

STATISTICKÁ FYZIKA A BARVENÍ GRAFŮ

LENKA ZDEBOROVÁ^{*)}

Fyzikální ústav AV ČR, v.v.i., Na Slovance 2, 182 21 Praha 8
LPTMS, UMR 8626 CNRS et Université Paris-Sud,
91405 Orsay CEDEX, Francie

Článek popisuje analogii mezi statistickou fyzikou skelných systémů a řešením složitých optimalizačních úloh. Příkladem zde je NP-úplný problém barvení grafů, který podrobně představíme. Diskutujeme chování sklených systémů spolu s modely, které se pro jeho popis používají. A konečně ukážeme, jak uvedená analogie vedla v hlubšímu pochopení a popisu struktury optimalizačních úloh a zejména k vývoji nových efektivních heuristických algoritmů.

Úvod

V tomto článku přiblížíme výsledky, které vznikly na pomezí statistické fyziky, kombinatoriky a algoritmické složitosti. Popíšeme, jakým způsobem jsou metody statistické fyziky užitečné pro porozumění problémům na první pohled fyzice velmi vzdálených. Příkladem pro tento článek nám budou barvení grafů. Uvedené výsledky však platí pro mnohem obecnější třídu problémů, zejména pak pro splnitelnost logických formulí (SAT). Z pohledu fyziky se zde setkáme se skelným chováním antiferomagnetického Pottsova modelu.

Na začátku budeme definovat barvení grafů a projdeme krátkou exkurzí o výpočetní složitosti. Dále uvedeme několik praktických aplikací barvení grafů a zamyslíme se nad algoritmy, které se v těchto aplikacích používají. V další části nastíníme otázky a problémy týkající se obarvování velkých náhodných grafů. Poté představíme, jak se na věc dívá statistická fyzika. Odbočíme k fenomenologickému popisu skel a poté vysvětlíme, co je na barvení grafů skelného. Tím se dostaneme do poněkud techničtější části článku, kde objasníme, jakým způsobem barvení grafů popsat a vyřešit. A to nejprve postupem jednoduchým, který se ukáže nesprávný v intervalu parametrů, v němž se chování barvení podobá reálným sklům, poté ukážeme, jak analogie se skly umožní popsat barvení grafů i pro tento interval parametrů. V předposlední části pak představíme algoritmus, který pochopení struktury problému přineslo a který je dnes nejlepším známým pro barvení náhodných grafů. V závěru krátce shrneme poznatky a otevřené otázky.

DEFINICE BARVENÍ GRAFŮ

Grafem máme na mysli množinu bodů (vrcholů) a hran spojujících některé dvojice vrcholů. Cílem je obarvit body daným počtem barev tak, aby žádné dva body spojené hranou neměly stejnou barvu. Známým příkladem tohoto problému je obarvení

politické mapy takovým způsobem, aby dva sousedící státy neměly stejnou barvu. Body v tomto případě odpovídají státům a hrany spojují sousední země. Každou mapu je možné obarvit čtyřmi barvami, matematický důkaz tohoto elementárního tvrzení, teorému čtyř barev, je ovšem velmi složitý a je známý jen od roku 1977 [1].

Uvědomme si, že mapy jsou velmi speciálním příkladem grafů, nazývají se grafy rovinné, jelikož se dají nakreslit do roviny, aniž by se hrany musely křížit. Nejjednodušším příkladem nerovinného grafu je pentagram, mající pět vrcholů a hrany mezi každou dvojicí. Pentagram je také nejjednodušším grafem, na jehož obarvení potřebujeme více než čtyři barvy.

O VÝPOČETNÍ SLOŽITOSTI

Nabízí se obecná otázka: Je možné obarvit daný graf daným počtem barev? Pokud bylo tak těžké dokázat teorém čtyř barev, nepřekvapí nás, že odpovědět na tuto otázku bude také těžké, ale jak moc těžké? Čímž se dostáváme k teorii složitosti [2]. Zjednodušeně řečeno, existuje třída tzv. NP problémů, jež mají následující vlastnost: Pokud mudrc navrhne řešení (v našem případě obarvení), my umíme v polynomiálním čase zkontrolovat, zda onen návrh řešením skutečně je. Polynomiálním časem máme na mysli, že čas, který pro kontrolu potřebujeme, neroste rychleji než polynomiálně s velikostí systému (velikostí máme na mysli počet bodů v grafu). Barvení grafů mezi NP problémy jistě patří; je-li počet bodů v grafu N , kontrola správnosti obarvení nám zabere jistě méně než N^2 kroků.

Potřebujeme mudrce, nebo umíme také v polynomiálním čase řešení najít? Existují tzv. P problémy, pro které je odpověď ano. Nejzajímavější ze třídy NP jsou tzv. NP-úplné problémy. Pokud bychom v polynomiálním čase uměli vyřešit libovolný z NP-úplných problémů, uměli bychom to pro celou třídu NP. Největší otevřenou otázkou teorie složitosti je, zda takový polynomiální algoritmus pro některý z NP-úplných problémů existuje. Na

^{*)} E-mail: zdeborov@lptms.u-psud.fr

toho, kdo dokáže, že ano či ne, či že takové tvrzení je nedokazatelné, čeká nebetyčná sláva, a zůstaneme-li při zemi, ceny v hodnotě několika milionů dolarů.

Barvení grafů je NP-úplným problémem. Chceme-li najít obarvení jakkoliv nezbedného grafu, nemáme o mnoho lepší způsob než vyzkoušet všechny možnosti. Pro graf o N bodech a počet barev q je takových možností q^N . O něco lepší způsoby tedy máme, ale v žádném známém se nezbavíme velikosti grafu v exponentu funkce, která udává potřebný čas. To je poněkud mrzutá situace. Proč? Jelikož rychle obarvit graf daným počtem barev je potřeba v mnoha praktických aplikacích. Uvedme několik z nich.

