



中科院计算所
INSTITUTE OF COMPUTING TECHNOLOGY, CAS

对信息领域科研规划的一些思考

李国杰

2009.08.22

前言

- 最近一年多参加科学院的《中国至2050年科学技术发展路线图战略研究》，对信息科学技术的长期发展趋势、15-20年内信息技术发展的重大障碍、我国全面进入信息社会的战略需求以及我国信息领域科技布局存在的问题等方面有一些体会与思考，希望与大家讨论，得到诸位专家的批评指正。
- 我国信息领域面临的最大挑战是从技术的**追随者**变成国际上信息技术**创新行列的重要成员**，863计划如果不能为此做出重要贡献，可能会被边沿化。



中科院计算所
INSTITUTE OF COMPUTING TECHNOLOGY, CAS

对信息科学技术长期发展趋势 和我国战略需求的基本判断

今后20-30年是 信息科学技术的变革突破期

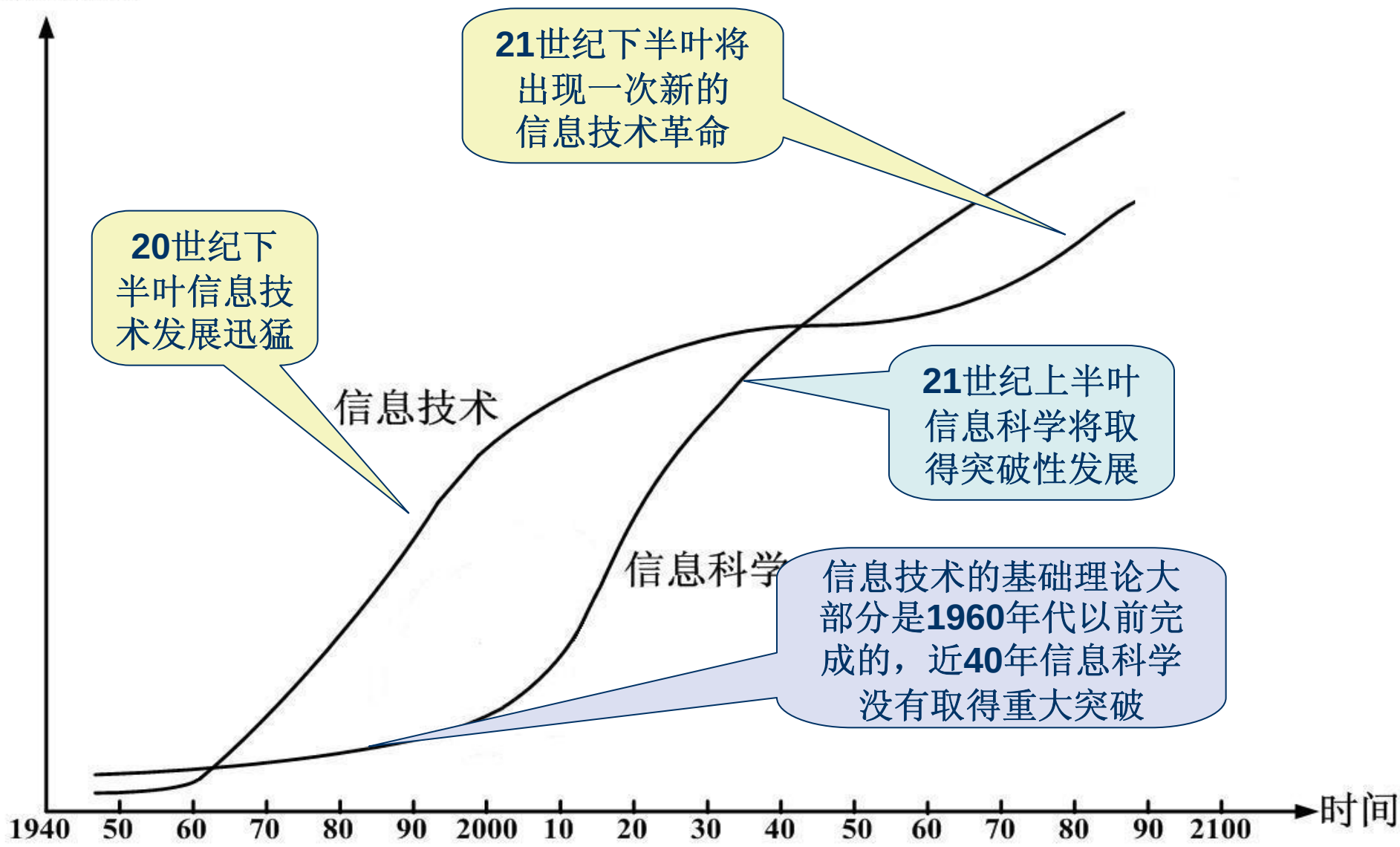
- 信息科学还是一门年轻的科学，过去的30年里，信息技术实际上走在信息科学的前面，许多重要的信息科学基本理论问题并没有解决。
- 近20年来集成电路和网络技术的飞速发展对信息科学提出了若干挑战性的理论问题，挑战问题将刺激信息科学突飞猛进地发展。
- 无论是集成电路、互联网、高性能计算机还是磁盘存储器，几乎所有现有的信息技术到2020年前后都会遇到靠渐进式的改进难以继续发展的重大障碍。2020到2040年期间必须在信息科学和信息器件、设备、软件上有原理性的重大突破。强烈的需求将激励信息科学技术在今后20—30年间有一次大的飞跃。

2020—2030年之间可能出现的“技术墙”

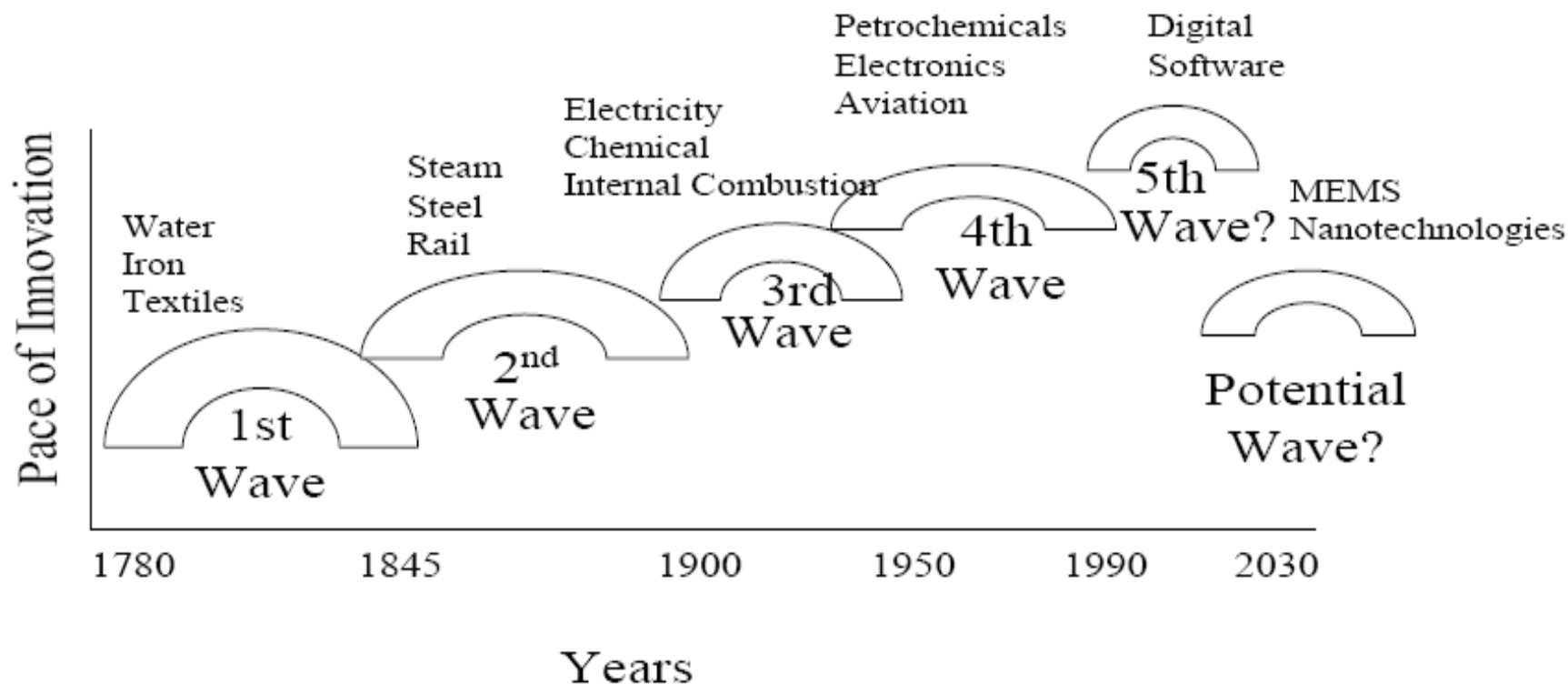
- 许多信息技术不约而同地将在2020-2030年之间出现难以逾越的障碍。
- 到2020年左右，摩尔定律将不再有效，集成电路正在逐步进入“**后摩尔时代**”，我们必须更多地从Beyond CMOS中寻找新的出路。
- 计算机正逐步进入“**后PC时代**”，终端设备将从“**高大全**”向“**低小专**”（“专”指个性化）转变，**降低功耗**是首要目标。
- 2020年以后，超级计算机的“**千倍定律**”将失效，只在现有的技术基础上做改进，2030年肯定做不出Zettaflops级（ **10^{21} flops**）水平的计算机。
- 进入“**后IP**”时代是不可避免的发展过程，可能需要20年时间才能真正突破TCP/IP协议的局限。

