

解释离婚情感动态的数学模型

Jose-Manuel Rey / 文 李玉田 / 译

摘要

背景

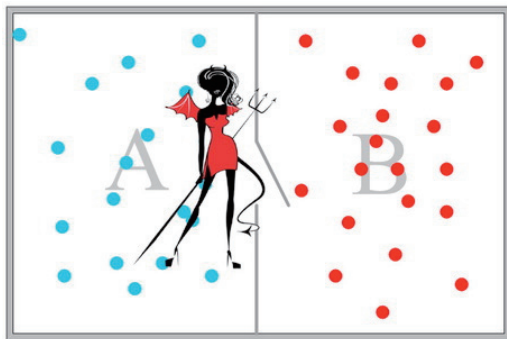
西方社会的离婚是普遍存在的。这无论在理论上还是解决方式上都引发了很严重的科学与社会学问题。学者和问题处理专家认为存在一种情感关系的热力学第二定律。仅有爱是不够的，还需努力来维持。

方法 / 主要发现

我们在一个简化的第二定律的基础上用最优控制理论作为建立情感动态模型的一个新处理。我们的分析结果与社会学数据相一致。我们揭示出，如果双方有相似的情感属性，存在一种最优的努力策略使得婚姻长久幸福。这一策略因两方面因素的组合而受到结构性不稳定的制约：一是努力差距，因为最优策略通常带来不愉快；一是由于动力系统的不稳定性使得努力降低到不可持续水平的趋势。

结论 / 重要性

这个模型的数学事实揭示了真实情境中配偶关系破裂的相关机制。在这个框架下，第二定律可以解释一个显见的悖论——一贯想天长地久的婚姻可能会破裂。



引 论

有着浪漫特性的情感关系通常被认为是西方社会中安定的幸福生活的基本组成部分。当人们被问及什么是幸福生活的必备要素时，大家通常将“爱”和“亲密关系”摆在首位。很难想象人类生活的其它方面会涉及如此多的文化、社会、心理及经济学问题。理论研究显示浪漫关系的最初阶段似乎是受化学作用控制，而维护一段情感关系则在较大程度上属于理性决策的范畴。人们通常在仔细考虑之后才投入长期的关系，最典型的就是婚姻。即使在一直是一夫一妻制的西方社会，配偶通常声称他们的意图是维持长久的关系并和睦相处。但欧美大量报道的高离婚率表明了他们在方案执行上的彻底失败。配偶关系破裂这一现象在美国被视为流行病，“每两对夫妇中就有一对以离婚收场”这一统计结果被媒体和学术报告频繁引用。欧盟 27 国的离婚率并不比美国好到哪里去，而且一些欧洲国家表现出了更高的离婚率。

诸多领域的学者们普遍认为，婚姻不稳定性日益加剧的主要原因应归结为二十世纪性别分工的变化所释放的经济力量。然而这一原因并不能解释过去几十年里不断增长、无处不在的婚姻破裂现象。事实上，人们并不理解为什么这么多配偶最终离婚而有些人却没有离婚。这一理解尤为重要，因为婚姻破裂所诱发的社会变化深深地影响着当代西方社会的社会结构以及福祉。

对大多数配偶来说，双方都打算维持长久的关系并为这一目标而努力，这个事实却与报道出来的高离婚率相



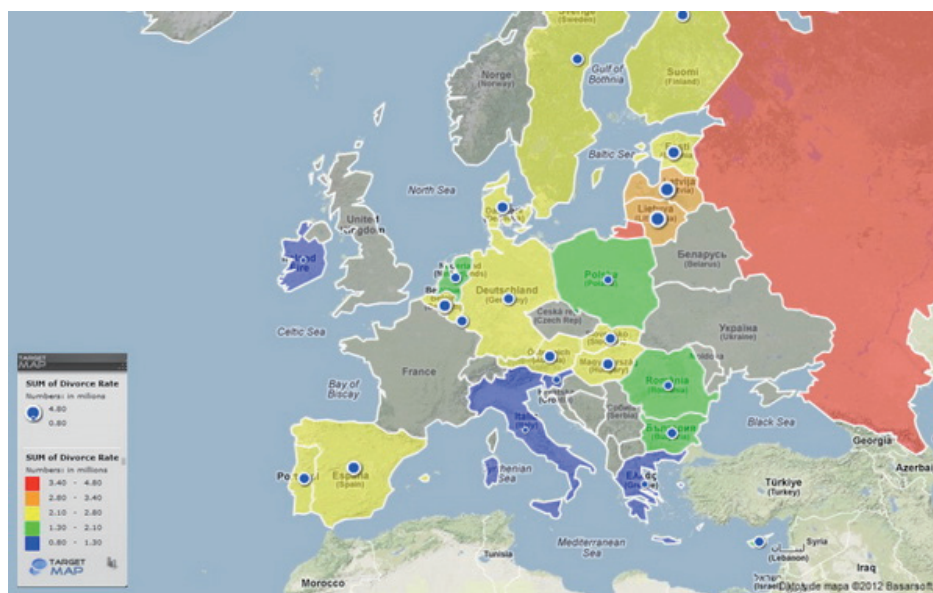
矛盾。本文称这种矛盾为**失败悖论**。根据戈特曼 (J. M. Gottman) 等人的研究，婚姻研究领域亟需一个理论，尤其是数学理论。本文旨在满足这一需求。尤其是本文将为这种失败悖论提供一个合理的解释。