KDE SE BARVENÍ POUŽÍVÁ?

Přemýšleli jste někdy ve svých školních letech, jak se dělá rozvrh hodin? V praxi se zřejmě nejváženější profesor rozhodne, kdy chce kterou třídu vyučovat, poté druhý nejváženější a tak dále. Pokud nezbyvá místo, zbude na mladého kolegu odpolední vyučování a může se stát, že některé třídy budou muset mít nesmyslné hodinové přestávky v průběhu dne. Zkusme se na problém rozvrhu hodin podívat trochu systematictěji. Jednotlivé vyučovací hodiny (přiřazené třídám) jsou body grafu. Hrany jsou mezi těmi, které jsou vyučovány jedním profesorem, a těmi, na kterých se překrývají některé (možná všichni) žáci. Minimální počet barev, kterými lze takový graf obarvit, je nezbytný počet hodin, ve kterých ve škole (či na univerzitě) pobíhá vyučování. Na rozvrh hodin můžeme samozřejmě klást další požadavky, ale pro náš účel to nedělejme.

Podobným způsobem můžeme přemýšlet o přidělování terminálů na letišti. Letadla jsou body v grafu. Hrany jsou mezi letadly, které se na letišti vyskytují v překrývající se dobu. Čísla terminálů odpovídají barvám, pomocí kterých chceme daný graf obarvit.

A do třetice příklad pro programátory. Při kompilaci kódu je potřeba používaným proměnným přiřadit registr. Pokud potřebujeme znát hodnotu dvou proměnných ve stejný čas, nemůžeme jim přiřadit stejný registr, nýbrž dva různé. A opět poznáváme barvení grafů, proměnné jsou tentokrát vrcholy a registry odpovídají jednotlivým barvám.

HEURISTICKÝ PŘÍSTUP A NÁHODNÉ GRAFY

Zdá se, že jsme se dostali do slepé uličky. V praxi potřebujeme najít obarvení velkých grafů, ale nejlepšimu algoritmu, který garantuje nalezení řešení, takový úkol zabere čas exponenciální ve velikosti grafu.

Existuje ale úniková cesta, zásadní je ono slůvko „garantuje“. Navrháme tzv. heuristický algoritmus, který nalezení řešení ve všech případech negarantuje, ale přesto ve velké většině praktických případů řešení najde v čase polynomiálním. Na tomto principu skutečně pracuje většina praktických aplikací. Je tedy zásadní takové heuristické

algoritmy studovat, vyvinout nejlepší možné a kvantifikovat, jak dobře fungují.

Přístup, jak s takovým studiem začít, se nabízí: Vyberme soubor grafů a zjistíme, kolik procent z nich umíme daným heuristickým algoritmem obarvit daným počtem barev v daném čase. Ty z grafů, které se nám obarvit nepodařilo, buď obarvit nejdou, anebo je algoritmus, který jsme použili, příliš pomalý nebo nešikovný. Celkový výsledek bude také samozřejmě záviset na souboru grafů, který jsme vybrali. Obecně nebude výjimečné, že i porovnání dvou různých algoritmů bude na souboru grafů záviset. Studovat praktické příklady souborů grafů je stále příliš složité, v tomto článku se zaměříme na soubor grafů náhodných [3].

Náhodný graf o N vrcholech vznikne tak, že mezi každé dva vrcholy umístíme hranu s pravděpodobností p . Střední počet hran jdoucích z jednoho vrcholu je pak $c = p(N - 1)$. K dalším vlastnostem takových grafů, zejména v limitě $N \rightarrow \infty$, se dostaneme.

ÚVOD DO BARVENÍ VELKÝCH NÁHODNÝCH GRAFŮ

Studium problémů na náhodných grafech je často prvním krokem k jejich plnému pochopení na grafech obecných. Uvidíme, že v případě barvení grafů je již tento první krok netriviální a otevírá cestu k mnoha neočekávaným spojitostem a výsledkům.

Fixujme počet barev q , který máme k dispozici, uvažujme počet barev mnohem menší než počet vrcholů $q \ll N$. Intuitivně, pokud je pravděpodobnost p velmi malá, v grafu nejsou skoro žádné hrany a bude jistě jednoduché najít správné obarvení. Na druhou stranu, pokud je pravděpodobnost blízká jedné, téměř jistě vytvoříme graf q barvami neobarvitelný. Pro přesnější popis spočítáme, kolik správných obarvení lze v průměru najít, tj. součin celkového počtu obarvení krát pravděpodobnost, že náhodné obarvení je správné.

$$\mathcal{N} = q^N (1 - 1/q)^{pN(N-1)/2} = e^{N[\log q - (c/2)\log(1-1/q)]} \quad (1)$$

Pravděpodobnost, že náhodný graf lze obarvit q barvami, je menší nebo rovna průměrnému počtu obarvení pro danou velikost grafu N a konektivitu c . Pokud je člen v hranatých závorkách v exponentu v rovnici (1) záporný, je střední počet obarvení téměř nulový pro velké grafy $N \rightarrow \infty$, a tedy i pravděpodobnost, že obarvení existuje, je skoro jistě nulová, pokud je střední konektivita grafu

$$c > \frac{2 \log q}{\log(1-1/q)} \geq c_q \quad (2)$$

Čímž jsme spočítali horní mez pro obarvitelnost náhodných grafů q barvami.

