

---

## Stellungstests: Was sie wirklich aussagen

---

**Viel wurde schon über, für und gegen Stellungstests geschrieben. Ob sie Rückschlüsse auf die Spielstärke erlauben, gehört zu den meistdiskutierten Fragen der Computerschach-Welt. Verwunderlich eigentlich, dass sich in diesen Diskussionen noch niemand Gedanken gemacht hat, wovon es abhängt, für welchen Zug sich ein Schachspieler entscheidet – welche Rolle spielen persönliche Vorlieben, gibt es überhaupt Entscheidungsfreiheit? Die Beantwortung dieser Fragen erweist sich als sehr bedeutsam für die Aussagekraft von Stellungstests.**

Eine Schachpartie ist der Kampf von zwei Gegnern, der sich dadurch äußert, dass diese abwechselnd Züge mit den vorhandenen Bauern und Figuren auf dem Schachbrett ausführen. Nach jedem Zug ergibt die bis dahin geschehene Gesamtheit der Züge Stellungen, die – ebenso wie die einzelnen Züge für sich – einen relativen Wert besitzen. Die Halbzüge sind die einzige Möglichkeit der Spieler, Einfluss auf das spätere Ergebnis zu nehmen. Daher muss ihnen im Einzelnen und im Verbund besondere Aufmerksamkeit gewidmet werden. Die Qualität von Halbzügen kann höchst unterschiedlich sein.

Bemerkenswert ist es in dem Zusammenhang festzustellen, dass menschliche Halbzüge meistens zu Recht Ausrufezeichen für herausragende taktische Leistungen erhalten, wogegen diese Halbzüge für Computer trivial sind. Dennoch pflegen menschliche Kommentatoren auch solchen Computer-Halbzügen gute Bewertungen zu geben, obwohl quasi jedes Programm so gezogen hätte. Hier liegt eindeutig ein subjektiver Bewertungsfehler des Menschen vor.

Viel wichtiger als Halbzüge, die mit Ausrufezeichen geziert werden, sind jene, die ein Fragezeichen haben. Der Grund liegt darin, dass ein noch so toller Zug mit „!!“, nicht den geringsten Einfluss auf die Bewertung einer Stellung haben kann, denn er dient lediglich dazu, den Wert der Stellung, den diese schon vor dem starken Halbzug objektiv besaß, beizubehalten.

Im Gegensatz dazu können Halbzüge fragwürdigen Charakters (?!, ?, ??) einen sehr hohen und immer negativen Einfluss auf die Bewertung der Stellung haben. Ein einziger „?-Halbzug“ kann eine gute Stellung ruinieren und den Verlust der Partie zur Folge haben.

Fragezeichen werden jedoch oft nach subjektiven Gesichtspunkten vergeben. So werden so genannte positionelle Fehler (z.B. unbegründete Randbauernzüge) oft überbewertet, da sie die objektive Bewertung der Stellung nicht wesentlich ändern.

Anhand vieler Datenbanken lässt sich problemlos nachweisen, dass eine Mehrzahl der Partien von Menschen zwischen 30 und 60 Zügen aufweist. Bei Engine-Partien ist die durchschnittliche Zugzahl etwas höher (ca. 60-70 Züge). Um eine weitere Eingrenzung vorzunehmen, möchte ich in der Folge als durchschnittliche Zugzahl für eine Schachpartie von der Eröffnung bis zur Aufgabe 50 Züge annehmen. Von den 50 Zügen ziehe ich weiter noch 10 Züge ab, da in diesen ersten Zügen oft eine theoretische Variante auf das Brett kommt, die sich für eine Bewertungsanalyse genauso wenig eignet wie die letzten offensichtlichen Züge vor dem Matt.

Von diesen 40 Zügen werden durchschnittlich die Hälfte dafür benötigt, um so genannte „Zwangzüge“ zu machen. Hierzu gehören nicht nur theoretisch einzig mögliche Züge nach einem Schach, sondern auch Schlagfälle, deren Unterlassen den erzwungenen Verlust der Partie zur Folge hätte. Somit verbleiben 20 Züge, auf die der Spieler einen direkten Einfluss hat. Das gilt aber nur für durchschnittliche Spieler, die mehr Freiheit genießen.

Meisterspieler sind im Hinblick auf die Auswahl ihrer Züge noch weiter eingeschränkt. Ich möchte zur Erläuterung dieses Tatbestandes an dieser Stelle einem wirklichen Experten, nämlich Emanuel Lasker das Wort erteilen [3]:

*„Schon im Schach des gewöhnlichen Spielers zeigt es sich, dass die Auswahl der Züge durch die Forderung ihrer Zweckmäßigkeit außerordentlich eingeschränkt wird. In der Praxis der Meister steigert sich noch diese Reduktion der Möglichkeiten und je weiter das Schachspiel fortschreitet, desto weiter geht auch diese Einschränkung der Wahl.“ (Lasker 1909)*

Demzufolge unterstelle ich, dass starke GM in einer Partie durchschnittlich keine größere Auswahl als 10 Züge haben, ihr einen individuellen Stempel aufzudrücken.

### Stellungen

Die Bewertung einer Stellung funktioniert in etwa so wie das Zusammensetzen eines Puzzles, wobei zu unterscheiden ist zwischen ein-, zwei- und dreidimensionalen Puzzles.