戈特曼等人的研究 (收录于 *The Mathematics of Marriage-Dynamic Nonlinear Models* 一书) 似乎是迄今为止对配偶关系所进行的唯一的一项数学研究。他们用一对非线性差分方程评估了双方在实验观察时的短期互动。斯特罗伽茨 (S. H. Strogatz) 率先提出了一个配偶互动的简单动态系统模型。我们这里采用不同的动力系统方法：配偶被视作一个单元 (不考虑内部互动)，他们的情感动态由其长期和睦共处的意愿来合理地描述。

鉴于配偶分手现象的普遍性，相对于情感关系中的具体瑕疵，寻找能解释分手现象的基本的确定性机制似乎更为合情合理。在社会学数据的基础上，我们提出了一个基于最优控制理论的数学模型来解释同类型配偶如何建立长期关系的理性计划。如果配偶双方有相似的特性，我们称之为同类型配偶。同类型婚配是西方社会情感关系中最普遍的类型。我们的模型实际上需要的是一种较弱的同类型性。我们通过基于情感互动的**热力学第二定律** (以下简称**第二定律**) 的动力学方程来描述情感关系的演化。第二定律指出，除非有“能量”供给，否则一段关系将会破裂。这一被广为接受的事实允许我们将情感关系建模成一个控制问题，以努力这一形式存在的能量充当着控制变量。最优控制理论已经被广泛地应用于应用科学，如工程学和经济学。我们的最优控制模型为分析婚姻和亲密关系提供了一个全新的数学方法。

给定一些可行性的条件，我们对于模型的分析表明维持长期成功的关系是可能的，并且可对应于动力系统的平衡路径。然而很明显的是，长期关系中不可或缺的是努力，这一模型最引人注目的发现是保持幸福关系的努力水平总是比先验的最优选择 (假如只考虑目前情形) 的努力水平要高。当这两种水平的**努力差距**在可容忍的范围内，维持一段关系是切实可行的。这一数学分析的最主要结果是：服从于第二定律的情感动力系统本质上是不稳定的。这说明，当努力放松时，情感的逐步恶化会轻易地发生。这一分析还确立了一个合理的机制用来解释逐渐退化将导致感情破裂抑或使人不满的情感生活。

本文的研究结果有助于解答关系失败悖论：根据第二定律，长久幸福关系的优化设计与动态不稳定性并由此导致的可能的分手相容。这一显著的发现解决了这一失败悖论，



欧洲各国和地区 2010 年的离婚率

因为真实的关系被认为有更多不稳定性和不确定性的源头。而且这一结果也许揭示了如何才能保持健康有活力的长久关系。

本文的第 2 部分将展示由社会学数据所支持的关键证据，它们将构成一个检验模型结果一致性的框架。失败悖论这一问题也由社会学证据来导出。随着对于基本假设的深入讨论我们将逐步提出这一模型的构成要素。对于该模型分析得出的主要预测将汇集在本文的第 3 部分，其中一些预测与第 2 部分中提到的事实保持一致。为了便于讨论，数学上的一些技术细节被放置到附录中。

方法

公认的事实

马丁和巴姆帕斯 (Martin and Bumpass) 使用 1985 年的数据表明，在 40 年之内，美国每三个婚姻中就有两个将以分居或离婚而结束。目前可能还没有达到这一比例，但是 2002 年的数据显示实际数字并不低多少。50% 的人在 40 多岁的时候已经离婚不止一次。2005 年的数据表明，报道中经常提到的 50% 这一数字只比欧盟 27 国的平均离婚率 (44%) 稍高，而在一些欧盟国家，这一比例甚至高达 71%。