20-21世纪信息科学与技术发展态势示意图

发展程度



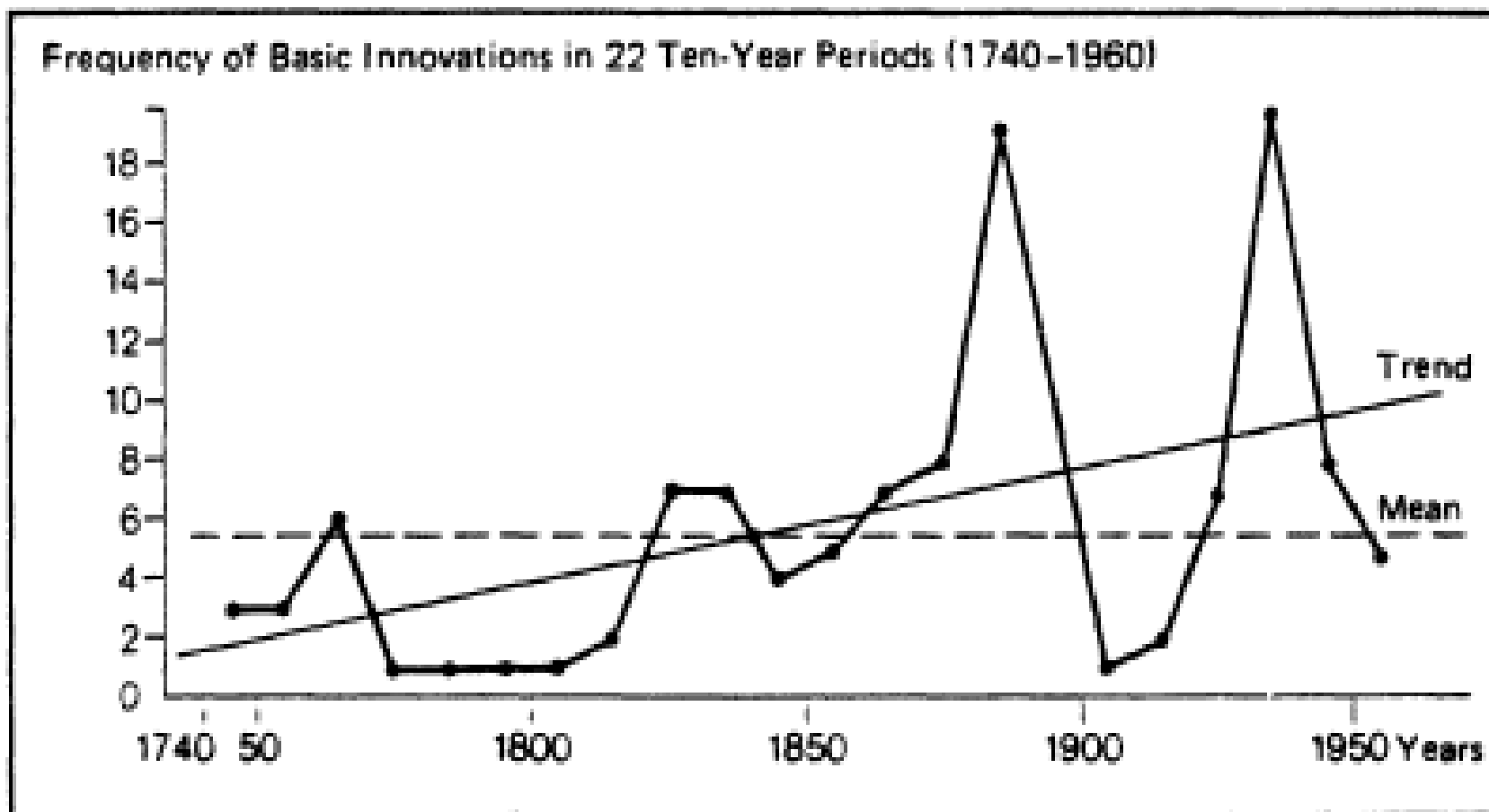
康德拉季耶夫经济长波示意图



R.. Fremann et al.. "Principle Concepts of Technology and Innovation Management", 2008

1740-1960年基本创新的频率

G. Mensch, 《Stalemate in Technology》 1979



技术创新模式的转变

工业化阶段	第一阶段	第二阶段	第三阶段
经济标志 人均GDP	小于300美元	300—4750美元	大于4750美元
技术标志 GERD/GDP	小于1%	1—2%	大于2%
技术创新阶段	使用技术为主	改进技术为主	创造技术为主