Die Qualität einer Bewertung hängt ab zum einen von dem Schwierigkeitsgrad der Stellung, zum anderen von den Fähigkeiten, die der Spieler mitbringt. Kasparow hat dem Schachvolk die Grundlagen für seine Herangehensweise in einem Interview mitgeteilt:

*„Ich nenne Schach ein dreidimensionales Spiel, weil es mit drei verschiedenen Themen zu tun hat: Material, Zeit und Qualität. Material wird von jedem Anfänger verstanden. Du weißt, du hast einen Bauern oder eine Figur mehr oder einen Turm weniger; so begreift jeder Amateur und jeder Computer die Stellung. Doch wenn man wächst, beginnt man auch den Zeitfaktor mit einzubeziehen. "Wenn ich einen starken Angriff habe und diese Figur opfere, dann müsste mein Bauer durchlaufen." Hierbei muss man Zeit und Material miteinander vergleichen. Viele Spieler können mit diesen beiden Faktoren umgehen. Doch dann ist da noch der schwierigste Faktor, die Qualität. Da geht es um einen starken Springer, um das Läuferpaar oder die bessere Bauernstruktur; jetzt muss man drei Dimensionen unter einen Hut bringen. "Wenn ich diesen Bauern opfere und ..., dann liege ich zwei Tempi zurück, habe aber im herannahenden Endspiel die bessere Bauernstruktur, und diese Figur ist wirklich stark." ... Dies ist ein großes Problem für Computer. Computer können nicht auf solch einem vagen Gebiet wie Qualität operieren, und Qualität zu verstehen ist genauso das Hauptproblem für die meisten Schachspieler, selbst für Topspieler. Denn Qualität hat keine genaue Beziehung zu jeder Stellung. Und deshalb kann man, wenn man in drei Dimensionen spielt, in jeder Partie Momente der Inspiration genießen.“ [2]*

Die drei Dimensionen für die Schachbewertung sind also laut Kasparow:

- 1. Dimension: Material
- 2. Dimension: Zeit
- 3. Dimension: Qualität

Einfache eindimensionale Stellungen lassen sich relativ gut erfassen. Es geht oft einfach darum materielle Zusammenhänge richtig zu deuten, z.B. die Verwertung eines Mehrbauern.

Schwieriger sind schon Stellungen zu bewerten, in denen als zweite Dimension der Zeitfaktor eine Rolle spielt. Das klassische Gambitspiel, bei dem ein Spieler Kompensation in Form von Entwicklung hat, ist hier als Beispiel zu nennen. Hier ist eine fehlerfreie Stellungsbewertung schon bedeutend schwieriger.

Unter der dritten Dimension sind strukturelle Merkmale zu verstehen, wie

- Königssicherheit
- Bauernstruktur
- Zentrum und Raum
- Figurenaktivität
- Dynamisches Potential

Die dritte Dimension (Qualität) ist mit Abstand am schwierigsten zu erfassen und demzufolge ist eine 3D-Stellungsbewertung auch für viele Spieler sehr diffizil. Fehler sind auch bei sehr starken Spielern oft die logische Folge [1].

Die erste Schwierigkeit eines Spielers vor Ausführung seines Zugs besteht darin, eine klare Stellungenbeurteilung durchzuführen. Hierfür muss er das Wesen der Stellung erkennen (eins-, zwei- oder dreidimensional) und das (oder die) hervorstechende(n) Stellungenmerkmal(e) bestimmen.

Ist ihm das gelungen, so kann der logische Zug durchaus schon auf der Hand liegen. Bei einer schwierigen Stellung, die mehrere – scheinbar gleichwertige – Züge zulässt, kommt häufig noch eine subjektive Komponente hinzu, bei der der Spieler die eigenen sowie die gegnerischen Stärken und Schwächen mit einbezieht. Diese subjektive Komponente ist in Schachkreisen auch unter dem Begriff „psychologisches Schach“ bekannt. Viele Menschen verarbeiten vor der Wahl eines Zuges solche „psychologische Erwägungen“. Jedoch sind sich auch gut programmierte Computerprogramme ihrer Stärken und Schwächen „bewusst“. Ein spezielles Beispiel für diese These ist Junior, der nur äußerst ungern seine Dame tauscht, weil er „weiß“, dass seine Spielstärke darunter leidet. Sogar eine Berücksichtigung des Gegners ist bei Computerprogrammen möglich z.B. durch Einstellungen an der Engine oder auf bestimmte Gegner spezialisierte Engines (z.B. Gambit Tiger gegen Menschen).

Die letzte Schwierigkeit besteht darin, aufgrund der objektiven und subjektiven Stellungenbeurteilung den richtigen Zug herauszufinden.

### Spielstärke

Was ist das überhaupt: Spielstärke? Kann man das überhaupt in einen absoluten Wert gießen (im Gegensatz zu den vielen relativen Werten, die ständig in der Schachliteratur diskutiert werden)? Antwort: Ja, man kann!

Spielstärke wird durch zwei Einflussmöglichkeiten bestimmt:

1. Die Spielstärke einer Engine (eines Menschen) leitet sich zum einen ab von der Summe der Qualität der Entscheidungen, die er während einer Partie zu treffen hat. Hat er die Fähigkeit, in einer Schachpartie die notwendigen und guten bzw. einzigen Entscheidungen zu treffen, dann ist der Spieler garantiert ein Meister bzw. ein modernes PC-Schachprogramm. Sind einige wenige der Entscheidungen der Partie nicht ganz so hochwertig, aber auch keine groben Fehler, dann handelt es sich wahrscheinlich um einen Turnierspieler oder um ein älteres Schachprogramm auf überholter Hardware.