如果包括未婚同居，这一数据将进一步上升，尽管基于同居状态的数据很难获得。一项最近的研究证实，总体上非婚姻同居比婚姻更不稳定：49% 的婚前同居在 5 年内分手 (10 年后分手率是 62%)，然而只有 20% 的婚姻在 5 年内以分居或离婚告终 (10 年后这一数字是 33%)。我们正在观察的这一现象的第一个公认的事实可以阐述如下：

断言 1 在爱情关系中普遍存在失败。

这种臭名昭著的情感关系不稳定性与婚姻或同居关系中的信任缺失并不相关，虽然信任是幸福的主要因素。相反地，人们普遍表示满意的情感关系是建立幸福生活的第一要素。并且人们还声称他们希望自己的伴侣伴随一生：

断言 2 配偶通常相信关系的持久性是追求幸福的重要组成部分。而且大部分人认为，他们自己的关系不会破裂。

现有的数据支持断言 2。当被问到使得他们幸福的要素时，78% 的美国大学生选择了“与自己理想的伴侣相爱并维持下去”。在美国的一次国家调查中，93.9% 的受访已婚夫妻认为他们离婚或分居的可能性很低 (19.9%) 或相当低 (74%)，而 81.1% 未婚受访者也有同样的回答 (32.4% 的受访者认为离婚可能性低，47.7% 的受访者则认为非常低)。

有趣的是，尽管公认分手的可能性非常高，绝大多数的人仍然认为他们自己的情感关系不会破裂。事实上，断言 1 与断言 2 提出了一个明显的悖论。根据上述的数据，一个新组成的伴侣声称有九成的把握走到最后。然而 5 年之后分手的同居者是 50%；10 年之后他们则更有可能不在一起。这一事实可以叙述如下：

失败悖论：打算长久持续的情感关系却怎么极可能破裂？

下面提出的模型表明，在合理的假设之下，断言 1 和断言 2 是相容的。为了进一步测试该模型的一致性，我们还将考虑其它两个公认的事实。

断言 3 配偶分手是一种逐渐恶化过程的结果。

已有的数据支持这一事实。根据加州离婚调解计划中80%受访男女的情况，他们离婚的主要原因是“逐渐产生的距离感和渐渐失去的亲密度，或许呆在一起但是情感上却逐步疏离，以致再也不能忍受孤独感”。

断言 4 配偶双方的主观幸福感在婚后逐渐降低。

尽管都认为婚姻比起单身有着较高的幸福感，但已婚人士自己感受到的满足感却在慢慢降低。研究人员发现，结婚后人们对于生活的满足感逐步下降。

模 型

依据以上情景，我们将建立一个简单的动力系统模型。

这一模型基于两个关键的假设，也就是下面假设 2 中要讨论的第二定律和上述断言 2 所说的配偶关系的长期规划。这些假设连同弱同类性（参见下述假设 1）以及情感状态的一个自然的成本 - 收益评估（下面的假设 3），允许我们将配偶的情感关系视为一个最优控制问题。

建模始于（时间 $t = 0$ ）浪漫的热恋期刚刚结束，对彼此的好感达到顶点之时。在这一起始时间，伴侣们有强烈的意愿希望能最终走到一起，并试图去做任何必需的事情以保证两人的长远未来。我们假设：

假设 1（弱同类性） 伴侣双方拥有下述模型所指的相同特质。或者说，一对伴侣作为规划问题的一个决策单元。

这一假设预示着模型中所有的参数、变量和效用结构都是定义在一对由两个类似的个体组成的伴侣上。人们倾向于被具有相同特质的人所吸引，这一事实已经在现有文献中获得广泛认同。西方社会中有充分的证据支持这一事实。所以假设 1 被视作一个法则，而不仅是一个期望。严格来说，我们的理论仅仅要求情感而不是性格上的相似性（参见下面的假设 3）会促成配偶关系，不论是一起约会或者是已婚夫妻。

根据前面的叙述，下面的假设对我们的模型至关重要。

假设 2（情感关系的热力学第二定律） 对于另外一方的感觉有一种消退的趋势。这种惯性必须由有意识的实践来抵消。

在诸多文献中对这一事实有着广泛的共识。没有精心维护的爱会随着时间消退似乎是一个自然法则，雅各布森(N. S. Jacobson)和马戈林(G. Margolin)认为这是婚姻不稳定的主要诱因。他们写道：“婚姻以幸福开始，但是时间的侵蚀是婚姻问题的源头。”流行的格言“爱还不够”反映了这一事实并含蓄地表明侵蚀是可以通过某种办法防止的。假设 2 这种定律的形式取自戈特曼等人的书里，在那里，情感被建议性地解释为“类似于婚姻关系的热力学第二定律的某种东西：一种缺了保持关系鲜活良好的能量就会分崩离析的东西”。