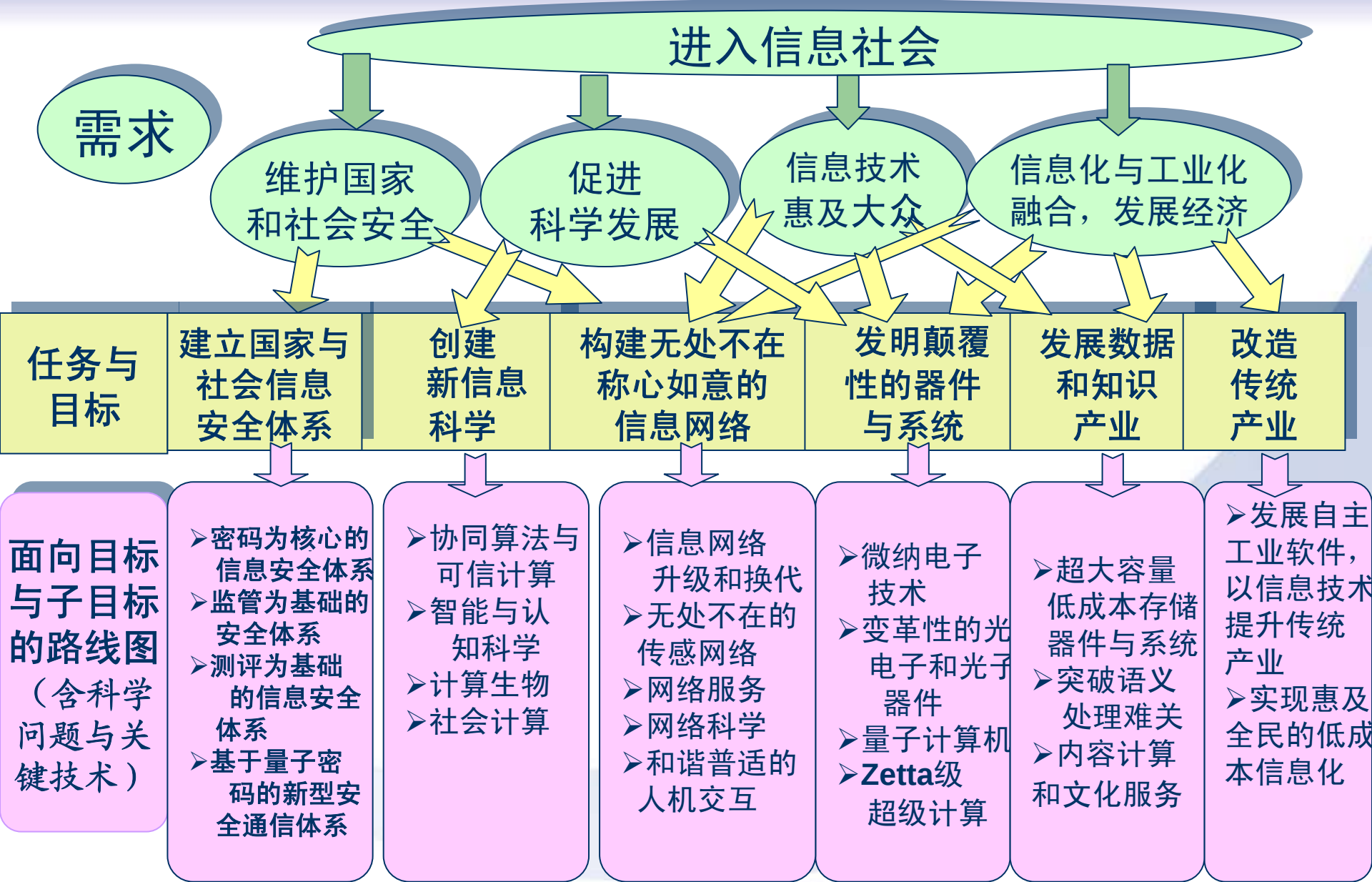
——引自穆荣平“中国技术资源开发与利用战略研究”报告

做出与国力相称的科技贡献

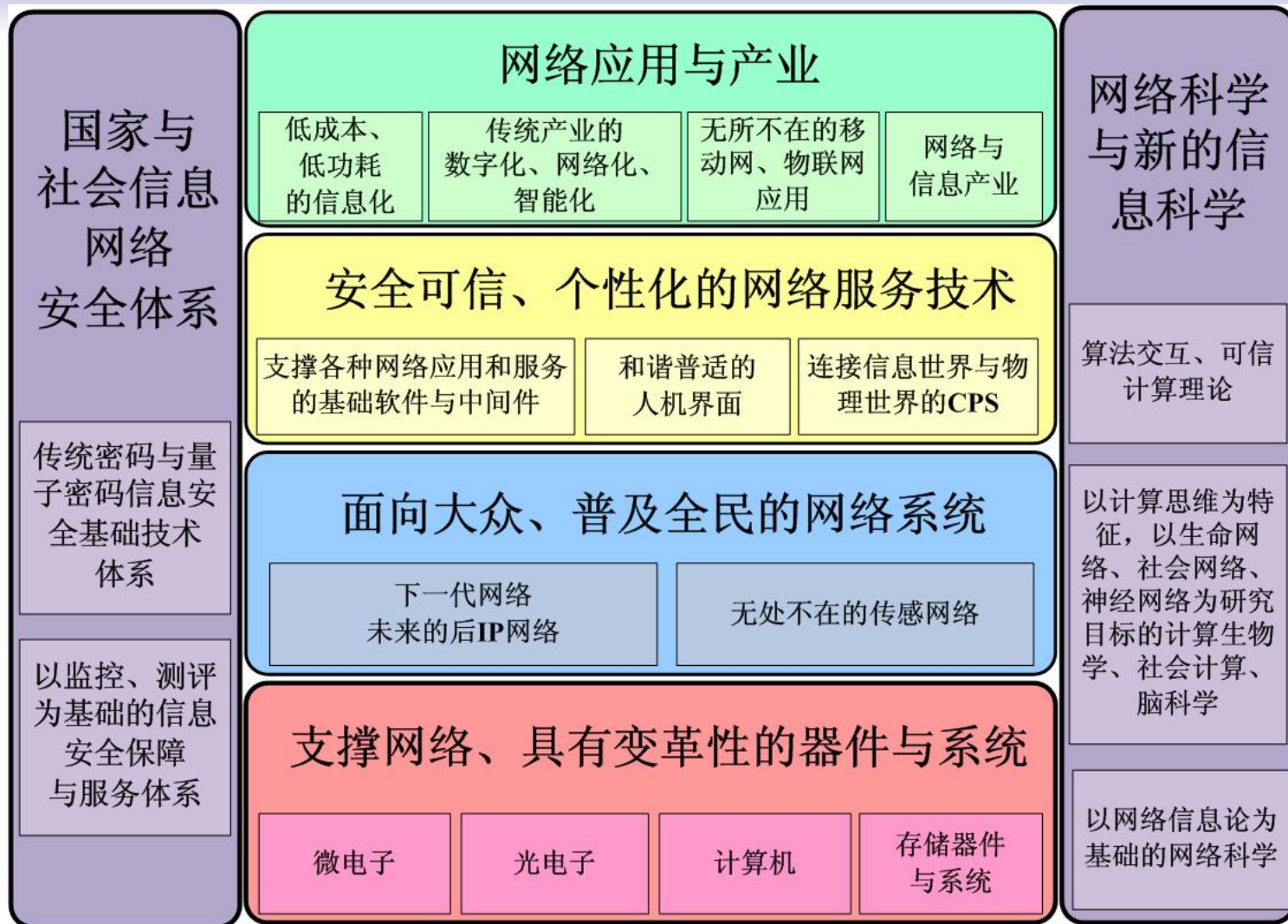
- 我国的人均GDP已经超过**3300美元**，深圳、上海、北京、广州、苏州、杭州、佛山等市的人均GDP已超过或接近**1万美元**。浙江、广东、山东、江苏等省的人均GDP也超过或接近**5000美元**。
- 多少年来，我们习惯于在国外的基础技术平台上做科研工作。信息领域过去30年没有认真考虑建立**自主可控的技术平台**，现在应想一想如何为建立这种平台而努力。
- 我国一定要争取对支持下一个经济长波的**基本创新**做出与国力相称的贡献，要致力于做“**改天换地**”的科研工作。
- 姚期智教授最近在《中国计算机学会通讯》上发表文章指出，我国大学争取世界一流只有最近10—15年的机遇期，15年以后，新的技术格局就会形成，我们就很难占一席之地了。