Dolní mez, tedy konektivitu takovou, že pro menší skoro jistě obarvení q barvami existuje, dostaneme například následujícím postupem: Odstraňme iterativně všechny vrcholy (a k nim spojené hrany), které mají stupeň menší než q . To, co z gra-

fu zbude, se nazývá q -jádro. Pokud se nám podaří obarvit toto q -jádro, obarvení odstraněných částí zkonstruujeme iterativně. q -jádro náhodných grafů je objekt v teorii grafů velmi dobře známý [4]. Bylo dokázáno, že velikost q -jádra je buď nulová, nebo pokrývá rovnou konečný zlomek ze všech vrcholů grafu. Kupříkladu konektivita, při které se objeví 3-jádro, je asi 3,35.

Chování, které jsme popsali, spolu s numerickými simulacemi vypovídá o existenci kritické konektivity c_q takové, že velké náhodné grafy s konektivitou $c < c_q$ jsou skoro určitě (pravděpodobnost se blíží jedné s $N \rightarrow \infty$) obarvitelné q barvami, a naopak velké náhodné grafy s konektivitou $c > c_q$ skoro určitě q barvami obarvit nelze. Existence (ale ne konvergence) kritické konektivity c_q byla dokázána v [5]. Existují lepší dolní a horní meze [6] na hodnotu c_q než ty, které jsme popsali výše.

Nabízí se zajímavá otázka: Existuje polynomiální algoritmus, který by našel obarvení pro skoro všechny velké náhodné grafy o dané konektivitě menší než c_q ? Uvědomme si, že najít takový algoritmus je mnohem jednodušší než najít polynomiální algoritmus pro všechny možné grafy. Nicméně ani tak nebyl zatím nikdo schopen odpovědět kladně. Podle výsledků numerických simulací se zdá, že pro konektivity blízké c_q se problém nalezení obarvení stává mnohem těžší než pro konektivity velmi malé. Nárůst složitosti se nedá vysvětlit tím, že blízko c_q existuje mnohem menší počet správných obarvení.

Navic tento fenomén se neobjevuje jen v problému barvení náhodných grafů, ale v celé řadě NP-úplných problémů, nejznámějším a nejdůležitějším je splnitelnost logických formulí (K-SAT) v mnoha variantách. Je tedy velkou výzvou a také je velmi užitečné pro praktické aplikace pochopit, co je příčinou nárůstu algoritmické složitosti ve všech těchto problémech.

JAKOU ROLI HRAJE STATISTICKÁ FYZIKA?

Představili jsme problém barvení náhodných grafů a nevyřešené otázky, které se ho týkají. Nyní se dostáváme k hlavní náplni tohoto článku, tzn. k roli statistické fyziky při popisu a pochopení barvení grafů a podobných problémů.

Denním chlebem statistické fyziky je Isingův model [7]. V původní formulaci [7] jde o model feromagnetu na mřížce, v každém bodě mřížky i je umístěn spin s_i , který může nabývat hodnoty $+1$ nebo -1 . Spiny feromagneticky interagují, tzn. že sousedé preferují mít stejnou hodnotu, energie dané konfigurace se pak zapíše jako

$$\mathcal{H}_{\text{Ising}} = \sum_{(i,j)} (1 - \delta_{s_i, s_j}). \quad (3)$$

kde suma je přes všechny sousedy a δ_{ab} je Kroneckerovo delta. Energie je nulová, pokud všechny spiny mají stejnou hodnotu, a kladná v každém jiném případě.

Jiným modelem Isingova typu je model Pottsův [8], kde spin s_i může nabývat q různých hodnot. Interakce mezi spiny může být také antife-

romagnetická, tzn. sousedé preferují být ve dvou různých stavech, odpovídající energie pak je

$$\mathcal{H}_{\text{Potts}} = \sum_{(i,j)} \delta_{s_i, s_j}. \quad (4)$$

Energie antiferomagnetického Pottsova modelu je tedy nulová právě tehdy, když konfigurace spinů odpovídá korektnímu obarvení mřížky. Umístíme-li antiferomagnetický Pottsův model ne na mřížku, ale na náhodný graf, převedli jsme barvení náhodných grafů na hledání základního stavu antiferomagnetického Pottsova modelu.

Ve fyzice velmi úspěšná metoda pro hledání základního stavu je simulované žíhání [9]. V simulovaném žíhání necháme systém ustálit (přejít do rovnováhy) za určité poměrně vysoké teploty. Teplotu pak velmi pomalu snižujeme a při každé změně čekáme, až se systém opět ustálí. Čím menší je teplota, tím menší je energie rovnovážného stavu. Nakonec při dosažení nulové teploty dosáhneme stavu s nejnižší možnou energií, tzv. stavu základního.

Není nic jednoduššího než vyzkoušet simulované žíhání jako algoritmus pro barvení grafů. Začneme s náhodným rozdělením barev a vysokou teplotou T . Krok, ve kterém systém přechází do rovnováhy, realizujeme Metropolitovým–Hastingovým Monte Carlo algoritmem [10]:

Vybereme spin a spočítáme, jak se změnila energie, pokud bychom změnili jeho barvu na nějakou jinou. Pokud se energie zmenší nebo zůstane stejná, změnu provedeme. Pokud se energie takto zvětší, provedeme změnu s pravděpodobností $e^{-\Delta E/T}$, kde T je teplota. Tento krok opakujeme mnohokrát, pak snížíme teplotu atd.

Simulovaným žíháním najdeme správné obarvení, pokud je konektivita grafu malá. Jiné heuristické algoritmy jsou sice rychlejší, ale nám jde o princip. Pokud se konektivita zvětšuje, potřebujeme stále více a více času na ustálení systému, uvedení do rovnováhy. Tento čas dokonce diverguje při určité konektivitě $c_d < c_q$. Podotkneme, že některé heuristické algoritmy, např. Walk-SAT [11], které se nesnaží najít rovnovážný stav při malé teplotě, mohou rychle nalézt obarvení i pro konektivitu větší než c_d . Nyní se pokusme pochopit, proč selhalo simulované žíhání.