为了将假设 2 转化为数学语言，我们用一个非负变量

$x(t)$ 来表示在 t 时刻 ($t \geq 0$) 的关系状态。它是一个感情变量，可以理解为伴侣对对方的情感。 $x(t)$ 是一个有序变量，它用来探测情感关系的质量水平。 $x(t)$ 的一个特定值并不包含任何信息，但是在不同时刻 t_1, t_2 的情感水平可以进行比较，比如 $x(t_1) \geq x(t_2)$ 还是 $x(t_1) \leq x(t_2)$ 。在 $t = 0$ 时刻共同的感情水平 $x(0) = x_0$ 被设定为一个非常大的数值。我们假定 $x(t)$ 降低到低于某个阈值 $x_{min} > 0$ 时，关系开始变得令人不满意，不同的伴侣的阈值不尽相同。

根据假设 2，对关系的细心处理可以抵消这种退化的惯性。这种处理用一个非负的有序变量 $c(t)$ 来表示，这个变量被称为努力变量，并假定这个变量是分片连续的（请看附录 1 中的讨论）。 $c(t)$ 包含所有日常生活中加强情感关系的实践，例如，治疗师提出的建设性行动（发问、积极地聆听、一起制定计划），以及宽容的态度（接受配偶的缺点、给她 / 他隐私、尊重不同的口味和习惯），这些只是其中一少部分建议的行动。不论是主动还是被动，努力或者说牺牲的重要性以及对情感关系的益处已经被人们广泛的注意到。

一个简化的第二定律可以用感情变量和努力变量组成的一个微分方程来表示

$$\frac{dx(t)}{dt} = -rx(t) + ac(t), \text{ for } t \geq 0, \quad (1)$$

这里 $r > 0, a > 0$ 。在没有介入的情况下（也就是 $c(t) = 0$ ），方程 (1) 意味着 $x(t)$ 以 r 的速度递减，这个速度 r 与每对配偶有关，用来表示这对配偶情感退化的力量。这个为人所熟知的简单的线性法则描述了很多自然界和社会现象。事实上，它的一个离散化版本在文献中被用来描述未受影响的配偶在短时间内的婚姻互动中的最初演化过程。对于任何速率，方程 (1) 中 $c(t) = 0$ 是对感情退化一个明显的假设。在方程 (1) 中“努力”被用来抵消这种感情退化。很明显，其中的参数 a 表示的是努力的效率。选定一个努力策略 $c(t)$ ，方程 (1) 就给出了情感的演化过程。方程 (1) 隐含地表明 $x(t)$ 是光滑变化的，除非在努力是不连续的情况下。

付出的努力 $c(t)$ 的强度由配偶对来决定，这与（非线性）变量 $x(t)$ 的水平不同， $x(t)$ 不能由配偶决定。努力变量 $c(t)$ 的理性特质允许我们将它解释为最优控制理论里的一个控制变量。在这一假设下，受控变量，也就是状态变量，是 $x(t)$ ，并且方程 (1) 是联系两个变量的状态方程。

我们接下来也是最后一个假设指的是对于努力和情感水平的成本 - 收益评估。我们考虑的是一个标准的效用性的途径。对于情感演化的感觉的数学表达是较为直接的（参见下面的假设 3）。然而，对于努力评估的表述却需要一番考量。最典型的努力是牺牲，为了亲密关系而忘掉自身利益，这种潜在的收益和成本已反复地在文献中被考虑过。对于牺牲和相关实践的实证研究表明，付出的努力可能同时蕴含着情感成本及其收益。这一明显的矛盾通过对牺牲行为的动机分析得以消弭，而这一动机分析则基于积极处理与刻意回避的不同态度。选择试图取悦伴侣的态度可能会产生积极的情绪影响，而刻意回避则可能引发紧张与痛苦。我们对于付出努力造成的情感差距的解释与努力的强度有关，因为我们认为在一定水平上的努力会获得感情回报，但是超出这一水平就会付出代价（痛苦）。这些形成了下面的假设。

假设 3（效用结构） 有两个不同的效用来源。一个来自于情感水平，另一个是努力强度的后果。

- i) 来自感情的效用可以用一个可微函数 $U(x)$ 来表示， $U(x)$ 满足 $U'(x) > 0$, $U''(x) < 0$ ，当 $x \rightarrow \infty$ 时， $U(x) \rightarrow 0$ 。换句话说，对于任何感情水平，它的边际效用是正的，但是随着感情越强边际效用递减到 0。
- ii) 努力 $c \geq 0$ 的负效用用一个可微函数 $D(c)$ 来表示，它满足 $D''(c) > 0$ ，存在一个 $c^* \geq 0$ 使得 $D'(c^*) = 0$ ，并且当 $c \rightarrow \infty$ 时， $D'(c) \rightarrow \infty$ 。也就是说，努力引发的不满在 c^* 达到一个最低水平，并且随着努力水平的不断增加，对婚姻的不满足会升高到无限大。