Neigt der Spieler dazu, grobe Fehler zu machen, sei es auch nur einen einzigen während der Partie, so hat dies eine erhebliche Spielstärkeeinbuße zur Folge. Während gelegentliche kleine Ungenauigkeiten sich auch gegen gute Gegnerschaft zuweilen ausgleichen lassen, ist das bei groben Fehlern nicht der Fall. Starke Spieler und moderne Computerprogramme bestrafen grobe Fehler i. d. R. mit einer Niederlage.

2. Spielstärke wird auch durch die Fähigkeit definiert, mehrdimensionales Schach über längere Zugfolgen fehlerfrei durchzuführen. Ich glaube, dass es Spieler mit Meisterstärke gibt, die vor allem durch die fehlerfreie Behandlung von eindimensionalen Stellungen auffallen. Jedoch spätestens auf GM-Ebene ist es damit vorbei. Spieler dieses Kalibers müssen fähig sein, längere Zugfolgen in allen drei Dimensionen relativ fehlerfrei zu arbeiten.

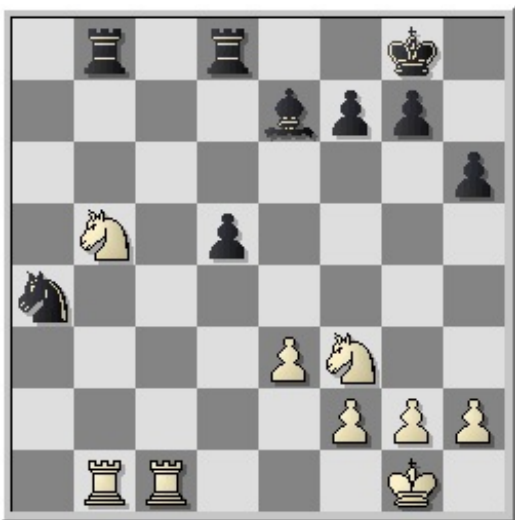
Was unterscheidet jedoch einen starken Großmeister, der quasi nur gute Züge macht, vom besten Spieler der Welt, dem Weltmeister? Ich denke, der Weltmeister (zumindest einer vom Kaliber Garri Kasparows) ist in der Lage, über eine ganze Partie immer die richtigen Züge zu finden und zwar auch in vielen aufeinander folgenden dreidimensionalen Stellungen. Es gibt viele Partien dieser Qualität, die er gespielt hat.

Manche Schachpartien, die in starken GM-Turnieren gespielt wurden, lassen sich beim Nachspielen auch von Spielern aus unteren Klassen begreifen, weil ein-dimensionales Schach gespielt wird bzw. einfach nicht viel passiert. Sogar in WM-Kämpfen gibt es mitunter ganze Serien solcher Partien, die von der beiderseitigen Absicht, keine Fehler zu begehen, geprägt sind. Hier folgt ein Beispiel aus dem 1. Wettkampf zweier der stärksten Schachspieler überhaupt:

#### Karpow – Kasparow

Moskau 1985

**1.Sf3 d5 2.d4 Sf6 3.c4 e6 4.Lg5 Le7 5.Sc3 h6 6.Lh4 0–0 7.e3 b6 8.Le2 Lb7 9.Lxf6 Lxf6 10.cxd5 exd5 11.0–0 Sd7 12.b4 c5 13.bxc5 bxc5 14.Tb1 Lc6 15.Lb5 Dc7 16.Dc2 Tfd8 17.Tfc1 Tab8 18.a4** Die Stellung ist das Ergebnis einer langen Eröffnungsvariante, bei der keiner ein großes Risiko eingeht. Eindimensionales Schach mit minimalem weißem Vorteil ist das Ergebnis. **18...Dd6 19.dxc5 Sxc5 20.Lxc6 Dxc6 21.Sb5 Le7 22.Sxa7 Da6 23.Sb5 Dxa4 24.Dxa4 Sxa4**



Die Abwicklung hat keine wesentlichen Unterschiede zur obigen Bewertung eingebracht. Die Tatsache, dass Schwarz einen Läufer hat, wird durch seine etwas schlechtere Bauernstellung mehr als aufgewogen. Somit kann Weiß gefahrlos noch etwas versuchen aber ohne wirklich Aussichten auf Erfolg zu haben. **25.Sfd4 Td7 26.Sc6 Tb6 27.Sbd4 Lf6 28.Txb6 Sxb6 29.Tb1 Sa4 30.g3 Sc5 31.Kg2 g6 32.Sf3 Td6 33.Scd4 Se6 34.Tb8+ Kg7 35.Sb3 Td7 36.Tb5 Lc3 37.Sc1 Sg5 38.Se2 Se4 39.Tb3 Tc7 40.Tb5 Td7 41.Sf4 Sf6 42.h3 h5 43.Tb3 Se4 44.Ta3 Kg8 45.Ta4 Sf6** Auch die letzten 20 Züge haben die Bewertung nicht im geringsten geändert. Entsprechend farblos und leicht verständlich plätschert die Partie dem Remis entgegen. **46.g4 hxg4 47.hxg4 Sxg4 48.Se2 d4 ½–½**

Partie online nachspielen

Repräsentiert diese Partie nun die volle Spielstärke der beiden Kontrahenten? Mitnichten, diese Partie hätte auch von zwei