ODBOČKA: JAK SE CHOVÁJÍ SKLA

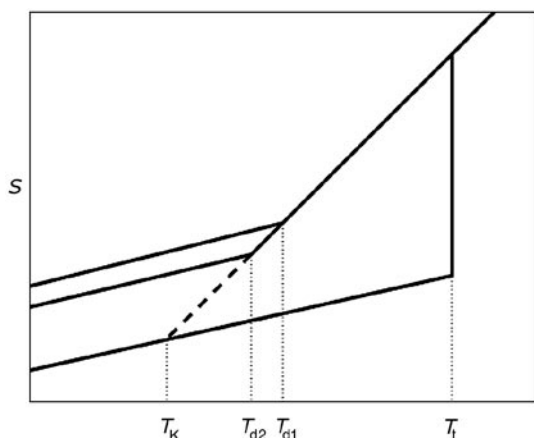
V této kapitole shrneme hlavní poznatky o skelných materiálech. Představíme dynamickou a Kauzmannovu teplotu. Spojitost s barvením grafů vyvstane vzápětí.

Tabulka skla ve vašem okně je příklad systému, který se nenachází v rovnovážném stavu, tím by byl krystal. O tom, jak přesně popsat skelný stav, se stále mezi fyziky vedou debaty [12]. Jedno je ale jasné, z mikroskopického pohledu se sklo (oxid křemičitý, SiO_2) více podobá kapalině než krystalu. Ale viskozita této kapaliny je tak velká, že bychom mohli pozorovat tečení, museli bychom čekat celou historii Vesmíru, a ještě mnohem déle, chceme-li se dočkat toho, až se sklo samovolně dostane do rovnovážného stavu (zkrystalizuje).

Je na místě vyvrátit lidovou pověru, že sklo teče sice pomalu, ale toto tečení se dá pozorovat v tabulkách oken starých katedrál. Sklo sice teče, ale mnohem pomaleji. To, že jsou tabulky některých starých katedrál silnější na dolním okraji než na horním, je dáno procesem jejich výroby (tenká tabulka se vyráběla na rotující plotně, nestejnорodost tloušťky je tedy dána odstředivou silou).

Popíšeme podrobněji chování skla při změnách teploty. Během experimentu měříme entropii systému. Přesněji měříme měrné teplo, ze kterého se integrací entropie dá spočítat, $c_p = T(\partial S/\partial T)_p$. Ohřejme sklo na teplotu tak vysokou, aby zkapalnilo. Pokud tuto kapalinu velmi pomalu chladíme (žiháme), její entropie bude klesat (molekuly mají menší volnost pohybu při menší teplotě), pod teplotou tání T_t kapalina zkrystalizuje a entropie klesne skokově na mnohem menší hodnotu (krystal je periodicky uspořádaný a jediný příspěvek k entropii pochází z kmitů molekul kolem rovnovážných poloh).

Zchladíme-li naopak kapalinu velmi rychle (zakalíme), při teplotě blízké teplotě tání nebude mít systém dostatek času najít výhodný krystalický stav, místo toho obdržíme podchlazenou kapalinu, jejíž entropie klesá s teplotou zhruba stejně rychle jako entropie kapaliny nad teplotou tání. Podchlazená kapalina se nachází v jednom z mnoha amorfních stavů, mezi těmito stavy jsou energetické bariéry. Čím nižší je teplota, tím déle trvá překonání těchto bariér, a ještě mnohem déle by trvalo nalezení krystalického uspořádání. Dosáhne-li teplota hodnoty T_d , systém není dále schopen energetické bariéry překonávat a z pohledu experimentátora zamrzne v jednom ze stavů. Entropie počítaná integrací měřeného specifického tepla začne při teplotě T_d klesat stejně pomalu jako entropie krystalu. Entropie skutečná, kterou měřit neumíme, by se téměř nelišila od entropie krystalu, jelikož by byla dána opět jen kmity molekul kolem jejich poloh.



1 Entropie kapaliny, krystalu, podchlazené kapaliny a dvou skel vzniklých různě rychlým chlazením. Krystal taje při teplotě T_t , přechlazená kapalina zamrzne do skelného stavu při teplotě T_{c1} při rychlejším chlazení, T_{d2} při pomalejším chlazení. Extrapolovaná entropie přechlazené kapaliny protne entropii krystalu při Kauzmannově teplotě T_K .

Problémem je, že hodnota T_d závisí na tom, jak rychle jsme systém chladili. Čím pomaleji, tím je T_d nižší. Kdybychom měli k dispozici nekonečně mnoho času, mohli bychom po rychlém zchlazení hluboko pod teplotu tání začít čekat a čekat a dynamická teplota T_d by byla stále nižší a nižší (ovšem pro její snížení o zlomek procenta bychom museli čekat tisíce let).

Obecně entropie systému odpovídá logaritmu počtu možných konfigurací. Entropii podchlazené kapaliny $s_{tot} = s + \Sigma$ můžeme tedy přirozeným způsobem rozložit na součet členu pocházejícího z lokálních fluktuací s , který se rovná přibližně entropii krystalu, a členu pocházejícího z počtu možných amorfních stavů $\mathcal{N} = e^{N\Sigma}$. Z tohoto důvodu se tento druhý člen (rozdíl entropie podchlazené kapaliny a krystalu) nazývá konfigurační entropie Σ . Můžeme extrapolovat konfigurační entropii pod dynamickou teplotu T_d , zaměříme se na teplotu, kde se extrapolovaná konfigurační entropie stane nulovou. Pro oxid křemičitý (sklo v okenních tabulích) se tak stane při nulové teplotě, všechna skla, která se chovají stejně, se nazývají skla silná [12]. Existuje druhá skupina skel, tzv. skla křehká [12], pro která konfigurační entropie vymizí při kladné teplotě T_K (podle Kauzmannova, objevitele tohoto paradoxu [13]).