注意我们并不需要 U 和 D 的数学表达式。这个理论对于满足以上性质的所有函数都是有效的。

效用这一术语可以与幸福感或者对生活的满足感互换。i) 的假设是消耗一些美好的事物所达到的效用的一个标准。定义在感情上的效用并非一定是一种上层结构：当 x （一个人的感情）与一段关系的情感有关系的时候， $U(x)$ 就产生一个关于这个感情水平 x 的评估，这个评估依据的是个体的判断并或许依赖于以往的经验以及人格特质。比如说，两对配偶对于类似的感情水平或许给予很不相同的评价，所以他们的评估会基于不同的效用函数。感情效用

函数的存在性这一假设可以像基于消费的效用的论证方式一样来进行明确合理的探讨。

函数 D 表达的是负效用，它基于的是额外付出的努力会导致对效用的一种消耗。因此它的负数 $(-D)$ 可以理解作为一种效用。图 1 给出了这两个函数的一个典型的形状。

在这个模型的动力学假设中， U 和 D 是用来测量瞬间（也就是此时此刻）感情和努力的效用和负效用。 D 是非单调的假设反映的是努力在一定的水平之下会自我觉得是值

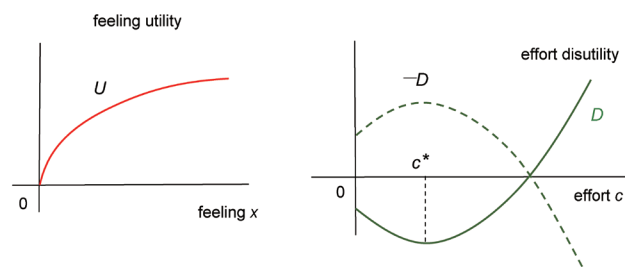


图 1 效用结构：效用与负效用函数的典型形状

得的这一事实。举例来说，计划与配偶的一次消遣活动，这不需要太大的努力并将产生一定的愉悦心情。尽管此刻的努力所达到的将来的益处隐含在感情效用中，因为此刻的努力为的是通过方程 (1) 来提升将来的感觉，如果 D 一直是非递减的，那么此刻努力所达到的即时的益处就没有被考虑到。

当努力水平很低的时候，做出小努力看起来是愉快的，然而当努力水平较高的时候，努力则会付出较大的情感代价。所以在假设 3 的 ii)，当努力在达到水平 c^* 之前所造成的效用是随着努力水平增加的，而在这个水平 c^* 之后随着努力水平的增加，效用却是减少的。这个参数 c^* 代表的是一对配偶的一个先验的最佳努力水平，它在分析中扮演着关键角色。这个理论也允许 D 是一直单调的特殊情况，也就是 $c^* = 0$ 的时候。这对应的是，即使努力水平最低，（此刻的）努力也会导致（此刻的）烦恼这一情形。所建议的 D 的结构似乎允许更加合理的情形。

一对配偶面临的问题就是如何制定一个努力策略，使得他们的关系持久并给双方最大限度的满足感。因此努力水平的演化是由追求最大幸福感的一个理想标准来决定的。这是一个优化问题，可以作如下地描述

问题 (P) 情感动力学的努力控制问题： 假设感情的演化过程由方程 (1) 来描述，效用结构由假设 3 来描述，最初的感觉水平 $x(0) = x_0 \gg 1$ ，用 $\rho > 0$ 表示不耐烦的因子。在这些条件下，寻找一个努力计划， $c^\vee(t) \geq 0$ 对于 $t \geq 0$ ，使得净效用最大并且双方的感觉和努力的演化是长期可持续的。

总的满意度是 $t \geq 0$ 所有瞬间净效用的总和，它可以用一种标准的方式来表达

$$W = \int_0^{\infty} e^{-\rho t} (U(x(t)) - D(c(t))) dt.$$

(指数项对应于对将来效用的一个评估。) 问题 (P) 是一个标准

的无限水平最优控制问题。由于断言 2，这一问题的计划时期设定为无限长。持久性的问题，作为配偶问题的一个关键要求，关注点有两个：可容许性与可行性。不仅长期的感觉和努力水平一定要在可容许的范围内（也就是说感情必须在水平 x_{min} 之上），而且向这些渐进水平的过渡一定要是可行的。