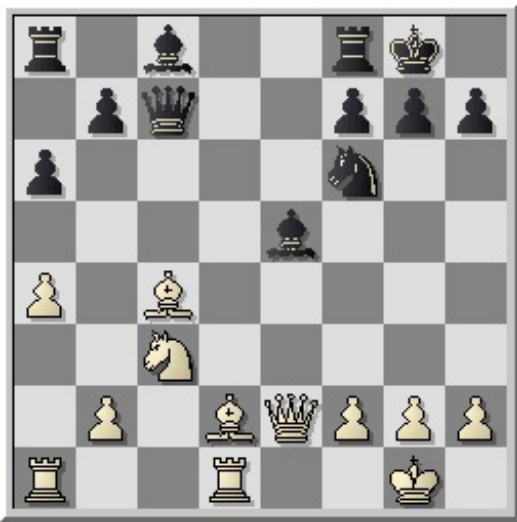
deutschen Oberligaspielern auf das Brett gebracht werden können, ohne dass einer besonders stolz auf seine Leistung hätte sein können. Die Partie ist qualitativ einfach nicht hochwertig genug, um eine Aussage über die Spielstärke der Spieler ableiten zu können. Sie haben sich nicht gefordert, ähnlich wie zwei Tennisspieler, die sich die Bälle von der Grundlinie gegenseitig rüberschaufeln. Zu diesem Zeitpunkt hatte es für die Gegner einfach absolute Priorität, nicht zu verlieren. Da Spieler dieses Kalibers sich in eindimensionalen Stellungen quasi fehlerfrei bewegen können, haben sie diese auch angestrebt und fehlerfrei zum Remis geführt. Demzufolge hat sich über die ganze Partie die Stellungsbewertung nicht geändert: Immer hatte Weiß minimalen Vorteil!

Partien solcher Qualität finden sich auch häufig bei Computerturnieren, z.B. in der nachfolgenden WM-Partie:

### Deep Sjeng - Shredder

World Computer Chess Championship (4), 2005

1.d4 d5 2.c4 dxc4 3.Sf3 Sf6 4.e3 e6 5.Lxc4 a6 6.0-0 c5 7.a4 Sc6 8.De2 Dc7 9.Sc3 Ld6 10.Td1 0-0 11.Ld2 cxd4 12.exd4 e5 13.dxe5 Sxe5 14.Sxe5 Lxe5



Nachdem drei Zentrumlinien geöffnet wurden, verbleibt eine qualitativ eher gering anzusetzende Stellung. 15.h3 Te8 16.Le3 Le6 17.Lxe6 Txe6 18.Tac1 Tc6 19.Ld4 Lxd4 20.Txd4 Te8 21.Dd2 Db6 22.Kh1 h6 23.Td1 Db3 24.Td3 Db6 25.Td4 Tce6 26.b4 Dc6 27.Tc1 Db6 28.Td1 Te5 29.Kg1 Dc6 30.Tc1 T5e6 31.a5 b5 32.Sd5 Dxc1+ 33.Dxc1 Te1+ 34.Dxe1 Txe1+ 35.Kh2 Sxd5 36.Txd5 Te4 37.Td6 Txb4 38.Txa6 Ta4 39.g4 Ta3 40.Ta8+ Kh7 41.h4 g5 42.h5 Kg7 43.a6 b4 44.Kg2 b3 45.a7 Ta4 46.Kf3 b2 47.Tb8 Txa7 48.Txb2 f5 49.gxf5 Kf6 50.Tb5 Ta3+ 1/2-1/2

Partie online nachspielen

Dass Kasparow und Karpow auch in der Lage waren, anders miteinander umzugehen, zeigt wohl am besten eine Partie aus dem „Wimbledon des Schachs“, Linares 1993:

### Karpow – Kasparow

Linares, 1993

1.d4 Sf6 2.c4 g6 3.Sc3 Lg7 4.e4 d6 5.f3 0-0 6.Le3 e5 7.Sge2 c6 8.Dd2 Sbd7 9.Td1 a6 10.dxe5?! Karpow begeht hier einen psychologischen Fehler, indem er Kasparow ein aktives Figurenspiel erlaubt. 10...Sxe5 11.b3 b5 12.cxb5 axb5 13.Dxd6 Sfd7 14.f4?! Wieder unterschätzt Karpow die schachlichen Fähigkeiten des Gegners. 14...b4! 15.Sb1 Sg4 16.Ld4 Lxd4 17.Dxd4 Txa2 18.h3 c5 19.Dg1 Sgf6 20.e5 Se4 21.h4 c4 22.Sc1



Ein bemerkenswertes Bild! Durch die Ausschöpfung aller seiner rechnerischen und intuitiven Fähigkeiten in drei Dimensionen ist es Kasparow gelungen die Figuren seines Widersachers komplett auf die erste Reihe zurückzudrängen. Die Verwertung des Vorteils erfolgt ganz im Stil Kasparows: dynamisch! 22...c3!? 23.Sxa2 c2 24.Dd4 cxd1D+ 25.Kxd1

**Sdc5 26.Dxd8 Txd8+ 27.Kc2 Sf2 0–1**

Partie online nachspielen

Worin besteht der Unterschied zu der oben gezeigten Partie der beiden Kontrahenten? Kasparow muss diese Partie sehr genossen haben, er badet geradezu in der von ihm beschworenen 3D-Inspiration. Allerdings musste Karpow ihm hierzu vor allem durch psychologisch ungenaue Züge erst die Möglichkeit geben.

