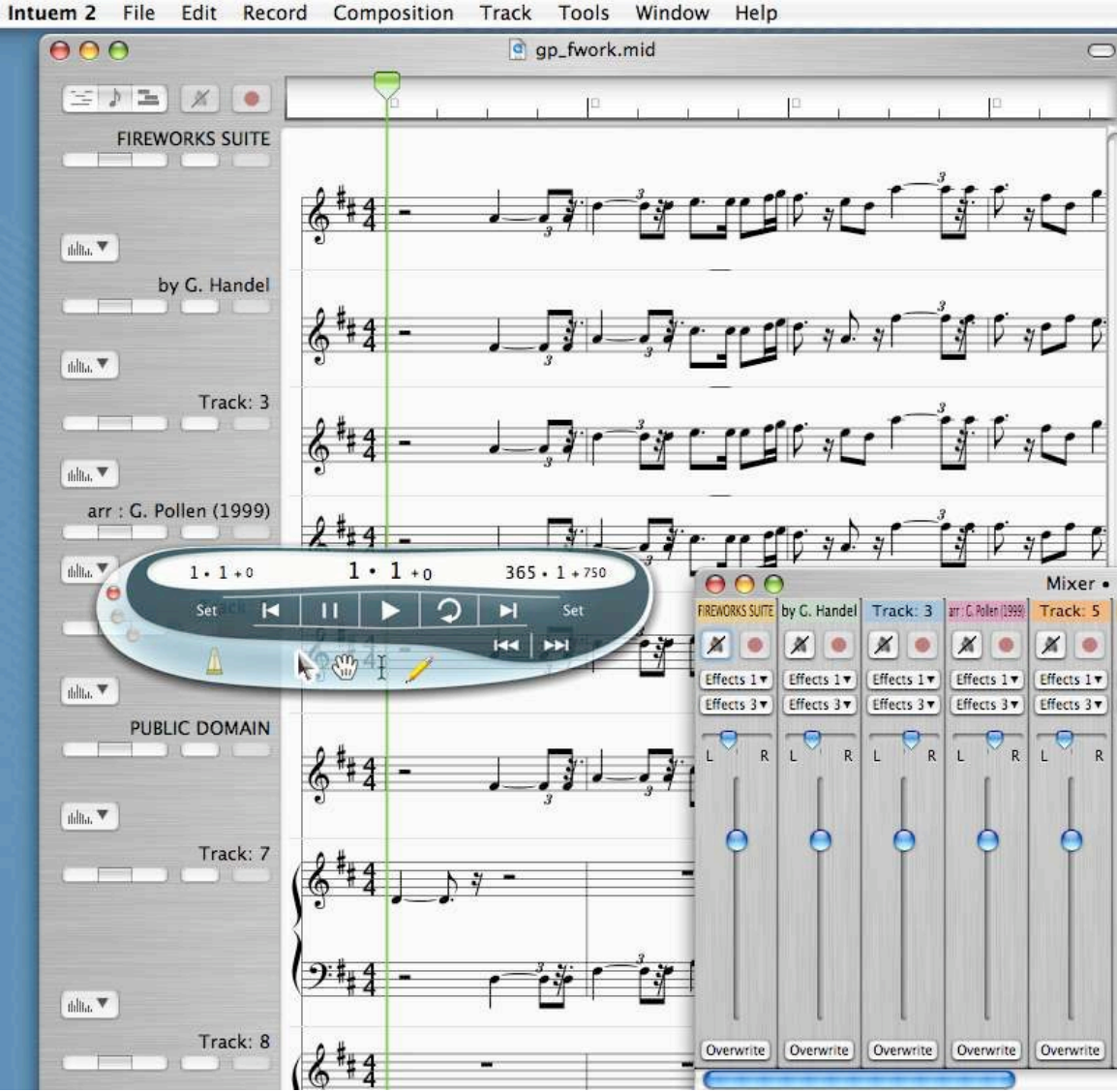
Time	Source	Message	Chan
16:18:40.621	To SXQDestinatio	Note On	2 E4 87
16:18:40.629	To SXQDestinatio	Note Off	2 D4 0
16:18:40.637	To SXQDestinatio	Note Off	2 D3 0
16:18:40.729	To SXQDestinatio	Note Off	2 E4 0
16:18:40.729	To SXQDestinatio	Note On	2 G3 85
16:18:40.729	To SXQDestinatio	Note On	2 G4 98
16:18:40.737	To SXQDestinatio	Note Off	2 E3 0
16:18:40.829	To SXQDestinatio	Note On	2 C3 87
16:18:40.829	To SXQDestinatio	Note On	2 C4 82
16:18:40.863	To SXQDestinatio	Note Off	2 G3 0
16:18:40.879	To SXQDestinatio	Note Off	2 G4 0
16:18:40.929	To SXQDestinatio	Note On	2 E4 100
16:18:40.937	To SXQDestinatio	Note On	2 E3 89
16:18:40.963	To SXQDestinatio	Note Off	2 C3 0
16:18:40.963	To SXQDestinatio	Note Off	2 C4 0
16:18:41.054	To SXQDestinatio	Note On	2 G3 82
16:18:41.054	To SXQDestinatio	Note On	2 G4 98

- MIDI-Dateien sind extrem kompakt.
- MIDI-Aufzeichnungen sind genauer als normale Notenschrift!

