

Schachdatenbanken

Schachdatenbanken als Trainingsutensil für
Vereinsspieler

Dr. Sven Hermann
sven@it-s-hermann.de
Karlsruhe 20.2.2018

20.2.2018

Dr. Sven Hermann für die SF Neureut

Inhalt

- Einleitung
- Datenbanken
- Eigenschaften von Partien
- DBs als Trainingsutensil
- Beispiele
- Fazit

Inhalt

- **Einleitung**
- Datenbanken
- Eigenschaften von Partien
- DBs als Trainingsutensil
- Beispiele
- Fazit

Partien (analog)

1 Partie
0,2 kB

The New York International Chess Master's Tournament 1924
ALAMAC HOTEL
71st Street and Broadway, New York

IV
21

Date 22/IV 192

Opening

WHITE Mr. Reti BLACK Mr. Capablanca

White	Black	White	Black
1. Sf3	1 Sf6	31 Td1-d5	31 Anfgabe
2 b4	2 b6		
3 Lb2	3 Lg7		
4 g3	4 b7		
5 Lge	5 Lb7		
6 d3	6 d6		
7 Sd2?	7 Sd7?		
8 d2	8 e5		
9 Tf1	9 e8		
10 a3	10 g5		
11 Sf1	11 b6		
12 b5	12 e7		
13 e3	13 d7		
14 d4	14 Lc9		
15 Sc3	15 e6		
16 Scd2	16 f6		
17 Lxd4	17 Scd		
18 Lg4	18 Sc4		
19 Sd2	19 Sc3		
20 Txd1	20 Sc2		
21 Se3	21 Sc7		
22 Sc4	22 Lxc2		
23 Kc2	23 Sc5		
24 Sc6	24 Sc5		
25 Sc3			
26 Sc3			
27 Sc3			
28 Sc3			
29 Sc3			
30 Sc3			



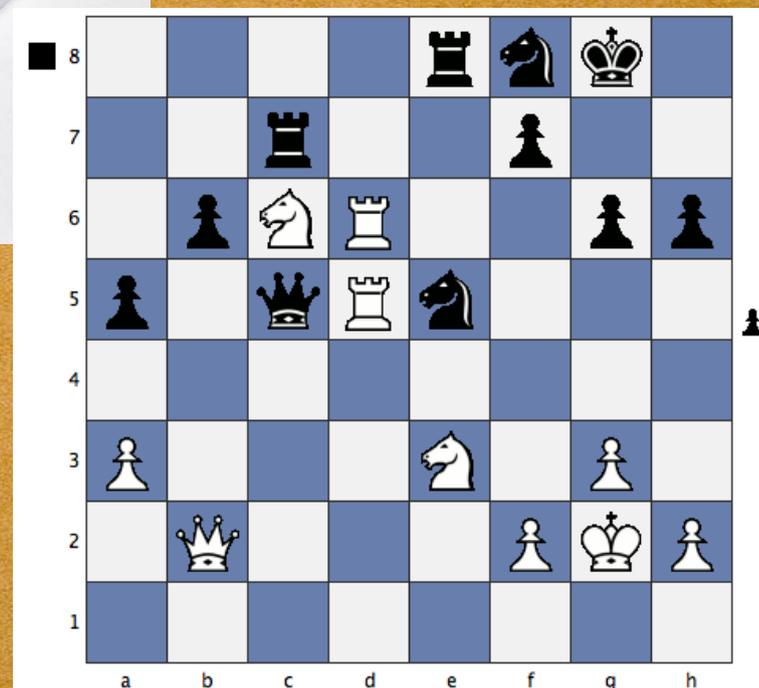
20.2.2018

Dr. Sven Hermann für die SF Neureut

Partien (digital)



5 Milliarden Partien
1.000.000.000 kB



Inhalt

- Einleitung
- **Datenbanken**
- Eigenschaften von Partien
- DBs als Trainingsutensil
- Beispiele
- Fazit

Datenbank

- System zur elektronischen Datenverwaltung
 - Speichert große Datenmengen effizient, widerspruchsfrei und dauerhaft
 - Stellt benötigte Teilmengen in unterschiedlichen, bedarfsgerechten Darstellungsformen bereit
- *Hier: Datenbank aus Schachpartien*

Elektronische Schachpartie

- Portable Game Notation (PGN)
 - als **Text lesbares** Datenformat für Schachpartien
 - **1994** von S. J. Edwards entwickelt
 - zum **Austausch** von Schachdaten zwischen verschiedenen Schachprogrammen
 - ISO-8859-1-Zeichensatz (inklusive **Sonderzeichen**)
 - aus **Metadaten** (Turnier, Spieler, ...) und **Zügen**
 - **nicht proprietär**
 - kann von **fast allen Schachprogrammen** gelesen werden.
 - auch **mehrere Partien** in einer einzigen PGN-Datei

Inhalt

- Einleitung
- Datenbanken
- **Eigenschaften von Partien**
- DBs als Trainingsutensil
- Beispiele
- Fazit

Wieviel Partien gibt es ...