Auch hier sei für die Spielweise der Computer ein Beispiel aus der vorletzten WM gestattet:

### Zappa - Deep Junior

World Computer Chess Championship (7), 2005

**1.d4 d5 2.Sf3 Sf6 3.c4 e6 4.Sc3 c6 5.e3 Sbd7 6.Dc2 Ld6 7.g4!?** Ein Zug, der mehrdimensionales Schach garantiert!  
**7...h6 8.Tg1 e5 9.cxd5 cxd5 10.g5 hxg5 11.Sxg5 e4 12.Sb5 Sb6 13.Ld2 Lf5 14.h3 Lh2 15.Tg2 Tc8 16.Db3 Lb8 17.Lb4!**



Zappa und Junior haben eine Stellung geschaffen, die alles andere als eindimensional ist, vielmehr viele Qualitätsmerkmale enthält. Jedoch gilt es zu erkennen, was das hervorstechende Merkmal der Stellung ist. Dies ist Zappa gelungen, denn mit seinem letzten Zug hat er die Sicherheit des schwarzen Königs ins Visier genommen. **17...Sh7 18.Sxh7 Txx7 19.Tg1 Tc6 20.Sc3 a6 21.a4 Sc4 22.Lxc4 dxc4 23.Dc2 Tch6 24.De2 Dc8 25.0–0–0!** Interessanterweise findet der weiße König einen sicheren Platz am Damenflügel, obwohl der erste Blick das nicht offenbart. **25...b5 26.d5 Db7 27.d6 Ld7 28.axb5 axb5 29.Tg5 f5 (29...Te6) 30.Td5 Da8 (30...Da6 31.Kb1 Txx3 [31...Lxd6 32.Dd1 Lxb4 33.Txd7]) 31.Kb1 g6? (31...Txx3 32.Te5+ Kd8 33.Dd1 T3h6 34.Tg1±) 32.Te5+ Kf7 33.Sd5 Kf8 34.Sf6 Tf7 35.Te8+ Kg7 36.Sh5+ Kh7 1–0**

Partie online nachspielen

### Stellungstests

WM-Test-Autor Michael Gurevich ist sehr streng hinsichtlich der Auswahl von Stellungen, die für die Aufnahme in einen Stellungstest in Frage kommen. Er stellt folgende Anforderungen:

Die Teststellung soll:

- einen einzigen besten Zug enthalten (der beste Zug bedeutet entweder den Gewinnzug in der ansonsten ausgeglichenen Stellung oder eine einzige Rettung in schlechter Position)
- einen stabilen Charakter der Engine-Lösung aufweisen (die Testergebnisse sollen demonstrieren, dass der gefundene beste Zug sicher bestimmte Zeit lang beibehalten bleibt)
- keine zufällige Lösungen oder die Lösung aus dem falschen Grund zulassen (es ist nicht gut, wenn eine Engine ahnungslos die Lösung zeigt, ohne die richtige HV-Anzeige zu liefern).
- die Lösung sollte nicht zu einfach und nicht zu schwer, also lösbar sein
- die Lösung sollte diverse thematische Charaktere demonstrieren; man braucht nämlich vielseitige Teststellungen aus allen Partiephasen
- keine spekulativen Lösungen mit Opfern enthalten.

Ich möchte diese Ausführungen noch um mindestens einen wichtigen Punkt ergänzen: In der vorliegenden Stellung sollte noch ein alternativer möglicher Zug existieren, der das Programm eventuell auf eine falsche Fährte führt. D.h. die Teststellung sollte einen leichten „Fallencharakter“ besitzen, um andere (falsche) Lösungen attraktiv zu machen.

Als Nächstes möchte ich mich mit der Fragestellung beschäftigen, wie oft Positionen vorkommen, die den oben angegebenen Anforderungen entsprechen. Jeder wird sofort einsehen, dass es eher unwahrscheinlich ist, dass der durchschnittliche Spieler im Verlauf seiner Entscheidungen diese Bedingungen erfüllt.

Gurevich findet nur vereinzelt entsprechende Stellungen. Ich gehe davon aus, dass die Wahrscheinlichkeit in einer Partie eine Teststellung zu entdecken, umso größer ist, je höher die Qualität der Partie ist, da in qualitativ minderwertigen Partien wie der oben angegebenen Partie Karpow-Kasparow (Moskau 1985) sehr oft mehrere in etwa gleichwertige Halbzüge möglich sind. Zudem findet man kaum einmal den von M. Gurevich geforderten „thematischen Charakter“. Ich möchte an dieser Stelle allerdings nicht versäumen, dass es Stellungstests gibt, wo es in erster Linie darum geht, den falschen Zug zu vermeiden und dass dabei auch Stellungen vorkommen, wo mehrere Lösungen zulässig sind.

Offenbar findet sich nur alle paar Partien mal wirklich eine Teststellung, die den genannten Kriterien entspricht. Um es zu quantifizieren behaupte ich jetzt, dass alle 10 Partien eine Teststellung vorkommt. Vielleicht ist es daher sinnvoll, dass Testautor M. Gurevich vor allem Partien von Weltmeistern auf Teststellungen durchkämmt, da es hier aufgrund der anzunehmenden Spielstärke wahrscheinlicher ist, dass sich eine finden lässt.

Dass es jedoch nicht unbedingt eines Weltmeisters bzw. eines starken Spielers bedarf, eine Teststellung zu produzieren, wollte ich selber mal beweisen. So überraschte ich M. Gurevich im April 2002 mit dem Hinweis, eine Stellung aus meiner letzten Partie dahingehend zu überprüfen, ob sie seinen Kriterien entspricht. Ich selber war jedenfalls felsenfest der Ansicht war, dass mir die „Produktion“ einer Teststellung gelungen war.