Chování křehkého skla je načrtnuto na obr. 1. Kauzmannův paradox spočívá v tom, že negativní konfigurační entropie není možná, jelikož podchlazená kapalina by byla uspořádanější nežli krystal. Co se stane s podchlazenou kapalinou pod Kauzmannovou teplotou? Odpověď na tuto otázku není známá, je mnoho teorií, ale vědecký konsenzus nebyl dosažen.

CO JE NA BARVENÍ NÁHODNÝCH GRAFŮ SKELNÉHO?

Vraťme se ale k barvení náhodných grafů a výpočetní složitosti a pokusme se vyložit spojitost se skelným chováním. Kontrolním parametrem zde není teplota, nýbrž průměrná konektivita c , jinými slovy počet vazeb (interakcí) připadající na dva vrcholy. Zvyšování konektivity je tedy analogické zvyšování hustoty materiálu, ve většině materiálů se hustota zvyšuje se snižující se teplotou. Popíšeme strukturu všech korektních obarvení velkého grafu o dané konektivitě. Jedno korektní obarvení si lze představit jako jednu možnou konfiguraci skla. Entropie zde je logaritmus počtu obarvení vydělený velikostí grafu.

Rozdílem je, že v barvení náhodných grafů neexistuje nic obdobného krystalu, jak ale můžeme tušit, samotná existence krystalového stavu není pro skla příliš podstatná. Při neexistenci krystalu by ve skle byla rovnovážným stavem podchlazená kapalina, v předchozí kapitole jsme ale viděli, že pro teplotu menší než T_d se dosažení rovnováhy nejsme schopni dočkat, namísto toho je systém zamrzlý v amorfním skelném stavu. Analogicky, jak jsme již popsali při výkladu simulovaného žihání, při barvení náhodných grafů se nedočkáme rovnováhy pro konektivitu větší než c_d .

Při konektivitě c_d se množina správných obarvení rozdělí na mnoho podmnožin (klastřů), odpovídajícím mnoha amorfním stavům podchlazené kapaliny. Konfigurační entropie tentokrát odpovídá logaritmu počtu klastřů (vydělenému velikostí grafu), které pokrývají skoro všechna obarvení. Zvyšujeme-li konektivitu c_K zcela vymizí. Odpověď na otázku: „Co se stane poté?“ je v tomto případě známá. Skoro všechna správná obarvení jsou koncentrována v poměrně malém (na velikosti grafu nezávislejícím) počtu klastřů. Následně, překročíme-li již zmíněnou konektivitu c_q , všechny klastry vymizí, a tím i všechna korektní obarvení (analogie tohoto přechodu v reálných sklech neexistuje).

Analogie je to hezká, ale proč by měla fungovat? Jak jsme řekli, strukturní skla jsou studována v mnoha pracích, ale konsenzus pro jejich teoretický popis neexistuje. Existují buď modely mikroskopickým vlastnostem skel věrné, ale neřešitelné (resp. nevyřešené), nebo modely mikroskopicky zjevně nesprávné, řešitelné, chovající se kvalitativně podobně jako skla. Na druhou stranu strukturu všech obarvení náhodného grafu spočítat a popsat umíme, sice ne na úrovni matematicky rigorózní, ale testem matematické rigoróznosti projde jen velmi málo výsledků statistické fyziky.

JAK OBARVENÍ POPSAT: METODA KAVITY

Vydejme se tedy do části techničtější, ve které popíšeme fyzikální metodu kavity [14], jež se pro popis obarvení náhodného grafu používá. Čtenář, který shledá tuto část nesrozumitelnou, může přeskočit na shrnutí v popisku obrázku 3, a pokračovat částí o algoritmických důsledcích.

Princip metody kavity je jednoduchý, vyjmeme spin (vrchol) i z grafu (odtud jméno kavity), dále předpokládejme, že v nepřítomnosti vrcholu i jsou jeho sousedé $V(i)$ nezávislí. Za tohoto předpokladu napíšeme iterativní rovnice pro pravděpodobnost $\psi_{s_i}^{i \rightarrow j}$, že spin i je obarven barvou s_i , podmíněnou nepřítomností spinu j ,

$$\begin{aligned} \psi_{s_i}^{i \rightarrow j} &= \frac{1}{Z_0^{i \rightarrow j}} \prod_{k \in V(i) - j} \sum_{s_k} (1 - \delta_{s_i s_k}) \psi_{s_k}^{k \rightarrow i} \\ &= \frac{1}{Z_0^{i \rightarrow j}} \prod_{k \in V(i) - j} (1 - \psi_{s_i}^{k \rightarrow i}), \end{aligned} \quad (5)$$

kde $k \in V(i) - j$ jsou sousedé vrcholu i kromě vrcholu j , a $Z_0^{i \rightarrow j}$ je normalizační konstanta. Najdeme řešení systému rovnic (5) a napíšeme pravděpodobnost, že spin i je obarven barvou s_i :

$$\psi_{s_i}^i = \frac{1}{Z_0^i} \prod_{k \in V(i)} (1 - \psi_{s_i}^{k \rightarrow i}), \quad (6)$$