我国对信息技术的基本战略需求



普惠泛在的信息网络体系（U-INS）



这一体系体现了21世纪上半叶我国要全面进入信息社会的重大战略需求，也包含了信息领域需要重点发展的科学技术。



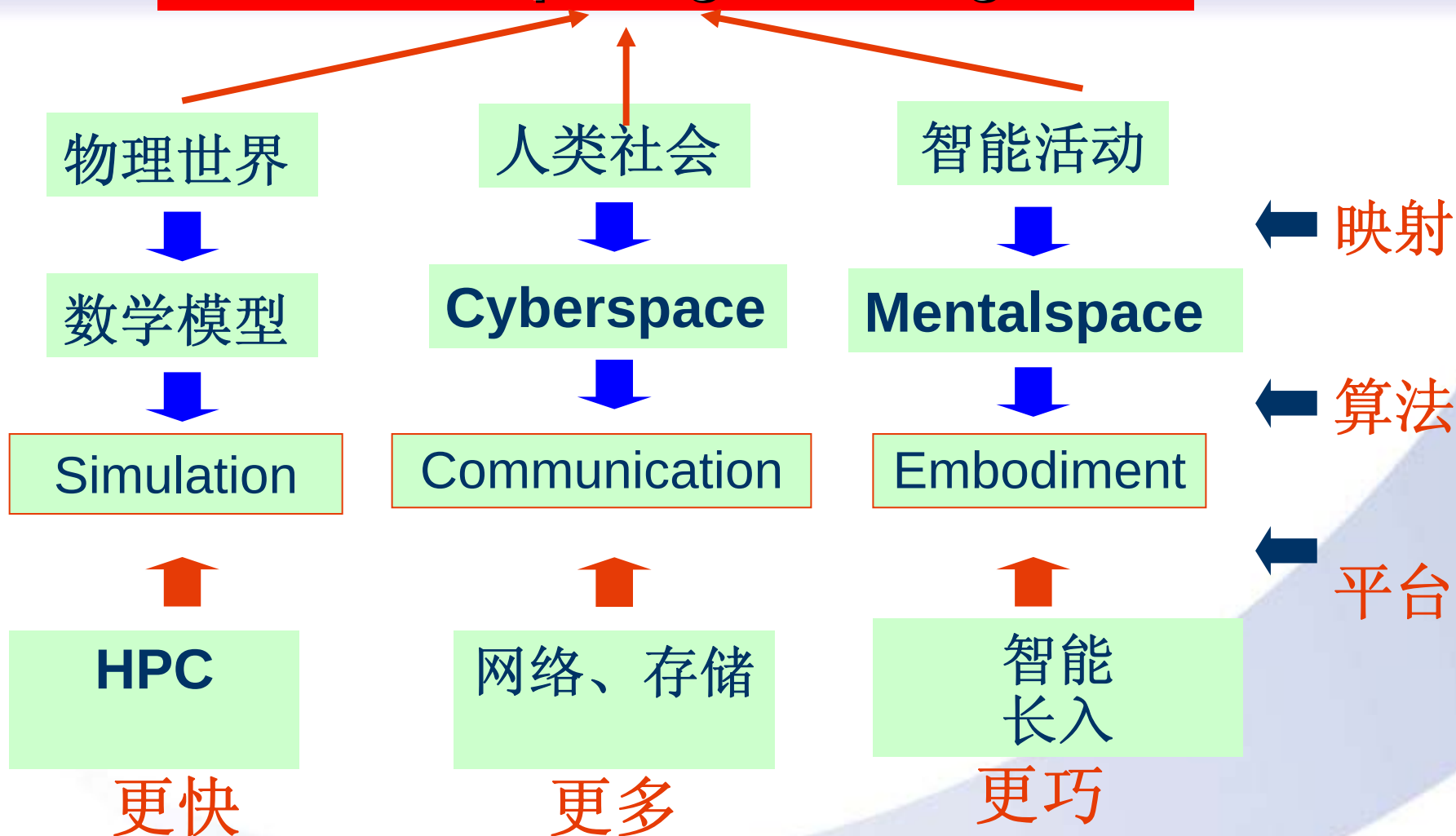
中科院计算所
INSTITUTE OF COMPUTING TECHNOLOGY, CAS

需要改变发展信息科技的某些观念

改变“狭义工具论”的旧观念

- 长期以来，计算机和信息网络被社会看成是一种高科技工具，计算机科学技术这门学科也被构造成一门专业性很强的工具学科。
- 这种社会认知很容易导致负面的“**狭义工具论**”，即认为信息科技只是一种高科技工具，“高科技”意味着认知门槛高、成本高，“工具”意味着它是一种辅助性学科，并不是能够满足国家经济社会发展、满足人民经济文化需求的主业。这种狭隘的认知是信息科技向各行各业渗透的最大障碍，对信息科技的全民普及极其有害。
- 需要在全社会传播和普及“**计算思维**”（Computational Thinking）