In der nachfolgenden Partie hatte ich interessanterweise schon vor Ausführung des 29. weißen Zugs das Gefühl, dass ich hier eine kritische Stellung vorliegen hatte. Mit 15 min auf der Uhr blieb mir auch keine große Zeit zum Rechnen, daher erfolgte meine Reaktion nach einer Minute intuitiv:

#### Vossen - Olbricht

Eschweiler, 2002

**1.d4 d5 2.c4 dxc4 3.Sf3 Sf6 4.g3 e6 5.Lg2 Lb4+ 6.Sbd2 Sc6 7.0-0 c3 8.Sc4 cxb2 9.Lxb2 0-0 10.Dc2 Te8 11.Tad1 Lf8 12.Sfe5 Sb4 13.Db3 Sbd5 14.e4 Sb6 15.Se3 c6 16.f4 De7 17.a3 g6 18.g4 Sfd7 19.f5 Lg7 20.fxc6 Sxe5 21.gxh7+ Kxh7 22.dxe5 Kg8 23.Tf2 Ld7 24.Tdf1 Tf8 25.a4 c5 26.a5 Sa4 27.La1 b5 28.axb6 Sxb6**



## 29. Sf5! Welches Programm findet diesen Zug? 29...exf5

Auch Ablehnen des Opfers war möglich: [29...Dg5!? 30.h4! Dxg4 31.Tf4+- Dh5 (31...exf5 32.Txg4 fxc4 33.e6 Lxe6 (33...fxe6 34.Lxg7 Kxg7 35.Dc3+ Kg6 36.Ta1! Tf7 37.Dxc5 Tg7 38.Ta5+- ) 34.De3 f6 35.Dxc5 Sd7 36.Dd6+- ) 32.Dg3 exf5 33.exf5 Kh8 (33...Kh7 34.Tg4 ) 34.Tg4 ]

**30. exf5 Lxe5** [30...Dg5 31.h4 Dxh4 (31...Dxg4 32.Tf4 De2 (32...Dh5 33.Dg3 Kh8 (33...Kh7 34.Tg4 (34.Lf3 Dh6 35.f6 Tg8 36.Le4+ Kh8 37.fxc4+ Dxg7 ) 34...Tg8 35.Lxa8 Dxg4 36.Dxg4 Lxe5 37.Dxg8+ Kxg8 38.Lxe5 Sxa8 ) 34.Tg4 Dh6 (34...Tg8 35.Lxa8 Dxg4 36.Dxg4 Lf6 37.Lf3 ) 35.Lxa8 Txa8 36.e6 Lxa1 37.Txa1 Le8 38.Tg8+ Kh7 39.Ta2! ) 33.Dg3 Kh7 34.Te1 Dd2 35.Lc3 Da2 36.Tg4 Tg8 37.Lxa8 Lxf5 38.Tg5 ) 32.Lxa8 Txa8 33.Th2 Dg5 34.Dh3+- Lb5 (34...Sd5 35.e6 De3+ 36.Dxe3 Sxe3 37.exd7 Sxf1 38.Te2 ; 34...Kf8 35.f6 ) 35.Te1 f6 36.Lc3 fxe5 37.Ld2 Dd8 38.Dh7+ Kf7 39.g5 ; 30...Dh4! 31.g5! (31.Lxa8 Txa8 (31...La4 32.De3 Txa8 33.e6 Lxa1 34.Txa1 Tf8 35.exf7+ Txf7 36.Txa4 Sxa4 37.De8+ Kg7 38.Dxa4 ) 32.e6 (32.g5 Lc6 33.g6 Dg4+ 34.Dg3 Dxg3+ 35.hxc3 ; 32.Df3 Tc8 33.e6 Lc6 34.De2 Lxa1 (34...f6 35.e7 Kh7 ) 35.Txa1 Tc7 36.f6 ; 32.Tg2 Lc6 33.Tg3 ) 32...Lxa1 (32...Lc6 33.exf7+ Kh7 (33...Kh8 34.Lxg7+ Kxg7 35.f6+ Kg6 36.h3 c4 37.Da3 Tf8 38.Tf5 Dh6 39.T1f2 ) 34.Tf4 Sd5 (34...Lxa1 ; 34...Ld5 35.Dc2 ) ) 33.Txa1 fxe6 (33...Lc6 34.exf7+ Kf8 35.h3 ) 34.fxe6 Lc6 35.h3 c4 ; 31.e6 La4 32.De3 Lxa1 ; 31.Lxa8 ) 31...c4 (31...La4? 32.Db1 Dxg5 33.f6 Lh6 34.Tf3 Kh8 35.Tg3 Dh5 36.Tf5 Dh4 37.Th3 Dxh3 38.Lxh3+- ; 31...Tab8!? 32.f6 Lh8 (32...c4? 33.De3 Lh8 34.g6 Le6 (34...Tfe8 35.Tf3+- ) 35.Le4 Lh3 36.gxf7+ Txf7 37.e6+- ) 33.Tf4+- c4 34.De3+- ; 31...Dg4? 32.Tf4+- c4 (32...Dxg5 33.T4f3 c4 34.Da3+- Lxf5 35.Tg3 Lg4 (35...Dh5 36.e6 f6 37.e7 Tfe8 38.Lxf6 ) 36.h3 ; 32...Dh5 33.Dg3! ) 33.Da3 Dh5 (33...Dxg5 34.T4f3 Lxf5 35.Tg3 ) 34.f6 Lh8 35.g6 Dxg6 36.T1f3 ; 31...Dxg5? 32.Tf3+- c4 33.Da3 Lxf5 34.Tg3 Dh5 35.Lf3 Dh7 36.e6 f6 37.Lxa8 Sxa8 38.e7 Te8 39.Dc5 ; 31...Tad8! 32.f6 Le6 (32...Lh8 33.Tf4!+- ; 32...c4 33.De3 Lh8 34.Tf4+- Dh5 (34...Dh7 35.De1 ) 35.Dg3 ; 32...Lxf6 33.Txf6+- ) 33.Dc2 Lh8 34.g6 c4 (34...Tfe8 35.Le4 Sc4 36.Tg2 Kf8 37.gxf7 Lxf7 38.Lg6 Td7 39.Lxf7 Txf7 40.Tg3+- ; 34...Td7!? 35.Lc6 Tc7 36.Tg2 fxc6 37.Dxg6+ Tg7 38.fxc7 Txf1+ 39.Kxf1 Lc4+ 40.Te2 Lxe2+ 41.Kxe2 Dc4+ 42.Kf2 Lxg7 43.e6 Df4+ 44.Lf3 Df8 45.Ke2 Sc8 46.Lb2 Se7 47.Dg5 a5 48.Le5+- ; 34...Sc4 35.Dc1 Sxe5 36.Lxe5 fxc6 37.Dxc5+- ) 35.Le4+- Sd5 (35...Lh3 36.Tg2! Lxg2 37.Dxg2 Dh6 38.e6 fxc6 39.e7 Kh7 40.exf8S+ Txf8 41.Tf4 ) 36.Tf3 Sb4 37.gxf7+ Txf7 38.Dg2+ Kf8 39.Lf5 Lxf5 40.Txf5 Te8 41.Tg5 Dh7 42.e6 Txe6 43.Tg7 ) 32.De3 Tac8 33.f6 Lh8 34.g6 Le6 35.g7 Lxg7 36.fxc7 Kxg7 37.Tf3 Th8 38.Txf7+ Kg8 39.T7f6 Sd5 40.Tg6+ Kh7 41.Lxd5 Kxg6 42.Tf6+ Dxf6 43.exf6 Lxd5 44.Dg3+ Kh6 45.Df4+ Kg6 46.f7 ; 30...c4 31.Dg3+- Dg5 32.h4 Dh6 33.g5 ; 30...La4 31.Dg3+- f6 (31...Dg5 32.h4 Dh6 33.g5 ) 32.g5]