结果和讨论

这个模型所蕴含的主要结论在以下导出并加以评论。引人注目的是，断言 3 和断言 4 中所述的一些实证证据将由这个模型的分析理论地推导出来。并且，在这一模型框架下，断言 1 和断言 2 是相容的，这在一定程度上解决了失败悖论。分析的数学细节在附录 1 中。时刻 t 的最优努力必须满足

$$\frac{dc}{dt}(t) = \frac{1}{D''(c)} ((r + \rho)D'(c) - aU'(x)), \quad t \geq 0. \quad (2)$$

方程 (2) 给出了最优努力的变化法则。方程 (1) 和方程 (2) 组成了一个描述最优的感情和努力水平演化轨迹的方程组。这个最优轨迹记为 $(x^\heartsuit(t), c^\heartsuit(t))$ 。

情感平衡

如果方程 (1) 和 (2) 的稳态解是可行的，那么基于一套固定的努力程序，则可保证达成长久幸福的感情生活。享受永久的情感回报，却无需担心付出努力所引起的情感波折，显然是情感动力学一个吸引人的特点。这使得平衡态成为一个维系长久关系的理想结构。

存在性和可行性 平衡态由方程 (1) 和方程 (2) 中关于时间的导数为 0 来描述。基于对该模型的详述，在附录 1 中证明了存在唯一的明确定义的情感平衡状态 $E = (x_s^\heartsuit, c_s^\heartsuit)$ ，该状态描绘在图 2 中。

x_s^\heartsuit 位于 x_{min} 上方，因此这个平衡态是一个可接受的解。模型分析的一个关键发现是平衡的努力水平 c_s^\heartsuit 位于 c^* 上方(参看附录 1)，如图 2 所示。这是一个很重要的推论，也就是额外的努力 $c_s^\heartsuit - c^* > 0$ 对于维持长久的情感动力系统平衡是必要的。如果这对配偶认为这个努力差距 $c_s^\heartsuit - c^* > 0$ 并不需付出太大代价，那么这个平衡解就是可行的。当 $x_s^\heartsuit - x_{min} > 0$ 是可接受的并且努力差距 $c_s^\heartsuit - c^* > 0$ 是舒适的时候，一段关系就处于平衡状态。只要执行固定的努力计划 $c(t) = c_s^\heartsuit$ ，关系就会长期处于平衡状态。因为这一水平是问题 (P) 以 $E = (x_s^\heartsuit, c_s^\heartsuit)$ 为初始值的唯一解，所

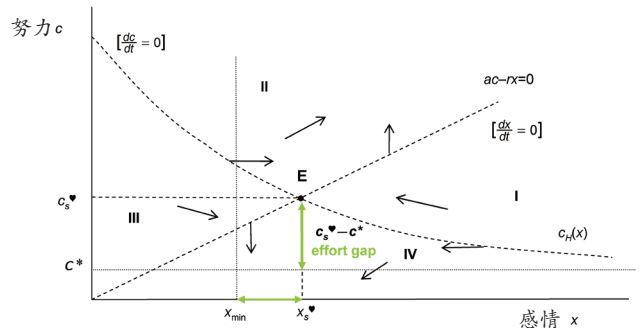


图 2 情感平衡态

以就达到了最大的幸福感。然而努力差距的存在或许会导致平衡解的不可行性。

不稳定性 一个基本的问题是扰动会随着时间消失还是放大。如果扰动会被放大，那么系统就是不稳定的。稳定性会促成一段长久稳固的关系，而不稳定性则是一个严重的缺陷。在不稳定的情况下，小的冲击，通常归咎于较低的努力水平，将会使得感情 - 努力状态远离平衡态。如果没有介入，扰动造成的最终的命运是情感关系的破裂。这将会在整体情感动力系统分析中予以明确。方程 (1) 和 (2) 的平衡解被证明是不稳定(参见附录 1)。此外，平衡态的局部动力系统属于鞍点类型。这对于整体动力系统有很重要的影响。一个可行的但是不稳定的情感平衡状态在理论上是可持续的，这需要配偶留心由于低于平衡努力水平所造成的扰动，并及时的增加额外努力使得系统恢复到平衡状态。

情感动力学和分手机制

通常来说，情感关系的最初状态并不处于平衡点，这是由于最初对对方的感情一般的来说远远地高于平衡水平 x_s^\heartsuit 。所以我们的讨论必须基于在初始感情 $x_0 \gg x_s^\heartsuit$ 的情况下考察动力系统 (1) 和 (2)。我们需要考察相空间的整体构造来解释向平衡点的过渡动态。