“一切皆可计算” Computing Thinking



—此图思想来自图灵奖得主Butler Lampson的报告

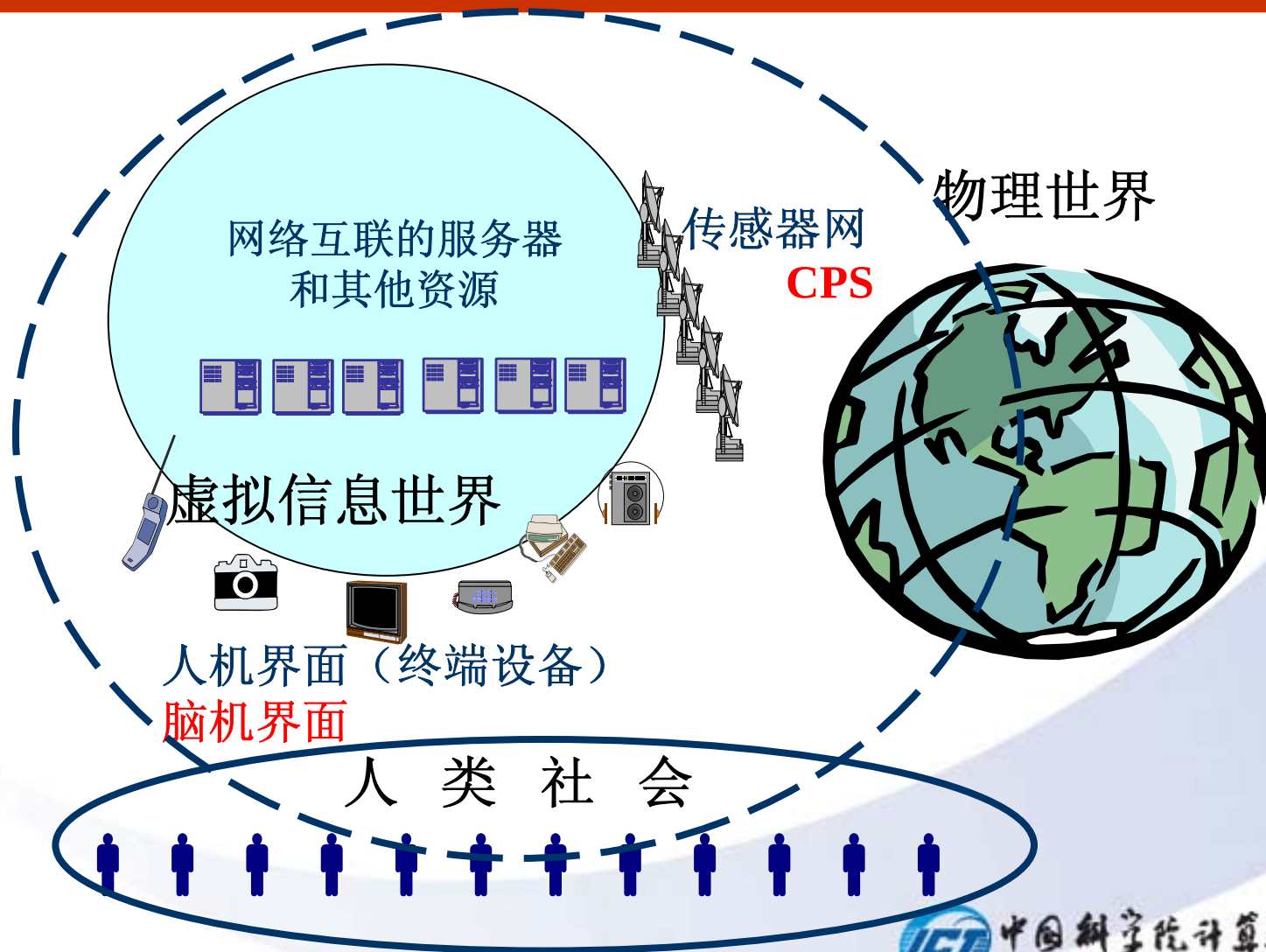


从人机共生模式转向人机物三元社会模式

- 今天使用的信息系统，在很大程度上仍然根基于40多年前提出的**人机共生**思想：人做直觉的（说不清的）、无意识的事，计算机做有意识的、确定的、机械性的操作。人确定目标和动机，计算机处理琐碎细节，执行预定流程。
- 今天的计算机世界已经与一人一机组成的、分工明确的人机共生系统不同。**人机物三元世界**是一个多人、多机、多物组成的动态开放的网络社会。
- 这个跃变促使信息科学发生本质性的变化。信息科学应当成为研究人机物社会中的信息处理过程。我们需要回答下述基本问题：万维网能被看成一台计算机系统吗？什么是万维网的可计算性？什么是物联网计算机的指令集？人机物社会中的“计算”如何定义？它还是图灵计算吗？
- 为了研究人机物三元世界的计算问题，传统算法科学的**集中式假设、确定起始假设、机械执行假设、精确结果假设**等可能都需要突破。



物理世界、信息世界、人类社会 组成三元世界—新信息世界观



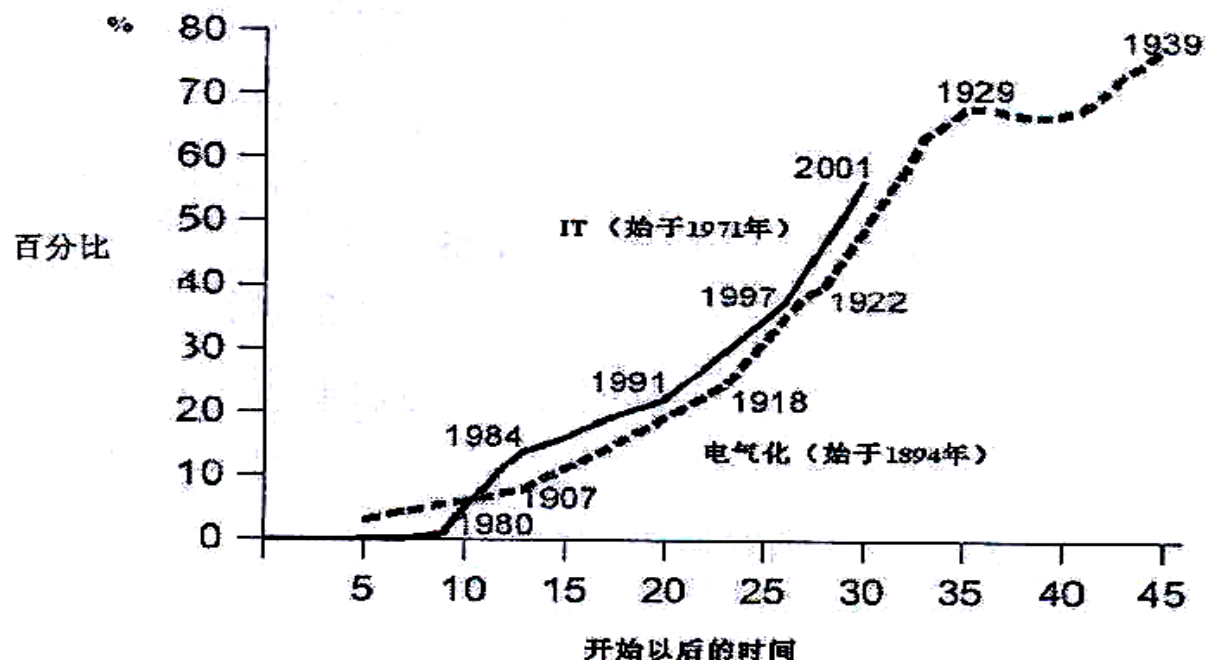
改变对低成本信息化的错误看法

- 从历史的长周期来衡量，电脑普及的速度和电力技术普及的速度差不多，要高度重视信息化普及的难度。
- 低成本信息化不是以降低实效、降低价值为代价，价值要与普及成正比。全民普及不仅意味着享受最低价值的用户增多，还意味着享受更高价值的用户也要按一定比例增加。
- 我们的预测表明：只有增值增长才是高实效的低成本信息化路线，才能推动信息产业良性持续发展。
- 成本信息化需要真正的高技术。



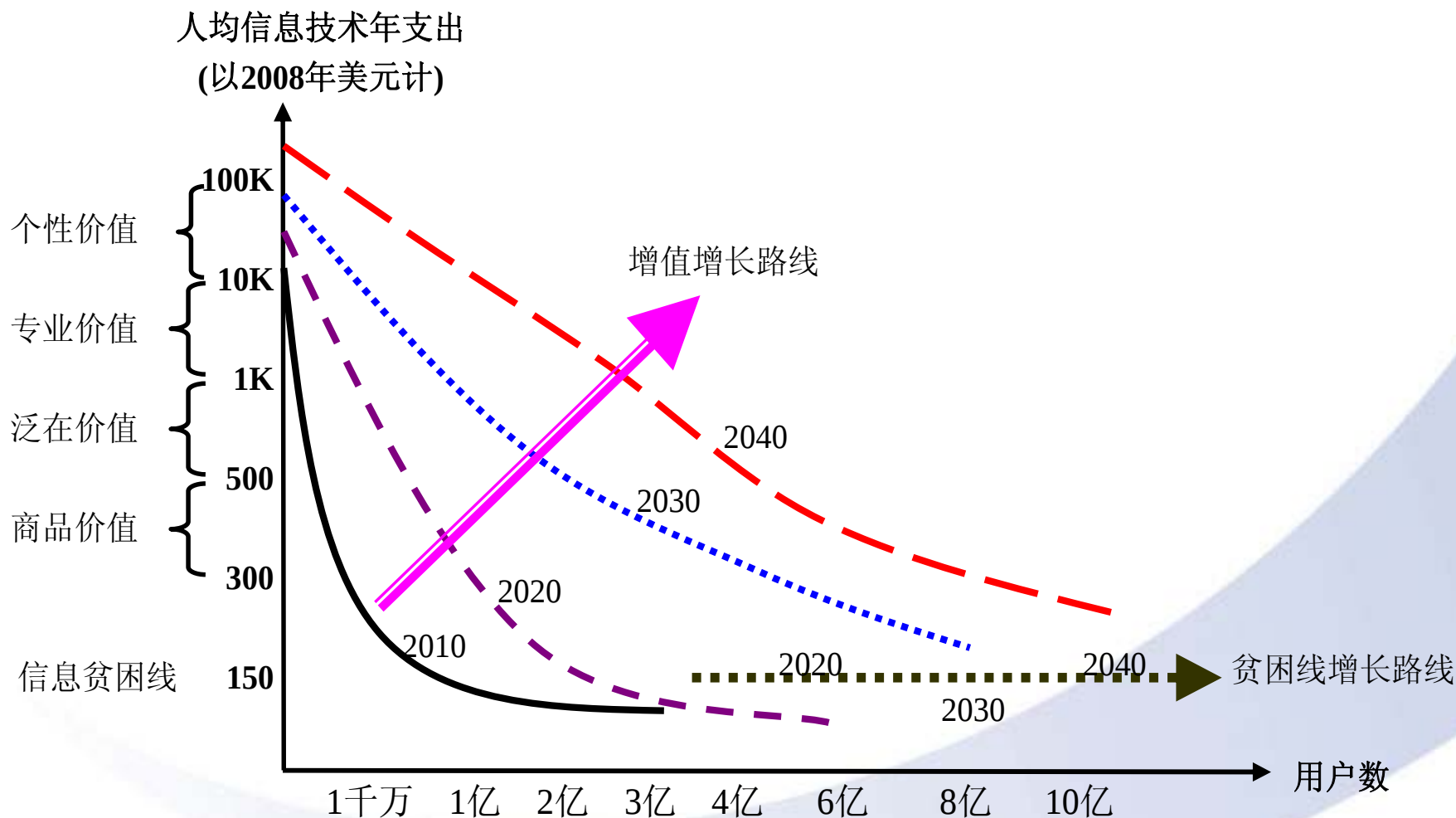
电脑普及与电气普及速度差不多

美国拥有电气和电脑的家庭比例:



资料来源: Jovanovic and Rousseau (2003), 通用科技

信息价值与普及必须同步增长

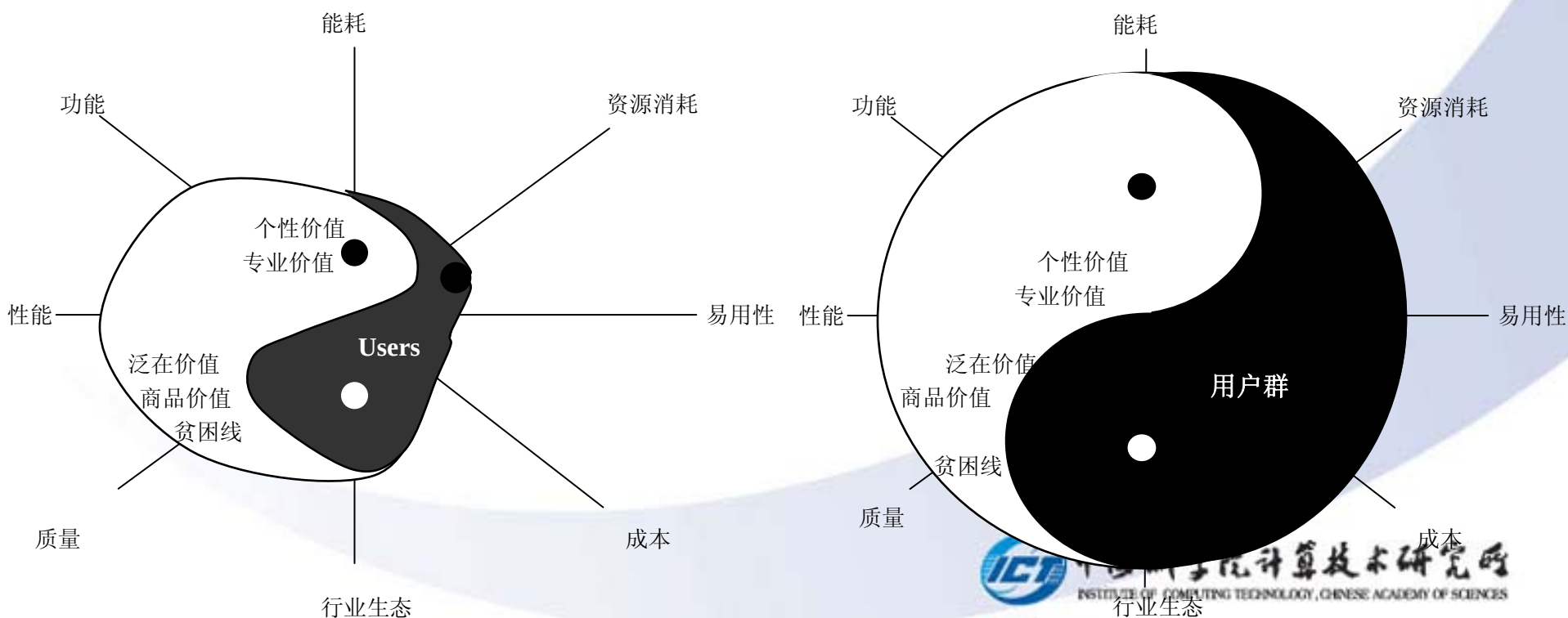


改变图灵计算模型不可突破的观念

- 目前的主流计算机科学教科书认为：图灵机不能做的事情将来的计算机也不能做。
- 图灵计算被看作是从输入到输出的函数，不终止的计算被认为是没有意义的。
- 在网络环境中，计算主体（进程）在与外界不断交互的过程中完成所指定的计算任务。对于这类交互式的并发计算，传统的基于‘函数’的计算理论不再适用。
- 如何为实际并发系统的设计与分析提供坚实的理论基础，在今后几十年内是计算机科学面临的重大挑战。
- 算法研究的重点应从单个算法的设计分析转向多个算法的交互与协同。

改变（平衡）科技创新的目标

- 功能、性能 → 成本、易用性、能耗、资源消耗
- 例子
 - 传统算法科学：研究和优化性能（时间复杂度）
 - 普惠计算科学：研究和优化**能耗复杂度、劳动复杂度**



改变重“支撑”轻“引领”的短视科技布局

- 中长期科技规划制定了十六字方针“**自主创新、重点跨越、支撑发展、引领未来**”，实际工作中往往只重视支撑发展，忽视重点跨越和引领未来，未来十年要纠正这一倾向。“**支撑**”与“**引领**”**两手都要硬**。
- 多少年来，我们总是认为自己底子薄、基础差，不具备创造技术的条件，习惯于在国外的基础技术平台上做科研工作。计算机领域基本上不敢跳出Wintel平台去思考创新。
- 今天不做引领性的科研，不重视“重点跨越”，明天就只能靠国外的技术支撑我国的产业。
- 支撑发展、重点跨越和引领未来可以理解为自主创新的三种不同层次。我们必须**统筹规划技术创新的不同模式**，不断提高技术创新的水平。



中科院计算所
INSTITUTE OF COMPUTING TECHNOLOGY, CAS

对未来10年863信息领域 总体思路的几点看法

(1) 863要做引领产业发展的前瞻性工作

- 企业的研发能力逐步提高，863若主要做3—5年内可产业化的技术研究，将失去引领作用。
- 研究的重点应逐步**向更前瞻方向移动**。十二五期间应安排足够的人力物力**做5-10年后才能产业化的前瞻研究**。
- 2020年以前的网络发展需要的技术基本上已提出，有些已纳入标准，如IPV6、LTE无线通信等。要重点研究泛在的IP后网络的理论基础和核心技术。

(2)应高度重视网络服务与应用技术研究

- 网络的升级换代不仅仅是基础设施的变革，网络服务与应用模式也会发生根本性变化，而且，新的服务与应用模式会影响基础设施的发展方向。
- 我国在器件与设备上落后于国外，但在适合国情的网络服务上不比国外差，腾讯、阿里巴巴、百度等都战胜了国外对手。这些公司的创新动力明显高于联想等设备制造公司，863要加大支持信息服务型公司的前瞻研究。
- 相对于P2P等分布式计算，云计算是更高层次的“分久必合”。云计算的核心技术是虚拟化技术，代表未来网络服务的发展方向，不完全是公司的炒作。863要重视云计算的关键技术研究。

云计算与User-oriented计算

- 云计算已成为当今最热的话题，其关键是资源集中和虚拟化技术，应当引起我们的重视。
- 云计算涉及国家信息基础设施的基本安全问题，不能掉以轻心。跟着国外大公司忽悠，可能完全受制于人。
- 网络信息技术的长远发展目标应该是真正以用户为中心，而不是以服务商为中心。变相的Client-Server结构或虚拟的Mainframe结构可能不是理想的结构。
- 信息不同于能量，信息的根本性质是可无限次共享而本身不减少，理想的信息服务模式可能不同于电力。
- 我们需要寻求符合信息本质规律、真正以用户为中心的网络体系结构。