- ... auf Wettkampf-Niveau?
- ... auf Großmeister-Niveau?
- ... seit Beginn der FIDE?
- ... seit Beginn der Menschheit?



Und passen die alle auf
meine Festplatte?



Partien: Wettkampf-Niveau

- Seit 1924: FIDE (modernes Schach)
0,004%* aller Menschen **spielen im Verein**
- Seither weltweit ca. $10 \cdot 10^9$ Geburten
- Grobe Annahme:

Jeder „Vereinsspieler“ spielt 2 ernste Partien pro Woche

Alle Partien seit 1924

Anzahl $10 \cdot 10^9 \cdot 0,004\% \cdot 2 \cdot 52 \cdot 93 = 4 \cdot 10^9$

Daten $4 \cdot 10^9 \cdot 0,2 \text{ kB} = \mathbf{800 \text{ GB}}$

* 2010 gab es 6,9 Milliarden Menschen wovon 280.000 in der FIDE als Schachspieler registriert waren
→ ca. 0,004% aller Menschen (am Beispieljahr 2010) spielen „Vereinschach“

Partien: Großmeister-Niveau

- Seit 1924: FIDE (modernes Schach)
0,0002‰* aller Menschen spielen wie GMs
- Seither weltweit ca. $10 \cdot 10^9$ Geburten
- Grobe Annahme:

Jeder „Großmeister“ spielt 1 ernste Partie pro Tag

Alle Partien seit 1924

Anzahl $10 \cdot 10^9 \cdot 0,0002‰ \cdot 365 \cdot 93 = 7 \cdot 10^7$

Daten $7 \cdot 10^7 * 0,2 \text{ kB} = 15 \text{ GB}$

* 2010 gab es 6,9 Milliarden Menschen wovon 1300 in der FIDE als Großmeister anerkannt waren
→ ca. 0,0002‰ aller Menschen (am Beispieljahr 2010) spielen „GM-Schach“

Eigenschaften von Partien

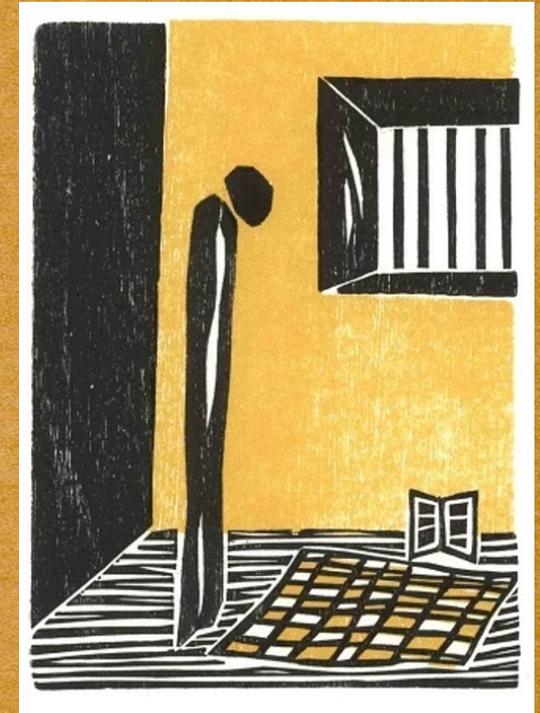
- Bestimmte **Stellung**
- Eröffnung/Zugfolge/**Variante**
- Spieler (weiß/schw., **ELO**, Name, Verein)
- **Ergebnis**
- Materialverteilung
- Muster (z.B. isolierter Damenbauer)
- Endspieltyp
- Datum, Turnier, ...

Inhalt

- Einleitung
- Datenbanken
- Eigenschaften von Partien
- **DBs als Trainingsutensil**
- Beispiele
- Fazit

Quellen für Partien

- Internet
 - TWIC (frei)
 - 365chess.com (Spende)
 - Turniere (frei)
 - bad. Schachbund u.ä. (frei)
 - höhere Schachligen (frei), ...
- eigene Partien
- Kommerzielle DBs (Chessbase Mega DB u.ä.)