**31. f6 Dd6?** Verliert schnell! [Viel stärker war 31...De6! und der Gewinn für Weiß ist noch weit weg 32.Dh3! Lxa1 (32...Dxg4 33.Dxg4+ Lxg4 34.Lxa8 kommt auf das selbe hinaus!) 33.Txa1 Dxg4 34.Dxg4+ Lxg4 35.Lxa8 Txa8 36.Tg2+/- Kf8 37.Txg4 Sd7 38.Ta6 Te8 39.Td6 Se5 40.Tg3 a5 41.h4+- Schwarz ist machtlos gegen den Vormarsch des h-Bauern!]

**32. De3!** Gegen die Doppeldrohung ist kein Kraut gewachsen.

**32...Lxh2+ 33. Kh1 Lf4 34. Txf4 Lc6 35. T4f3 Lxf3 36. Dg5+ 1-0**

Partie online nachspielen

Leider wurde ich von M. Gurevich nicht in eine seiner Testgalerien aufgenommen. Beim Weltmeister-Test habe ich das ja noch glatt eingesehen ...

Wie sah es aber mit der grundsätzlichen Tauglichkeit der Stellung nach dem 28. Zug von Schwarz als Teststellung aus? Nun, nach der Partie habe ich mich selber intensiv um die Stellung gekümmert. Ich denke, ich konnte zweifelsfrei nachweisen, dass der Zug eindeutig der beste der gegebenen Stellung ist (Ich bitte alle Andersdenkenden mich analytisch eines anderen zu belehren!).

Zufällige Lösungen sind garantiert hier auch nicht möglich, ebenso ist der thematische Charakter des Schlüsselszuges (Angriffopfer) sehr offensichtlich. Offenbar scheitert die Stellung daran, dass sie zu schwierig ist, als dass sie ein damaliges Programm (2002) hätte lösen können. Aber disqualifiziert sie das prinzipiell als Teststellung? Ich denke nicht! Die Zeit spielt für die Stellung, aber offenbar schafft die heutige Generation der Computer es immer noch nicht, oder?

Es ist nicht so, dass ich in einer meiner Partien nicht sonst schon mal den Eindruck hatte, dass die jeweils vorliegende Stellung als Teststellung geeignet ist. Dieses Gefühl beschleicht mich aber wirklich nicht jede Partie, eher dann wenn ich gerade dabei bin, jemandem eine Falle zu stellen. Bei meinen simplen Taktikfallen lasse ich den Gedanken natürlich

meistens schnell wieder fallen (Ausnahme: Stellung aus Vossen - Olbricht), anders bei Positionsfallen. Dennoch, die Wahrscheinlichkeit, dass eine solche Falle die oben formulierten Kriterien erfüllt, ist eher gering.