(3)要重视变革性器件和系统研究

- 10-15年内信息器件和系统将遇到难以逾越的功耗、可靠性等技术墙，863要组织力量**从原理上突破低功耗、自修复等变革性技术**，发明新的器件，提出设计复杂信息系统的新思路。
- 我国纳米技术研究基本上没有做信息器件，863要促进信息与材料领域的交叉融合，大力加强Beyond CMOS的变革性器件和系统研究。
- 片上光互联是多核和众核芯片进一步发展的必然选择，应引起高度重视。

(4) 要持续支持高性能计算研究、重点突破海量信息处理技术

- 我国在高性能计算上有较好的基础，863要发挥这一优势，改变只重视以Linpack为目标的传统习惯，重点研制支持信息服务和云计算的“Warehouse Scale Computer”，重点支持海量数据挖掘、大规模信息获取和分析。使我国基于大规模数据的高精度仿真计算走到国际前列。
- 863的支持重点要从HPC转到HTC，即高吞吐量计算机（High Throughput Computer）



(5) 要强调跨学科合作， 设定让人望而生畏的高目标

- 863要设定一些让人乍一看认为不可能实现的高目标，开展旗舰式的重大任务研究。
- 重点支持跨学科研究，让信息科学技术成为提升其他学科研究水平的杠杆。

必须重视脑科学与信息技术的结合

- 国外的科学家普遍认为：“Neuro”已成为新的”Nano”,2009年是脑科学和神经工程学（**Neuroengineering**）历史上关键的一年。
- 神经工程学，脑反向工程是脑成像技术、神经系统成像技术、计算神经学等领域的交叉科学，近年来发展迅猛。
- **我国在经络研究上已落后于韩国。**
- 我国制定中长期规划时，**脑科学没有得到足够的重视**，应该与量子调控、蛋白质科学放在同样的地位。
- 国内各学科画地为牢，不希望其他学科打进自己的领地，严重影响交叉学科发展，未来十年要打破学科领域的限制，高度重视脑科学与信息科学的交叉融合。

新兴汇聚科学

- NBIC, (**N**anotechnology, **B**iotechnology, **I**nformation technology and **C**ognitive science)
- **GNR** (**G**enetics, **N**anotechnology, **R**obotics).
- **GRIN** (**G**enetic, **R**obotic, **I**nformation, and **N**ano processes)
- **GRAIN** (**G**enetics, **R**obotics, **A**rtificial **I**ntelligence, and **N**anotechnology)
- **BANG** (**B**its, **A**toms, **N**eurons, **G**enes)



中科院计算所
INSTITUTE OF COMPUTING TECHNOLOGY, CAS

信息领域今后10年 值得重视的若干课题（举例）

1、构建IP后（Post-IP）的新网络体系

- 今后10—40年发展信息技术的第一位任务是要建设让大众非常便捷地获取信息和知识、更有效地协同工作、生活更加高品质的信息网络。
- 近10年内网络技术经历宽带化、移动化和三网融合将走向基于IPv6的下一代互联网，2020年以后世界各国将逐步形成共识，共同构建**IP后（Post-IP）的新网络体系**。
- 美、欧、日、韩都在积极探索IP后的未来网络。
- 在改进现有互联网和构建重叠网的同时，我国必须尽快开展IP后网络系统结构和关键技术研究。

IP后网络与无处不在的传感器网络

- 互联网具有固有的缺陷，如**缺乏服务保障、可管理性差、安全性差等**，在保持互联网开放自治的前提下，需要研究**超出IP协议的未来网络**。移动网络的用户已超过固定的互联网用户，研究未来网络需要重点考虑**移动网用户的需求**。
- **传感网络**是数字世界与物理世界的桥梁，将实现人与机器，机器与机器之间的信息交互，**数字物理系统（CPS）**是数字世界与物理世界交互的网络系统，具有监视与控制功能。IBM公司大力提倡的“**智慧地球**”强调的就是数字物理系统（CPS）。传感器和CPS是典型的多学科交叉，涉及通信、光学、微机械、化学、生物等诸多领域。
- 研究IP后的未来网络不仅要考虑人与人通信，网页与网页互联，还要考虑物物通信（物联网）



实现人和服务的连接

- 网络要具有**SET特性**，即简单（**Simplicity**）、高效（**Efficiency**）、可信（**Trust**）。
- 从Networking of network 转变为Networking of service and people，体现以用户为中心的网
络构建原则，旨在实现**人和服务连接**而不仅仅是
机器的连接。
- 采用通用计算账号（**Universal Computing Account**）技术实现个人信息环境与信息终端和
网络服务松绑。
- 力争经过**15**年的努力，使我国在未来网络升级换代和向**U**社会的过渡中取得主动。

美国FIND计划致力于以下问题

- 网络是否要继续采用分组交换；
- 端对端原理是否要改变（端对端原理的提出者 Clark 教授在2009 的一次FIND工作会议上建议：将EtoE 原理改为trust-to-trust原理，即应用的智能不一定在终端，而是放在有足够的信任能正确完成任务的设备上）；
- 路由和包转发是否要分开；
- 拥塞控制和资源管理问题；
- 身份认证和路由问题。

2、艾级（Exa级， 10^{18} ）超级计算 和大规模数据处理技术研究

- 生命、环境、能源、材料、高能物理等基础科学研究对超级计算提出了强烈的需求，通过超级计算模拟生命现象、环境变化、仪器看不到的微观世界等，**为科学研究提供强大的计算仿真工具**，其作用如同粒子物理领域的对撞机。
- 从原理上突破**艾级计算（每秒 10^{18} 次）**的核心技术，为国家构建艾级计算系统提供技术保障。推动我国计算生物学、计算物理、计算化学、生物信息学、纳米信息学、脑科学与认知科学等众多新兴交叉科学的发展。使超级计算广泛应用于基因检测、新药、新材料、新能源发现，造福于全国人民。
- 部署HTC研究，研究适合云计算的Warehouse Scale Computer 系统结构和处理器芯片。



超级计算机和软件发展趋势

- 到2050年，超级服务器的发展需要支持各种各样的个性化应用负载，突破低能耗、海量并行、可靠性、低成本等技术障碍，40年内超级计算机的性能将增长 $10^8 - 10^9$ 倍，达到每秒 10^{24} 次运算速度。在这个进程中，重大难点和技术突破会发生在从 Exaflops (10^{18} flops) 过渡到 Zettaflops (10^{21} flops) 阶段。
- 发展信息技术的一个重要目标是使软件业和服务业也产生类似摩尔定律的走势，即同样功能和性能的软件开发成本平均每两年降低50%，同样质量的服务所需的成本每两年降低50%。

并行编程是必须突破的关键技术

- 新的摩尔定律变成：
 - 在基本不变的主频下，单个芯片上的处理器核的数目每一代（约两年）增加一倍。
- 单处理器性能的提高遇到阻碍以后，计算机要提高性能基本上只剩下一条路可走——依靠并行处理。这一迫不得已的转变，对软件界既是挑战也是机遇。
- 转向增加并行性不是基于重大突破的高歌猛进，而实际上是由于传统硅实现方法遇到巨大障碍而采取的一次退却。
- 今后几十年中，算法研究也将从单一算法转移到对多个算法相互影响的研究，为研究分布式系统奠定坚实的基础。并行算法和并行编程是今后几十年计算机科学和软件界必须突破的科学问题和关键技术。。



3、低成本低功耗信息器件和系统 核心技术与低成本信息化示范

- 到2020年，器件**性能功耗比要提高100倍**，CPU芯片性能达到每秒10万亿次运算，功耗低于10瓦。**计算机系统性能功耗比提高1000倍**，超级服务器性能达到艾级（每秒 **10^{18}** 次运算），功耗低于10兆瓦。
- 低成本医疗设备和远程医疗示范。基于上述技术，为**1亿人提供低成本医疗服务**，包括前端的体检和后端的海量（100 PB）医疗信息库与个性化医疗信息网络服务。
- 为“核高基”重大专项提供低成本低功耗前瞻技术突破，为建立我国自主可控、激励自主创新、跨越国外垄断的**计算机产业基础平台**奠定科技基础，使我国在国际计算机科学和核心技术创新中占有一席之地。