Von Elke Rehder - Mirdsson2, CC BY-SA 3.0.
<https://commons.wikimedia.org/w/index.php?curid=22766804>

Datenbanken (inhaltlich)

- Referenzdatenbank
 - Enthält *alle* Partien
 - Dient als Pool für Suchen/Filter
 - Kann alle jemals in der FIDE gespielten Turnierpartien und mehr enthalten
- DBs für
 - einzelne Spieler oder Vereine
 - Eröffnungen, Analysen ...
 - Endspiele, Taktikaufgaben, ...

Suchen in der DB (Bsp. SCID)

- Stellung (*Stellungssuche*)



- ELO & Ergebnis (*Partiedaten*)



- Variante (*Zugbaum*)

▼ Zugbaum [french]									
Zug	Häufigkeit		Pkte	Remis	Elo	Erflg	Jahr	ECO	
1: Sc3	102435: 41.1%		56.0%	29%	2257	2285	1997	C10a	
2: Sd2	72425: 29.1%		57.0%	33%	2261	2297	1999	C03a	
3: e5	40884: 16.4%		50.1%	24%	2139	2139	1999	C02a	
4: exd5	30525: 12.2%		46.1%	37%	2048	2081	1998	C01a	
5: Ld3	1503: 0.6%		51.8%	27%	2162	2173	2001	C00y	
6: Le3	528: 0.2%		52.5%	24%	1995	2029	1994	C00x	
7: c4	122: 0.0%		46.3%	19%	1945	1945	1999	D30b	
8: f3	82: 0.0%		28.6%	13%	1754	1679	2000		
9: Sf3	39: 0.0%		38.4%	21%	1751	1798	2001		
10: [end]	32: 0.0%		50.0%	88%	2179	2139	1997	C00x	
11: c3	10: 0.0%		10.0%	0%			1995		
12: Lf5	10: 0.0%		20.0%	0%			2000		

Inhalt

- Einleitung
- Datenbanken
- Eigenschaften von Partien
- DBs als Trainingsutensil
- **Beispiele**
- Fazit

Beispiel 1

- Frage: Wie bereite ich mich gezielt auf eine Eröffnung vor?
- Lösung:
 - Partien sammeln
 - Datenbank für Eröffnung anlegen
 - Alle Partien der betreffenden Farbe hinein kopieren
 - Vergleichspartien von Großmeistern f ltern
- Ergebnis: Essenz für eine Eröffnung

Beispiel 2

- Frage:
Wie spiele ich diese* Stellung auf Sieg?
- Lösung: *Filter* (Suche)
 - nach einer Stellung **und**
 - nach Spielstärke (ELO > 2400) **und**
 - nach Ergebnis
- Ergebnis: Einige dutzend Partien
so spielen Großmeister sie zum Gewinn!

* frühe oder späte

Inhalt

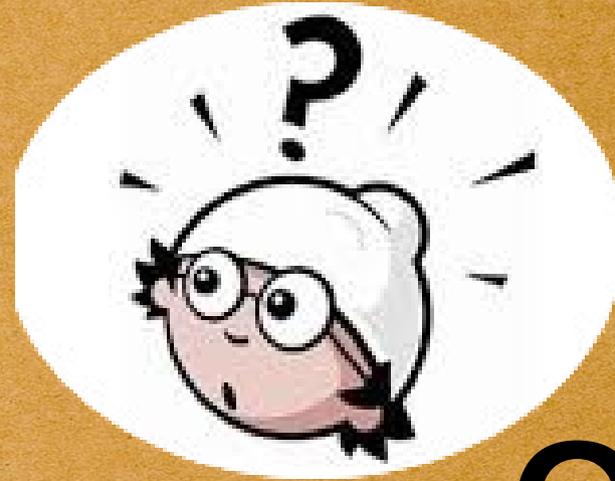
- Einleitung
- Datenbanken
- Eigenschaften von Partien
- DBs als Trainingsutensil
- Beispiele
- **Fazit**

Fazit

- Unmengen Partien frei verfügbar
- Abfragen der Datenbank
 - gehen in Sekundenschnelle
 - bei Millionen Partien
 - bei komplexen Parameterabfragen
 - Stellungssuche komfortabel
 - Suche nach ELO liefert Meisterpartien
- DBs für jeden Zweck (Eröffnung, Gegner, ...)

Diskussion

!?



?!

20.2.2018

Dr. Sven Hermann für die SF Neureut