### Stellungstests und Spielstärke

Gurevich betont bei dem Thema die Bedeutung der Ausgewogenheit eines Stellungstests. Die Kriterien, die eine Testsuite seiner Meinung enthalten sollte, sind:

- diverse Partiephasen (Königsangriff, Positions- und Endspiel)
- leichtere und schwierigere Aufgaben
- ausreichende Stellungszahl

Die beiden Kernfragen lauten:

*Ist es möglich, eine direkte Beziehung der Fähigkeit eines Programms, Stellungstests zu lösen, auf die gesamte Spielstärke herzuweisen?*

Ich denke, diese Fragestellung kann gewiss mit einem klaren Ja beantwortet werden, denn Teststellungen kommen – wie mehrfach nachgewiesen wurde – ab und zu in praktischen Partien vor. Jede Einzel-Verbesserung des praktischen Spiels hebt bekanntermaßen die Gesamtspielstärke an.

Zudem haben Teststellungen einen besonderen Wert, da sie eindeutig zu lösen sind im Gegensatz zu vielen anderen Stellungen.

*Wie groß ist der Einfluss von Stellungstests auf die Spielstärke eines Programms?*

Bei der Beantwortung dieser Fragestellung möchte ich auf die Erkenntnisse aus dem Kapitel „Eingrenzung der Zugzahl“ zurückgreifen. Die Anzahl der Züge, auf die der Spieler individuellen Einfluss hat, habe ich mit durchschnittlich 10 pro Partie beziffert. Dagegen habe ich abgeschätzt, dass ungefähr alle 10 Partien (wenn überhaupt) mal eine Teststellung, die allen strengen Kriterien genügt, vorkommt. Das bedeutet, dass bei ca. 100 kritischen Stellungen eine dabei ist, die Teststellungscharakter besitzt.

Wie ich oben schon ausführte, ist eine Teststellung zweifelsohne von besonderem Wert, und die Fähigkeit, sie zu lösen, deutet gewiss auf eine hohe Spielstärke hin. Aber die Seltenheit von Teststellungen (ca. 1 %) gibt doch zu denken.

Wie in den vorangegangenen Kapiteln ausgeführt, existieren viel mehr Stellungen, die überhaupt keinen Stellungstestcharakter haben, als solche mit. Diese „Nicht-Test-Stellungen“ sind aber dennoch von großer Bedeutung für den Ausgang der Partie. Häufig haben sie keine eindeutigen Lösungen, sind nicht „schön“ und haben ausgesprochen defensiven, ja destruktiven Charakter. Bei der Auffindung von guten Zügen in diesen Stellungen ist es sogar oft nicht wichtig, „den richtigen Zug“ zu finden, sondern keinen Fehler zu begehen. In zweiter Linie kann es auch bedeutungsvoll sein, den (unter Berücksichtigung des Gegners) „psychologisch“ richtigen Zug zu machen.

Meiner Meinung nach bestimmen diese Aspekte die Spielstärke eines Schachspielers viel mehr als die Fähigkeit, eine Teststellung lösen zu können. Das ist m. E. der Grund dafür, dass Teststellungen zwar eine Aussage über die Spielstärke einer Engine zulassen, diese aber nicht zu groß ist!

### Fazit

Gibt es einen – womöglich sogar proportionalen – Spielstärkezuwachs eines Schach-Computerprogramms gegenüber einem anderen, wenn es in der Lage ist mehr Stellungstests zu lösen? Diese Fragestellung hat mich bei der Erstellung des Artikels intensiv beschäftigt. Ich möchte daher versuchen, eine Antwort zu geben.

Ich bin der Meinung, dass die Fähigkeit einer Engine, Stellungstests zu lösen, durchaus einen Einfluss auf die Spielstärke hat. Es ist aber nicht so, dass es möglich wäre, eine Proportionalität von der Anzahl der gelösten Stellung auf die Höhe der Spielstärkesteigerung herzustellen. Wie ich ausführlich dargelegt habe, ist der Begriff „Spielstärke“ grob gesagt eine eigentümliche Kombination aus der Fähigkeit Fehler in höchst unterschiedlichen Stellungen zu vermeiden und die eigene Spielqualität zu verbessern. Hierbei fällt auf, dass mehr als 99 Prozent aller Stellungen für Tests nicht geeignet sind. Ich bin daher geneigt, die Fähigkeit eines Programms, viele Stellungstests zu lösen, als eine Art qualitatives Merkmal zu klassifizieren, ähnlich wie die Fähigkeit, bei Stellungen mit Läuferpaar dessen Vorzüge zu betonen oder die Beherrschung des Spiels mit einem Isolani.

Natürlich hat eine solche qualitative Aufwertung eine Spielstärkesteigerung zur Folge, es verbleibt jedoch die Frage nach der Basis-Spielstärke. Im Umkehrschluss bin ich der Überzeugung, dass es möglich ist, Engines zu programmieren, die ausgezeichnet Stellungstests lösen können, deren grundsätzliche Spielstärke jedoch eher gering ist. Das bedeutet aber auch, dass es Engines geben muss, die nicht unbedingt zu den besten Testlösern gehören, jedoch eine enorme praktische Spielstärke aufweisen. *(Peter Vossen)*

Partien online nachspielen

Literatur:

[1] Wahls, Matthias: Die drei Dimensionen des Schachs; Lehrgangsunterlagen, Berlin 2003

[2] Kasparow, Gari: Kasparow's winning moves; Interview mit Brian Redhead, 1993

[3] Lasker, Emanuel: Kampf; Nachdruck der Erstausgabe von 1907, Verlag für Berlin-Brandenburg, Potsdam 2001

---