Z této pravděpodobnosti můžeme spočítat všechny termodynamické veličiny, zejména entropii. Řešením rovnice (5) je $\psi_{s_i}^{i \rightarrow j} = 1/q$, kde q je stále počet barev. Numerickým řešením rovnic (5) se můžeme

presvědčit, že pro téměř všechny velké obarvitelné grafy je toto řešení jediné, které lze iterativně nalézt. Odpovídající entropie se rovná té, kterou jsme spočítali v (1), tedy

$$s = \log q - (c/2) \log(1 - 1/q). \quad (7)$$

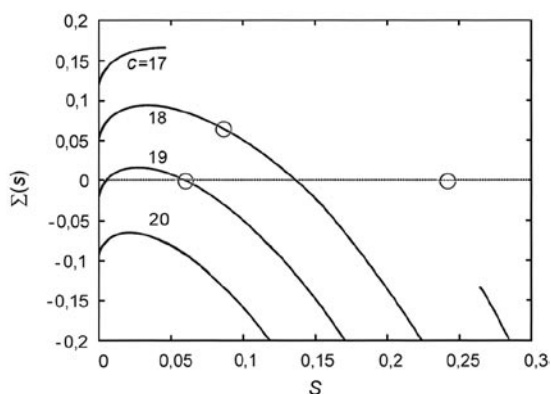
Poznámkou pro informatiky budiž, že iterativní rovnice (5) jsou známé pod názvem šíření domněnek (belief propagation) [15], tyto rovnice se používají v algoritmech pro samoopravování kódů [16].

Principu podobného metodě kavity se ve fyzice používá již dlouho v různých formách, zejména pro přibližné řešení modelů na mřížce je znám pod názvem Bethého přiblížení [17]. V našem případě se ovšem nebude jednat o řešení přibližné. Dá se zformulovat a ověřit podmínka [18], při které je výše uvedeného řešení pro velké náhodné grafy přesné. Zjednodušeně řečeno, pokud by graf byl stromový, vrchol i byl kořenem a my bychom znali hodnoty všech spinů na okraji, vrchol i musí zapomenout veškerou informaci, kterou by mu mohly okrajové spiny předat (matematici formalizují tuto podmínku do pojmu: příslušná Gibbsova míra je extrémální [18, 19, 23]). Pro barvení náhodných grafů je uvedené řešení přesné pro konektivity menší než c_d , pro konektivity větší než c_d je řešení z podstaty chybné, ač jak uvidíme, vztah pro entropii je platný také pro všechny konektivity menší než c_K .

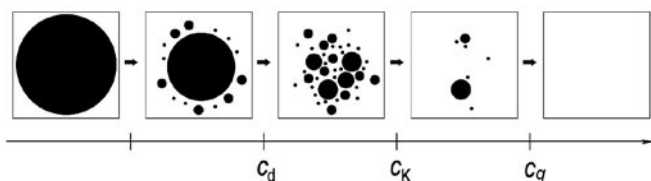
KLASTRY SPRÁVNÝCH OBARVENÍ

Pro konektivitu grafů větší než c_d tedy musíme výše uvedený postup opravit. Ukazuje se, že řešením je uvažovat rozdělení všech obarvení na mnoho podmnožin (klastřů, stavů). Uvnitř každého stavu opět použijeme rovnice metody kavity a na závěr zprůměrujeme výsledky přes všechny stavy. Podotkneme, že rozdělovat množinu konfigurací (přesněji pravděpodobnostní míru na množině konfigurací) na více částí (stavů) je ve statistické fyzice standardní procedura, například při řešení Isingova modelu při vysoké teplotě popisujeme systém pomocí jednoho jediného stavu (paramagnetu), kdežto při teplotě nízké pomocí dvou stavů (feromagnetických). Kritická teplota je i zde dána hodnotou, při které průměrná hodnota spinu uprostřed mřížky začne záviset na okrajových podmínkách. Rozdíl (a zdroj těžkostí pro matematiky) v našem případě je, že počet potřebných stavů závisí exponenciálně na velikosti grafu, $\mathcal{N} = e^{N^2}$. Nicméně metoda kavity si i s takovým případem umí poradit. Tomuto postupu se říká jedнокrokové narušení symetrie replik, jelikož byl původně odvozen pomocí mnohem hůře interpretovatelného replikového triku [20].

Podobně jako zavádíme teplotu ve statistické fyzice, abychom se mohli ve výpočtech zaměřit na systém o dané střední energii, zavádíme zde také Lagrangeův multiplikátor, abychom se mohli zaměřit na stavy o dané velikosti (vnitřní entropii s , logaritmu počtu obarvení náležících do stavu vyděleného velikostí grafu). Výsledkem je systém



2 Konfigurační entropie stavů (logaritmus počtu stavů vydělený velikostí grafu) v závislosti na jejich vnitřní entropii pro obarvení náhodných regulárních grafů o konektivitě c šesti barvami. Převzato z [23].



3 V tomto obrázku si každý bod lze představit jako jedno obarvení grafu. Směrem doprava roste průměrná konektivita vrcholů. Pro konektivitu menší než c_d patří skoro všechna řešení do jediného stavu, odpovídá kapalné fázi ve sklech. Pro konektivitu větší než c_d , ale menší než c_K jsou obarvení rozdělena do exponenciálně mnoha stavů, konfigurační entropie je pozitivní a sklo je dynamicky zamrzlé v jednom ze stavů. Pro konektivitu větší než c_K , ale menší než c_q jsou téměř všechna obarvení koncentrována do několika stavů, konfigurační entropie je nulová, ve sklech bychom možná mohli tuto fázi pozorovat, pokud bychom dokázali kapalinu podchládit pod Kauzmannovu teplotu. Konečně nad konektivitou c_q neexistují žádná korektní obarvení. Převzato z [23].

funkcionálních rovnic pro funkci $\Sigma(s)$, který umíme numericky vyřešit. Výsledek pro obarvení šesti barvami náhodné grafy o konektivitě c je na obr. 2.