整体情感动力系统 图片 3 是感情 - 努力的相空间的一幅定性的图片，使用标准技术获取(参看附录 1)。这个图

片对于所有满足假设 3 的效用和负效用函数都是近似有效的。这个动力系统的构造是一个非线性的鞍点。定向曲线代表的是最优路径（每条路径都使得满足一定初始和最终条件的总的净效用达到最大）。

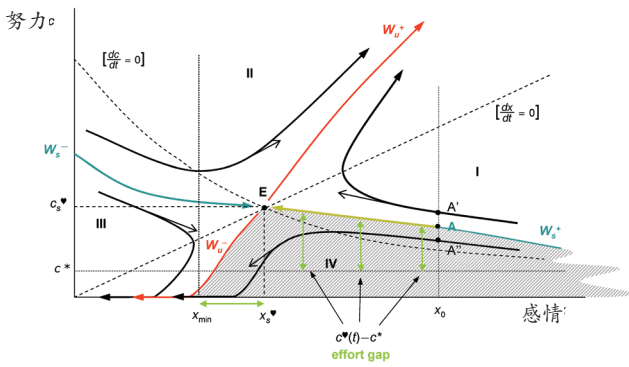


图 3 长久关系

稳定和不稳定流形，由所有接近和远离平衡点的点 (x, c) 组成，在图 3 中由曲线 W_s 和 W_u 表示。一旦轨迹到达 x -轴，在这和之后的努力必须设定在 $c^\heartsuit(t) = 0$ （参见附录 1）。因此，到达线 $c = 0$ 的一条轨迹按照

$$\frac{d}{dt}x(t) = -rx(t)$$

沿着 x -轴衰退到 $x = 0$ 。

长久关系的瞬时动力学 关键问题在于，给定一个初始的感觉 x_0 ，是否存在一个努力策略可以达到平衡态，如果存在，这个努力策略的特点是什么。稳定流形就是使得轨迹达到平衡点的唯一的曲线。任何其它的轨迹要么是 unacceptable 的，要么对应于非长久的关系。事实上，位于 W_s 上方区域 I 和 II 中（见图 3）的轨迹是不可接受的，因为它们会导致越来越高的努力水平并最终导致难以忍受的负效用水平（参见附录 1）。另一方面，在有限时间内到达 x -轴的轨迹会导致此后放弃努力并最终达到的感情水平只能是以结束关系收场。给定 x_0 ，存在一个合适的水平 c_0^\heartsuit ，使得点 $A = (x_0, c_0^\heartsuit)$ 位于 W_s^+ 上演变到平衡点 E（见图 3）。只要 x_s^\heartsuit 大于 x_{min} ，这个目标轨迹 AE 就是成功维系永恒情感关系的（唯一）秘诀。因为目标轨迹 AE 嵌入在 W_s^+ 上，而整个处于 $c = c^*$ 之上，所以在通向平衡点的过程中必须付出一些额外的努力（大过 c^* ）。因此，最优轨迹 AE 能够成功需要两个必要的条件，即这种感情盈余 $x_s^\heartsuit - x_{min} > 0$ 必须是有益的并且努力差距 $c^\heartsuit(t) - c^*$ 在所有的 $t \geq 0$ 时刻都必须是可以忍受的。沿此目标路径存在的这一努力差距或许是配偶关系破裂的源头，因为在许多情况下它或许是难以忍受的。

幸福感降低 因为目标路径位于区域 I 中，在达到平衡

值之前的路径上 $c^\heartsuit(t)$ 递增而 $x^\heartsuit(t)$ 递减。由假设 3 和链式法则可以推导出

$$\frac{d}{dt}(U(x) - D(c)) = U'(x)\frac{dx}{dt} - D'(c)\frac{dc}{dt} < 0$$

这意味着幸福感沿着最优路径 AE 一直递减直到达到 E 点。这一模型的理论预测与第 2 节中陈述的断言 4 一致。

分手机制 如上解释，典型的动态出现在图片 3 中的阴影区域，因为 x_0 很大并且轨迹导致的努力水平的增加是貌似不合理的。沿着阴影区域的轨迹，努力会最终降低直到撞到 x -轴并最终设定为 $c^\heartsuit(t) = 0$ 。这使得情感关系的维持只能是短暂的。因为不稳定性——由努力的减少引起的——原本沿着 W_s^+ 一条轨迹的偏离会使得系统状态落在阴影区域，而那里的最优路径偏离了目标曲线。这一关键性的特征是情感不稳定的主要源头。

