



BERGISCHE UNIVERSITÄT
GESAMTHOCHSCHULE WUPPERTAL

Die Entropie von Symbolfolgen

Dissertation

zur Erlangung des Doktorgrades im Fach Theoretische Physik
an der Bergischen Universität Gesamthochschule Wuppertal

vorgelegt von

Thomas Schürmann
aus Düsseldorf

22. März 1996
WUB-DIS 96-4

Inhaltsverzeichnis

1	Einleitung	1
2	Informationstheorie	9
2.1	Der statistische Informationsbegriff	9
2.2	Die Shannon-Entropie	10
2.3	„Finite Sample“ Korrekturen der Entropie	11
2.4	Entropieschätzung nach Lempel und Ziv	15
2.5	Präfix-Bäume und Lempel-Ziv Schätzer	17
3	Entropie und „Gambling“	21
3.1	Die optimale Verteilung von Kapital	21
3.2	Suffix-Bäume und Kontexte	24
3.3	Modellierung von Wahrscheinlichkeiten	27
3.3.1	Bayes'sche Wahrscheinlichkeiten	27
3.3.2	Rissanen's Methode	28
3.3.3	Superposition von Wahrscheinlichkeiten	29
3.3.4	Globale Schätzer für Wahrscheinlichkeiten	30
4	Anwendungen	33
4.1	Chaotische Systeme und Zellulare Automaten	33
4.2	Natürliche Sprache	37
4.2.1	Vorhersage von Buchstaben in Texten	42
4.2.2	Vorhersage von Buchstaben in Worten	43
5	Zusammenfassung und Ausblick	49

5.1	Danksagung	50
A	Die Kolmogorov-Sinai-Entropie	51
A.1	Die Entropie dynamischer Systeme	51
B	Entropie 1-dimensionaler dynamischer Systeme	53
B.1	Dynamik des hybrid dynamischen Systems	54
B.1.1	Symbolische Dynamik	56
B.2	Invariante Dichten und Entropien	56
B.3	Zusammenfassung und Ausblicke	59
C	„Accessibility Sampling“	61
C.1	„Importance Sampling“ und ähnliche Methoden	61
C.2	Die Methode des „Accessibility Sampling“	63
C.3	Die Informationsdimension	67
C.4	Experimentelle Zeitreihen	68