Interpretujme tento výsledek. K tomu potřebujeme pojem termodynamicky dominantní stavu, to jsou stavy, ve kterých se téměř jistě octneme, vybereme-li náhodně jedno obarvení, tedy stavy odpovídající maximu celkové entropie $s_{\text{tot}} = s + \Sigma$. Celková entropie nabývá maxima, pokud $\partial\Sigma(s)/\partial s = -1$ (pokud v tomto bodě funkce $\Sigma(s)$ existuje). V závislosti na konektivitě je několik možností:

- Při velmi malé konektivitě funkce $\Sigma(s)$ sestává z jediného bodu $\Sigma = 0$ pro entropii odpovídající rovnici (7). Množina správných obarvení pak sestává z jediného stavu. První část na obr. 3.
- Při vyšší konektivitě, ale stále menší než c_d , existuje sice exponenciální počet stavů, ale termodynamicky dominantní je stále jediný veliký stav. Příklad je na obr. 2, konektivita $c = 17$, dominantní stav označen kroužkem. Druhá část na obr. 3.
- Pro konektivitu větší než c_d , ale stále menší

než c_K je dominantních klastrů exponenciálně mnoho a odpovídají podmínce $\partial\Sigma(s)/\partial s = -1$. Celková entropie je stále dána výrazem (7). Příklad je na obr. 2, konektivita $c = 18$, dominantní stavy označeny kroužkem. Třetí část na obr. 3.

- Pro konektivitu větší než c_K , ale stále menší než c_q bod $\partial\Sigma(s)/\partial s = -1$ odpovídá negativní konfigurační entropii, takové klastry se vyskytují jen s pravděpodobností exponenciálně malou. Celková entropie je tedy dominována konečným počtem klastrů odpovídajícím $\Sigma(s) = 0$ a je menší než entropie ve výrazu (7). Příklad je na obr. 2, konektivita $c = 19$, dominantní stavy označeny kroužkem. Čtvrtá část na obr. 3.
- Pro konektivitu větší než c_q je celá funkce $\Sigma(s)$ negativní, a tedy žádné klastry nezůstaly. Příklad je na obr. 2, konektivita $c = 20$. Pátá část na obr. 3.

V popisku obr. 3 jsme situaci shrnuli, popsali jsme také ještě jednu souvislost s chováním reálných křehkých skel. Zastavme se zde u několika bodů, které si zaslouží zmínku. Antiferomagnetický Pottsův model (při nulové teplotě ekvivalentní barvení grafů) vykazuje skelné chování nejen při změnách konektivity, ale také při změnách teploty. Není to jediný takový model, historicky prvním podobným modelem, který vykazuje stejnou fenomenologii jako reálná skla, bylo p-spinové sklo [21] a později modely mikroskopicky sklům bližší, tzv. skelné modely na mřížce [22]. Všem těmto fyzikálním modelům je společné to, že jsou řešitelné jen na náhodných grafech (nebo na grafech úplných), reálným sklům by spíše odpovídalo řešení takových modelů na třídídimenzionální mřížce. Jeden ze zásadních rozdílů spojených s tímto faktem je následující: Pro reálná skla není teplota skelného přechodu T_d dobře definována, nýbrž závisí na rychlosti chlazení. Pro náhodné grafy je odpovídající konektivita c_d naopak definována dobře. Debatu, jak moc jsou tyto modely relevantní pro popis skel, je stále otevřená. V barvení grafů by tento problém přešel na otázku, jak moc jsou výsledky pro náhodné grafy relevantní pro grafy pocházející z praktických aplikací.

ALGORITMICKÉ DŮSLEDKY

Vraťme se nyní k otázkám položeným na začátku. Jak pochopení struktury všech obarvení a souvislost s reálnými skly pomůže při pochopení algoritmické složitosti a při návrhu nových efektivnějších heuristických algoritmů?

Metoda kavity umožňuje nejen popsat klastry řešení a spočítat všechny kritické konektivity, ale také napsat iterativní rovnice podobné (5), které berou v úvahu existenci mnoha klastrů. Pro daný klaster definujeme vrcholy zamrzlé do barvy s jako ty, které ve všech obarveních patřících do tohoto klastru mají barvu s . Pravděpodobnost, že v náhodném klasteru je vrchol i zamrzlý do barvy s , podmíněnou nepřítomností vrcholu j , označíme $\eta_s^{i \rightarrow j}$, vztah $\sum_{s=1}^q \eta_s^{i \rightarrow j} + \eta_0^{i \rightarrow j} = 1$ definuje člen $\eta_0^{i \rightarrow j}$. Iterativní rovnice pro pravděpodobnost $\eta_s^{i \rightarrow j}$ pro obarvování třemi barvami je

$$\eta_s^{i \rightarrow j} = \frac{\prod_{k \in V(i)-j} (1 - \eta_s^{k \rightarrow i}) - \sum_{p \neq s} \prod_{k \in V(i)-j} (\eta_0^{k \rightarrow i} + \eta_p^{k \rightarrow i}) + \prod_{k \in V(i)-j} \eta_0^{k \rightarrow i}}{\sum_p \prod_{k \in V(i)-j} (1 - \eta_p^{k \rightarrow i}) - \sum_p \prod_{k \in V(i)-j} (\eta_0^{k \rightarrow i} + \eta_p^{k \rightarrow i}) + \prod_{k \in V(i)-j} \eta_0^{k \rightarrow i}}. \quad (8)$$