由于努力的疏忽，导致情感关系日渐恶化的一个可能的

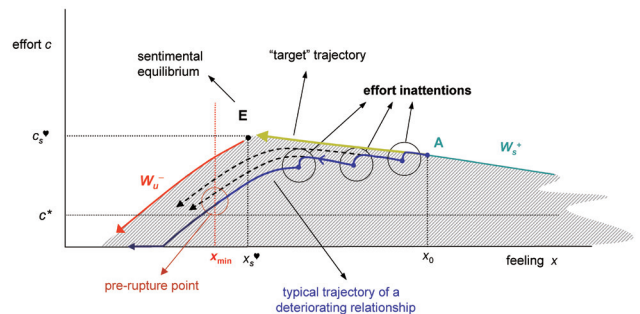


图 4 分手机制

机制在图片 4 中展示出来，并且可以进行如下地描述。假设关系起始于状态 A 点并且沿着位于稳定分支 W_s^+ 上的目标轨迹 AE 进行发展。如果在某点处发生了努力上的疏忽，也就是说如果努力水平降低了，状态就偏离了 W_s^+ 。如果努力不回到正确的水平并且系统遵循最优的动力系统方程 (1) 和 (2)，新的偏离状态就会作为一个新的初始条件，它导致的轨迹会逐渐偏离目标轨迹。这个新的轨迹可能会持续一段时间，直到新的努力疏忽出现，驱使状态处于更低努力水平的一个新位置，并沿着一个新的、更加远离 W_s^+ 的退化轨迹运行。经过一系列的努力的疏忽，不稳定性导致退化轨迹穿过阈值水平 x_{min} （参看图 4）。这是一个预破裂点，因为感情低于满意水平并且放弃努力只是时间问题。关系或许会持续一段时间但最终会突破承受的极限。在这一最后阶段，情感依恋逐渐消失，而这与专业文献中对离婚现象的描述基本相符。

如果系统沿着一条退化轨迹，通过提高努力水平仍然可以回到最优路径上。然而，做出反应及纠正偏差的耗时越长，状态就离目标轨迹越远，回到目标轨迹的困难就越大。如果

努力被长期忽视，这或许变得不可逆转。相当数量的被报道的不幸婚姻似乎符合这一判断。上述的恶化过程与第2节所

述的断言3一致。

总结

本文的数学理论给出了一个机制，用来解释在计划长久维持的情感关系中大量出现的恶化和破裂。两种力量一起导致了恶化过程的出现。首先，因为在成功路径上为了持久的关系必须付出额外的努力，假使这个差距是令人不舒服的，伴侣或许会放松并降低努力水平。接着，不稳定性出场，使得感情-努力状态偏离持久成功的动态。

一个更重要的发现是这一事实——伴侣建立关系并视其为明确目标与证据显示的他们的关系有可能破裂其实是相容的——这在模型中是很典型的动态过程。这消除了失败悖论，配偶分手的可能性是由于最佳选择下第二定律起作用的后果。

模型的分析或许可以给伴侣提供如何保持长期关系的建议。持久的关系只有在努力差距是可以接受的情况下才是可能的，最佳的努力策略是持续的关注使得其位于目标动态上，或许可以描述为在稳定分支附近的一条轨迹，并徘徊于平衡状态附近，时刻留心让努力保持在正确的水平上。这些

类型的关系是常见的，虽然还有一些例外。这与模型中持久成功关系的例外是一致的。

两个明显的事实用来验证本文提出的理论：(i) 模型建立在被接纳的证据之上（亦即，第二定律和配偶打算长久相处的意向）以及(ii) 模型的机制显示的与离婚及分手的实证事实相一致，这些事实也就是衰退关系中典型的逐渐恶化（第2部分中的断言3）和婚后幸福感的降低（第2部分中的断言4）。对于验证模型的进一步研究要——通过实验室试验或者实地调查——测试这个理论的两个主要发现，亦即努力差距的存在性和感情-努力动力系统的不稳定性。

对于配偶持久性的悲观结论应该也适用于并非理想的情景，只要第二定律的提法仍被认为是有效的。更一般的假设，如（弱）不同型配偶、外部冲击的出现或者次优的行为，或许可以作为加强不稳定性的因素来加以考虑。努力差距和揭示出来的不稳定性构成可能的情感失败的基本内在机制。

附录 I： 本文理论分析的相关数学推导。详见 <http://www.plosone.org/article/fetchSingleRepresentation.action?uri=info:doi/10.1371/journal.pone.0009881.s001>



作者简介：何塞-曼纽尔·雷 (José-Manuel Rey)，从马德里获得数学博士学位，在圣安德鲁斯大学和伦敦大学学院做过博士后研究，现为西班牙马德里孔普卢顿大学 (Universidad Complutense in Madrid) 教授。



译者简介：李玉田，吉林大学数学系毕业，香港城市大学博士，香港浸会大学助理教授。