# Typische Funktionen von MIDI-Sequenzern

- „Sequencer“ = Software zur Bearbeitung von synthetisierter Musik, z.B. mit MIDI
- Aufnehmen und Wiedergeben von Tonspuren
- Verschiedene Ansichten der gleichen Information:
  - Partitur, Keyboard-Matrix
  - Zeitgenaue Liniendarstellung
  - Darstellung von Zusatzinformation (z.B. *velocity*)
- Musik-Editor:
  - Komponieren (Noten einsetzen und verschieben, Längen verändern, Transponieren, ...)
  - Instrumente variieren
  - Effekte einfügen
  - Synchronisieren von Spuren und Abmischen
  - Oft integriert mit klassischer Mischpult-Funktionalität
  - Oft integriert mit Notensatz-Funktionalität

# Beispiel: MIDI-Sequencer



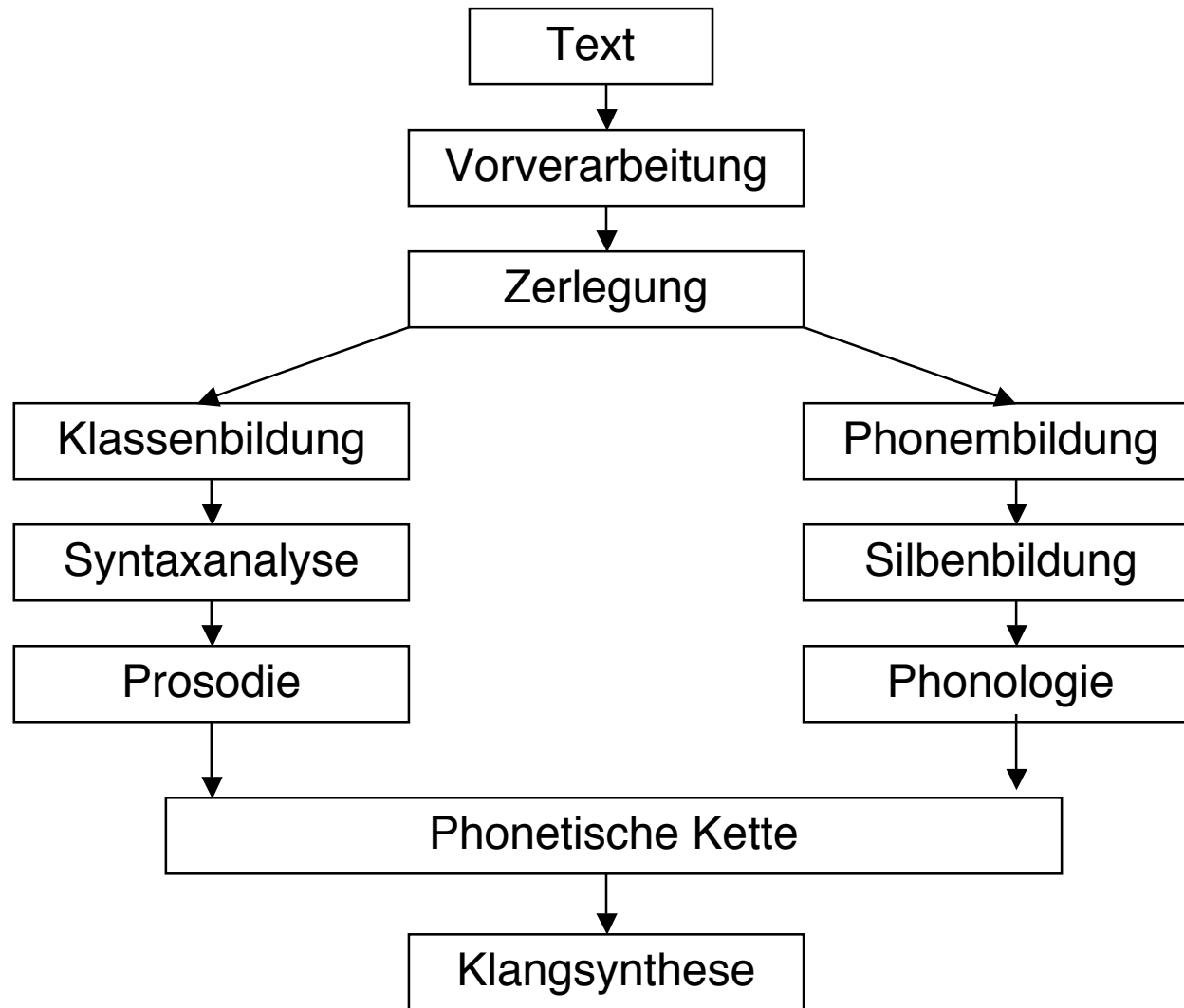
Intuem



# Sprachanalyse und Sprachsynthese

- Ein- und Ausgabe in natürlicher Sprache
  - Alter Traum der Informatik
  - Grenzgebiet zu Computerlinguistik, KI
- Sprachausgabe:
  - relativ stabile Technologie
  - Bestandteil vieler Standard-Betriebssysteme
- Spracheingabe:
  - immer noch relativ wenig beherrscht
  - Trainingsfreie Systeme noch störanfällig
  - Trainingsgebundene Systeme existieren mit akzeptabler Leistung

# Sprachsynthese: Grob Ablauf



# Weiterentwicklung im Bereich Klangerzeugung

- MPEG-4 Standard:
  - *Structured Audio Format* ermöglicht Spezifikation von Klangerzeugern
  - *SAOL (Structured Audio Orchestral Language)* zur Beschreibung von elektronischen Instrumenten und Audioeffekten
  - *SASL (Structured Audio Source Language)* erlaubt differenzierte Formulierung von Spielanweisungen (über MIDI hinaus)
- Anwendungsfeld Interaktion:
  - Akustische Signale in Spielen und Softwaresystemen tendieren dazu, den Benutzer durch stupide Wiederholung zu ermüden
  - Softwaresynthese von Klängen eröffnet die Möglichkeit, situationsabhängig neue Klänge zu generieren, wo erwünscht