V čitateli je pravděpodobnost, že žádný ze sousedů vrcholu i kromě j není zamrzlý do barvy s , ale alespoň jeden je zamrzlý do všech ostatních barev. Ve jmenovateli je pravděpodobnost, že existuje barva, do které není zamrzlý ani jeden ze sousedů vrcholu i kromě j . Rovnice (8) se říká šíření průzkumů (survey propagation) a byla poprvé odvozena pro problém splnitelnosti logických formulí v [24]; pro barvení grafů v [25]. Šíření průzkumů se dá použít jako algoritmus pro nalezení korektního obarvení. Jeden ze způsobů je iteracemi najít řešení rovnice (8), zafixovat barvu u zlomku z vrcholů, u kterých je $\eta_s^{i \rightarrow j}$ pro jednu barvu výrazně větší než pro ostatní barvy, a pokračovat tak dlouho, dokud rovnice (8) mají netriviální řešení. Na obarvení zbytku grafu se pak použije nějaký ze známých heuristických algoritmů (v praxi se používá Walk-SAT [11]).

Algoritmus šíření průzkumů je schopen najít správná obarvení pro velké náhodné grafy (až do několika milionů vrcholů) pro konektivity tak blízké c_q jako žádný jiný známý algoritmus. A právě tento fakt upoutal pozornost teoretických informatiků, kteří se začali zajímat o použité metody a myšlenky statistické fyziky.

ZÁVĚR

Popsali jsme fascinující cestu od snahy pochopit chování reálných skel po vývoj neefektivnějšího známého heuristického algoritmu pro obarvování grafů, problému důležitého jak z teoretického, tak z praktického hlediska.

Výzkum v tomto oboru je v posledních několika letech velmi aktivní. Hlubší porozumění a vývoj ještě efektivnějších algoritmů je motivací pro spolupráci mnoha fyziků, matematiků a informatiků.

Je několik důležitých stále otevřených otázek. Matematikové usilují o rigorózní důkaz popsaných výsledků. Informatiky zajímá, zda mohou fyzikální metody pomoci při analýze známých heuristických algoritmů a jak vylepšit algoritmus šíření výzkumů, aby pracoval rychleji a ještě blíže konektivitě c_q . Fyzikové by rádi studovali důsledky dělení klastřů na podklastry atd., což je relevantní pro určité intervaly parametrů. A konečně všichni by rádi pochopili, jak zobecnit použité metody pro analýzu nenáhodných grafů.

LITERATURA

- [1] K. Appel, W. Haken, *Illinois J. Math.* **21**, 421 (1977); 491 (1977).

- [2] S. Cook: *Theory of Computing*, Proc. 3rd Annual ACM Symposium. ACM, New York 1971, s. 151.
M. R. Garey, D. S. Johnson: *Computers and Intractability: A Guide to the Theory of NP-Completeness*, Freeman, New York 1979.
C. H. Papadimitriou: *Computational complexity*, Addison-Wesley, 1994.
- [3] P. Erdős and A. Rényi, *Publ. Math.* **6**, 290 (1959).
- [4] B. Pittel, J. Spencer, N. Wormald, *J. Combinatorial Theory, Series B* **67**, 111 (1996).
- [5] D. Achlioptas, E. Friedgut, *Random Structures and Algorithms* **14**, 1, 63 (1999).
- [6] D. Achlioptas, A. Naor, Y. Peres, *Nature* **435**, 759 (2005).
- [7] S. G. Brush, *Rev. Mod. Phys.* **39**, 883 (1967).
- [8] F. Y. Wu, *Rev. Mod. Phys.* **54**, 235 (1982).
- [9] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, *Science*, **220**, 671 (1983).
- [10] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [11] B. Selman, H. A. Kautz, B. Cohen: *Proc. of AAAI-94*, Seattle 1994.
- [12] P. G. Debenedetti, F. H. Stillinger, *Nature* **410**, 259 (2001).
- [13] W. Kauzmann, *Chem. Rev.* **43**, 2, ■■ (1948).
- [14] M. Mézard, G. Parisi, *Eur. Phys. J. B* **20**, 217 (2001); *J. Stat. Phys.* **111**, 1 (2003).
- [15] J. Pearl: *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. (2. vyd.). Morgan Kaufmann Publishers, San Francisco 1988.
- [16] R. G. Gallager: *Low Density parity Check Codes*. MIT Press, Cambridge, MA, 1963.
- [17] H. A. Bethe, *Proc. R. Soc. Lond.* **150**, 552 (1935).
R. Peierls, *Proc. Roy. Soc.* **154**, 207 (1936).
- [18] M. Mézard, A. Montanari, *J. Stat. Phys.* **124**, 1317 (2007).
- [19] H. O. Georgii: *Gibbs Measures and Phase Transitions*, Berlin 1988.
- [20] G. Parisi, *J. Phys. A* **13**, L115 (1980).
- [21] T. Kirkpatrick, D. Thirumalai, *Phys. Rev. Lett.* **58**, 2091 (1987).
- [22] G. Biroli, M. Mézard, *Phys. Rev. Lett.* **88**, 025501 (2001).
- [23] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, L. Zdeborová, *Proc. Natl. Acad. Sci.* **104**, 10318 (2007).
- [24] M. Mézard, G. Parisi, R. Zecchina, *Science* **297**, 812 (2002).
M. Mézard, R. Zecchina, *Phys. Rev. E* **66**, 056126 (2002).
- [25] R. Mulet, A. Pagnani, M. Weigt, R. Zecchina, *Phys. Rev. Lett.* **89**, 268701 (2002).
A. Braunstein, R. Mulet, A. Pagnani, M. Weigt, R. Zecchina, *Phys. Rev. E* **68**, 036702 (2003).