4、探索变革性的新器件

- 传统的信息器件和设备在复杂性、成本、功耗等方面已遇到巨大障碍，**急切期待颠覆性的新技术**。
- 目前尚未确定一条像近30年CMOS集成电路一样的主导技术路线，量子、自旋、纳米等技术发展呈现出不确定性和多样性，**确定新的主流器件技术可能需要15-20年的努力**。
- **石墨烯**纳米带晶体管可能成为延续摩尔定律的重要推动力，又可能成为超越硅基CMOS的很有希望的研究方向。
- 电子计算技术和光电子、光计算技术的融合最有可能成为未来开发汇集计算、存储、通信和信息处理于一体的新一代芯片技术，可实现**片上光互联和片上大规模光计算**。
- 我们不但要借助纳米、超导等变革性技术发展微电子、光电子和光子器件，还必须从**计算模型和计算机系统结构**层面考虑如何有效地利用量子、生物等新兴技术。



5、网络信息安全理论与技术

- 网络的开放性与安全性是一对矛盾。越开放的网络越需要安全保证。安全技术渗透到网络技术每一个层次，从器件、设备到系统软件和应用软件。**网络安全是一个体系**，除了传统密码和量子密码等基础安全机制，还需要考虑监控和评测。
- 网络安全涉及一个国家的根本利益，为了防御网络攻击，需要**网络“沙盘”或“靶场”**来分析国家网络安全态势。

6、人机交互和脑机交互

- 人机交互是计算机科学/工程研究的核心之一。在未来几十年中多模态人机交互将占据桌面、膝上和掌上系统，三维用户交互、实体交互、可交流情感的个性化交互、**脑机交互**将得到普及。在人机交互研究领域，自然语言理解、图像语义理解是需要长期攻克科学难题和关键技术。
- **人机交互**是人类世界与信息世界的界面。到2020年，手势、语言为代表的多模态交互将成为主流。普适交互环境的关键科学问题包括环境模型、主动察觉感知技术、情景感知等。

7、发展语义互联网

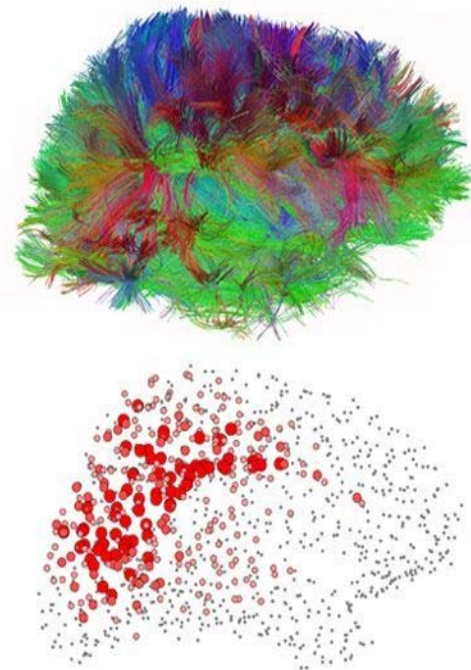
- 互联网也好，信息服务产业也好，将来的瓶颈都在**计算机对语义的理解**。发展语义互联网技术是实现全民应用互联网的重要途径。
- 我们必须挖掘和利用中华文明的特色，研究支持语义、内容和文化的科学技术与普惠信息网络基础平台，使**网上中文信息内容超过全世界网上信息总量的10%**，为发展中国特色的数据知识产业提供科技基础。

8、生物计算与社会计算

- 从计算的角度为细胞的发展过程建立模型，不仅有助于理解生物系统中的大量的基因和蛋白质如何协调工作控制细胞的新陈代谢及DNA修复等基本问题，而且对于**通信协议设计、并行计算模型**和机制的研究也具有重要意义。通过对生物分子和DNA等层次生命活动中信息转化过程的分析，可能**产生与基于硅的电子计算机原理完全不同的计算系统**。
- **社会计算**已成为继科学计算、生物计算之后新的国际前沿研究和应用方向。以认知科学、智能科学和复杂性科学为基础，开展社会计算研究和应用，已成为确保国家安全、建设和谐社会刻不容缓的任务。

9、智能和认知科学技术

- 探索智力的本质，了解人类的大脑和它的认知功能是当代最具挑战性的基础科学问题之一。基于认知机理的智能信息处理在理论与方法上的突破，有可能带动未来信息科学与技术的突破性发展。
- 智能科学是脑科学、认知科学、人工智能等共同研究形成的交叉学科。智能的研究不仅要运用推理，自顶向下，而且要通过学习，由底向上，两者并存。
- 近年来国外在脑科学研究、脑反向工程、神经工程等方面取得不少进展，相对国外而言，我国对脑科学和认知科学的研究还不够重视，今后要加大力度。



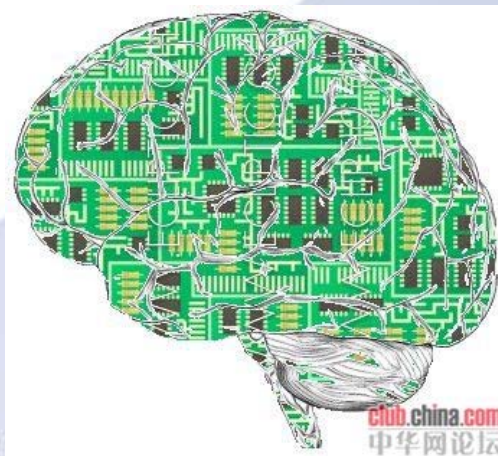
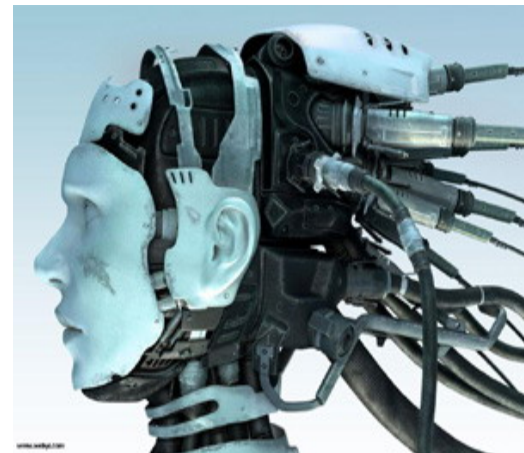
蓝色大脑计划

- 2005年瑞士洛桑高等理工学院（EPFL）启动了一项雄心勃勃的科研计划——**Blue Brain**。该计划使用了IBM的 **Blue Gene** 计算机(每秒22.8万亿次浮点运算)，模拟老鼠的大脑新皮层单元中**10000**个神经元行为。
- 人类大脑中复杂的物理结构以及随时变化的化学成分增加了模拟的难度，模拟记忆“**液态计算**”技术也许会为硅片电路带来新的发展方向。
- IBM以采用蓝色基因超级计算机对大脑皮层II/III 层的**22,000,000**神经细胞和 **110,000,000,000** 突触进行模拟。今年IBM 从DARPA得到**490万**美元的资助，与斯坦福等大学合作研制**类脑计算机**。



基于脑科学成果的片上神经计算机

- 德国海德堡大学正在协调欧盟支持的FACETS项目，该项目汇集了来自7个国家15个研究院所的科学家的，构建一台像大脑一样工作的神经计算机
- 第一步就是在单个芯片上建立一个由300个神经元和50万个突触组成的网络。第二期的新网络中将包含了20万个神经元和5000万个突触，用模拟电路实现神经元、数字电子实现通信。研究小组在20厘米的单硅片上建立了该网络
- 这个系统要比生物等效法快10万倍，比软件模拟快1000万倍。



Thank You!
请批评指正!

